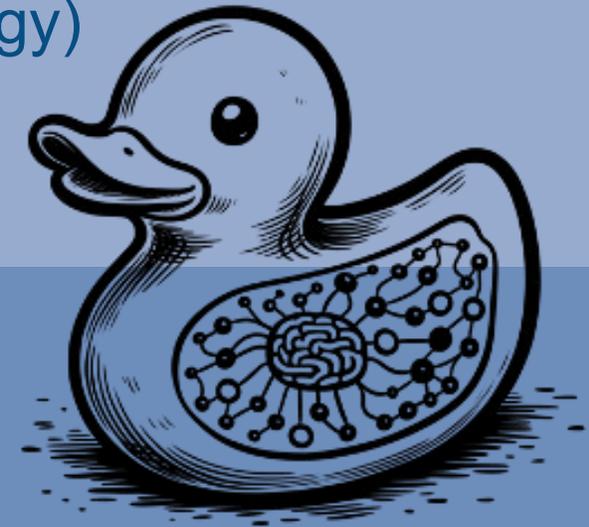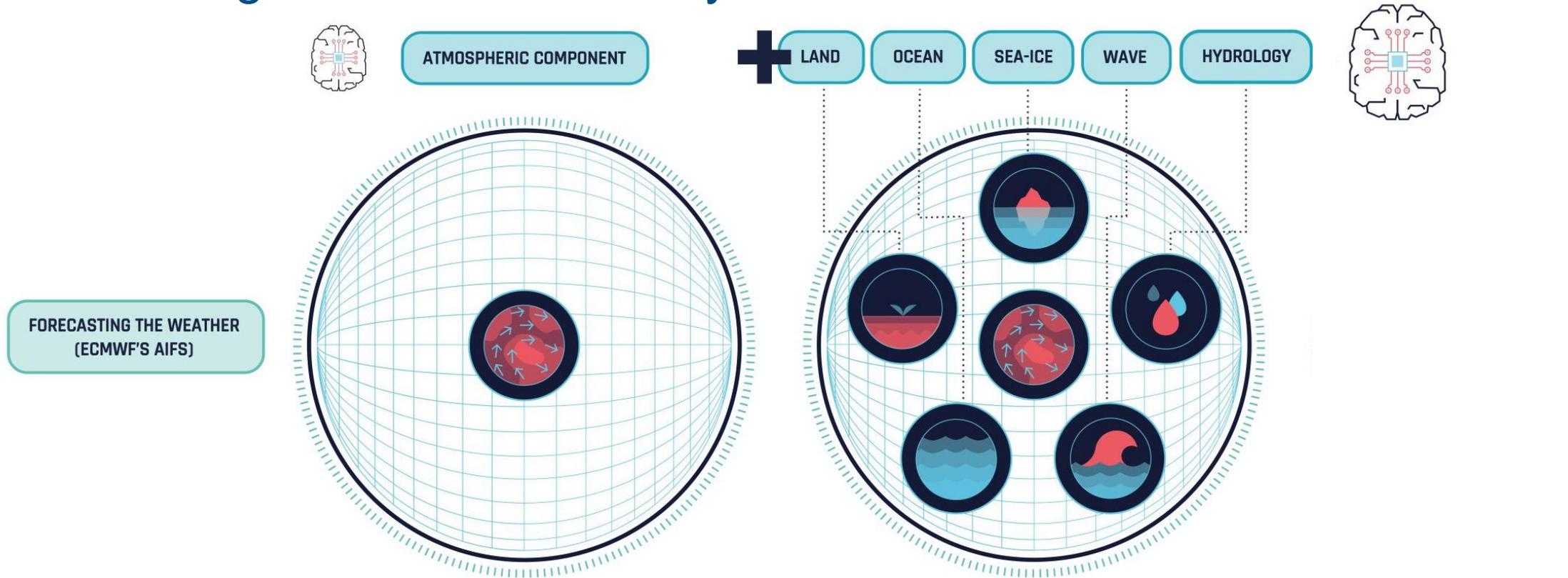# Destination Earth: Earth System Modelling

Ocean, Sea Ice, Land, Waves, (and Hydrology)

Rachel Furner   rachel.furner@ecmwf.int
on behalf of the core ML Earth System Components group:
Rachel Furner, Sara Hahner, Ewan Pinnington, Nina Raoult, Mario Santa Cruz, Maria Luisa Taccari, Kenza Tazi, Lorenzo Zampieri
and many other contributors

ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Advancing towards an Earth system ML Model with DestinE

ATMOSPHERIC COMPONENT + LAND OCEAN SEA-ICE WAVE HYDROLOGY

FORECASTING THE WEATHER (ECMWF'S AIFS)

Maria Luisa   Kenza   Ewan   Nina   Sara   Lorenzo   Rachel   Mario
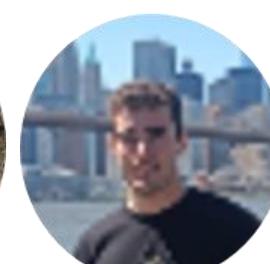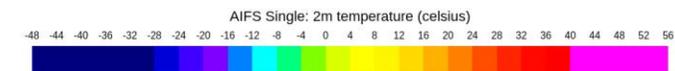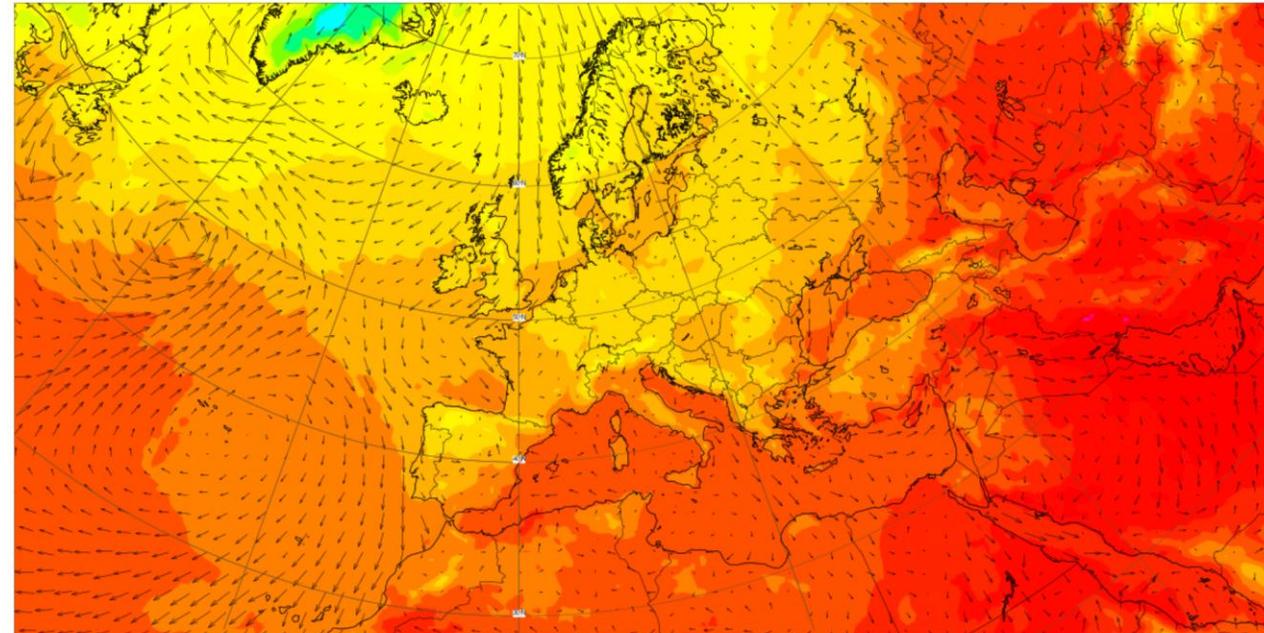
ECMWF
EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Why has ML worked so well for NWP?

- **Data abundance**:
  Decades of high-resolution reanalyses and observational products.

- **Predictability & formulation**:
  NWP to be posed as an initial-value, sequence-to-sequence learning problem.

- **Bounded horizon**:
  Forecasts usually cover days to ~2 weeks, limiting error accumulation.

- **Community Momentum:**
  Broad engagement and open-sourced toolkits

AIFS Single: 2 m temperature and 10 m wind

Base time: Tue 19 Aug 2025 06 UTC Valid time: Tue 19 Aug 2025 06 UTC (+0h) Area : Europe
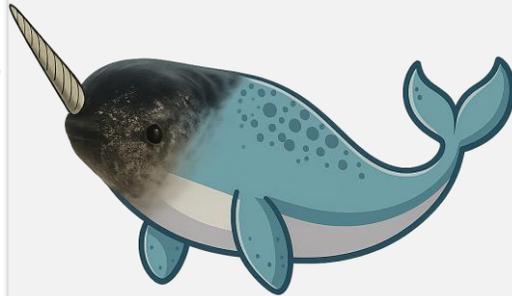

AIFS Single: 2m temperature (celsius)

-48 -44 -40 -36 -32 -28 -24 -20 -16 -12 -8 -4 0 4 8 12 16 20 24 28 32 36 40 44 48 52 56

AIFS Single: 10m wind (m/s)

→ Black wind arrows

Datasets

Common challenges
- Data quality & consistency
- Limited validation data
- Capturing extremes
- Scale mismatch
- Incomplete observability

Atmosphere
ERA5 reanalysis product
¼°, 6 hourly, resolution
1979-2025
~ 7 TB

Ocean and sea-ice
ORAS6 reanalysis product
¼°, 6 hourly, resolution.
2005-2025
~3 TB

Waves
Hindcast of wave model
with altimeter assimilation
¼° resolution.
1979-2025
~700 GB

Land
ERA5 reanalysis product
¼°, 6 hourly, resolution.
1979-2025
~1 TB

Hydrology
CARAVAN (global
large-sample hydrology)
(observed streamflow +
ERA5/IFS)
Catchment scale (area-
averaged)
1980-2024
~50 GB

Temporal scales & Memory

- Many land/ocean processes have **long memory** (groundwater, soil carbon, deep ocean).

- Large **variability between timescales** of different variables

- **Slowly evolving or sub-surface variables** are hard to capture.

- **Climate-scale processes** are limited by sparse observations.

- **Solutions:**
    - **Tendency scaling** for slow variables
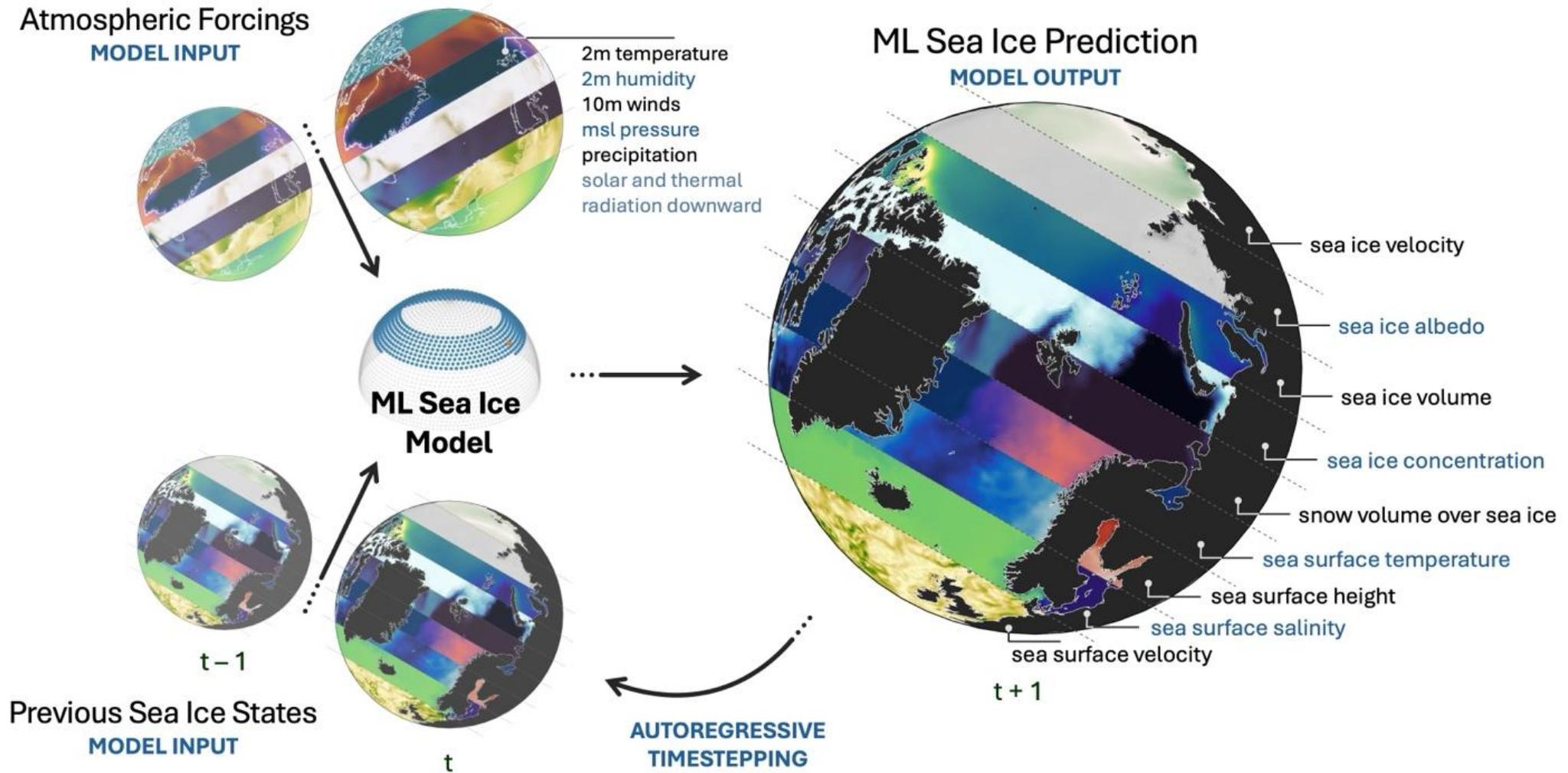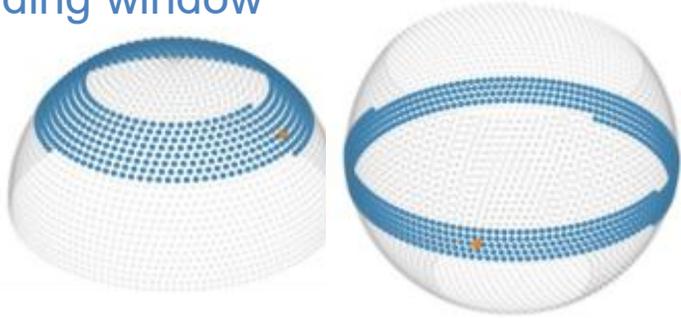   - **LSTMs / memory-aware ML** for temporal dependencies....
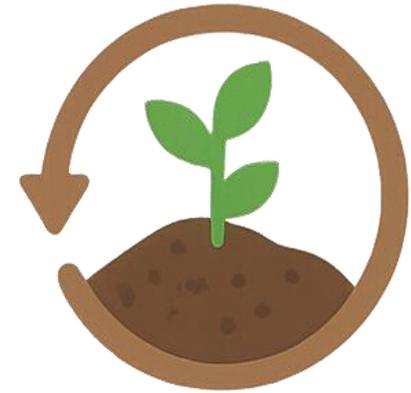
Soil carbon cycle
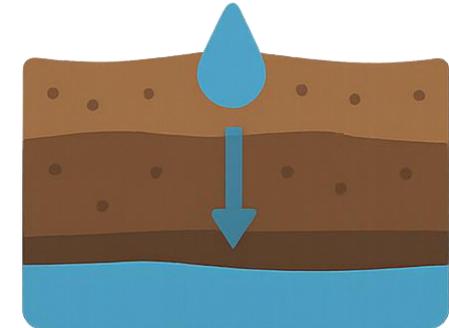
Groundwater recharge
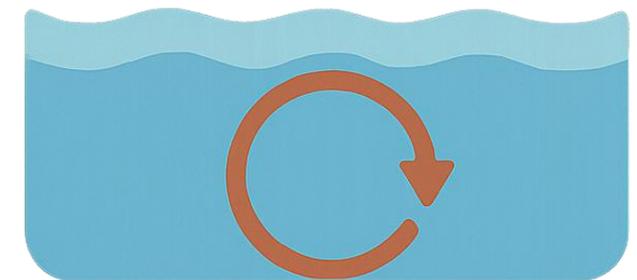
Deep ocean circulation

ML model design with dynamic forcings
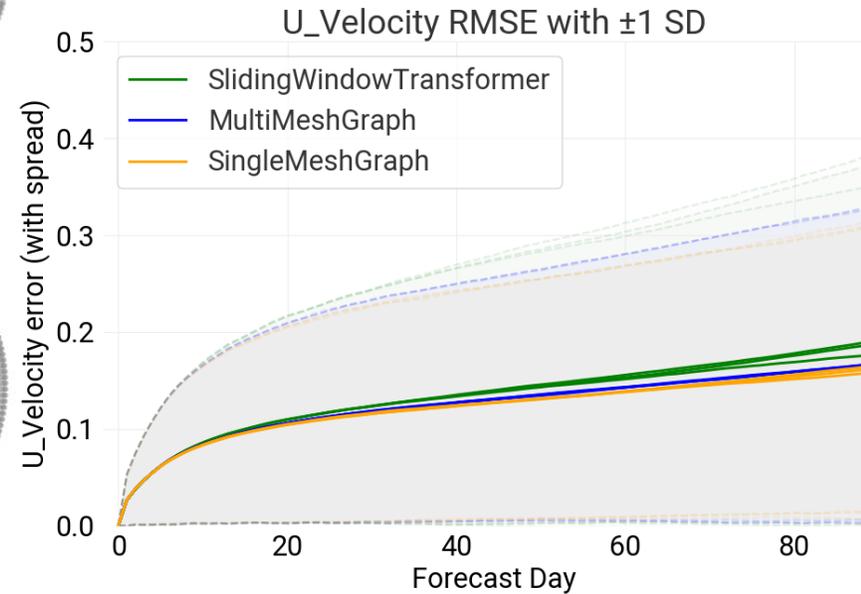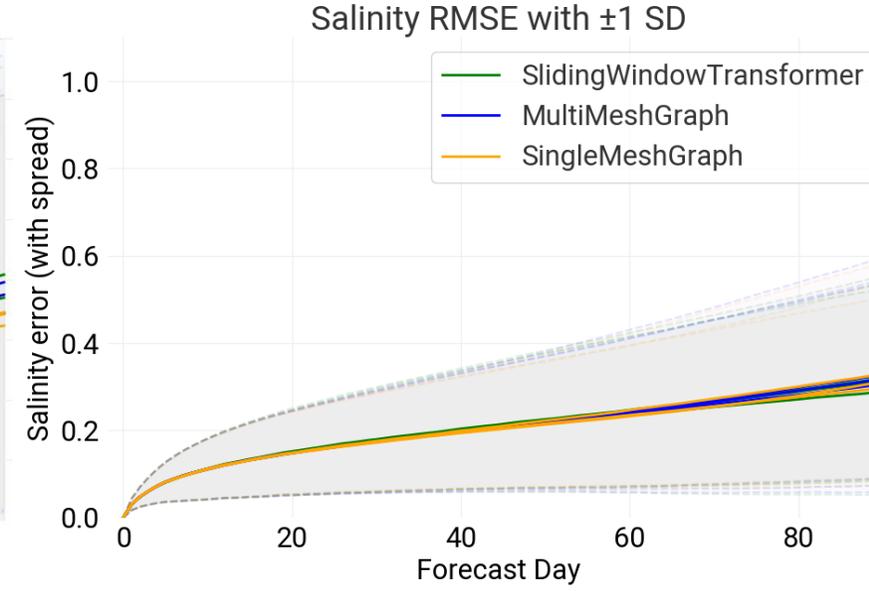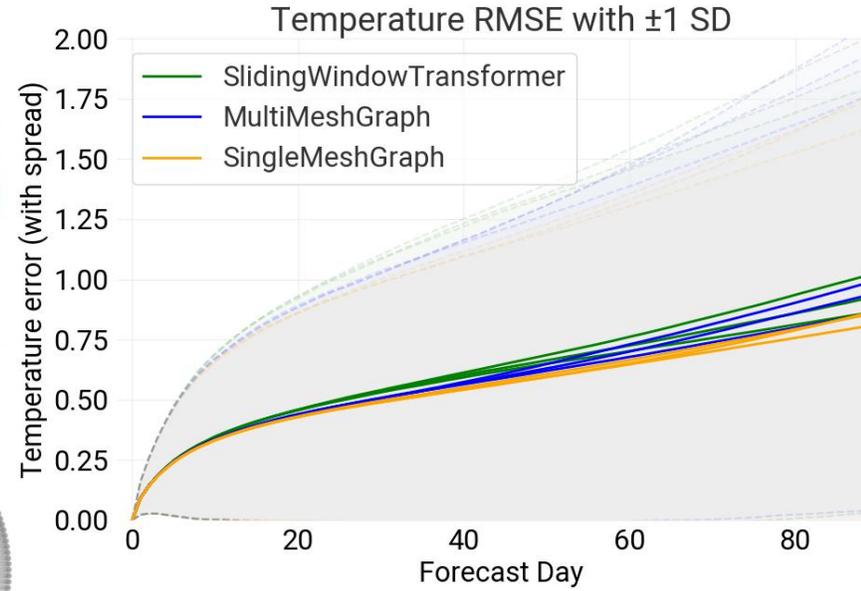
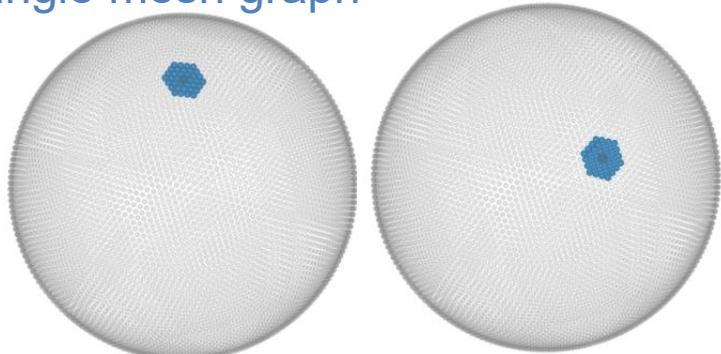# Spatial influences and architecture choice – The ocean
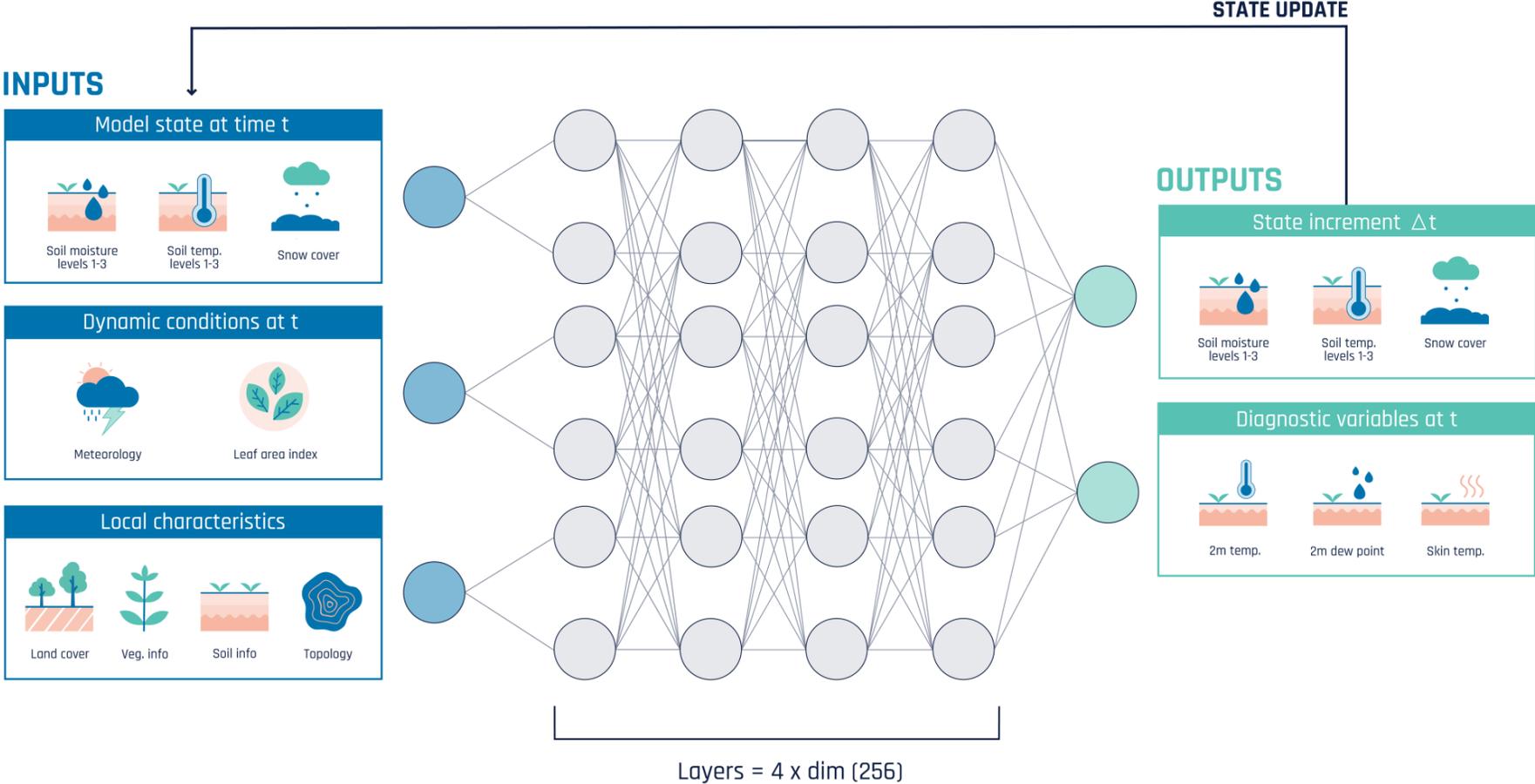
Sliding window

Multi-mesh graph
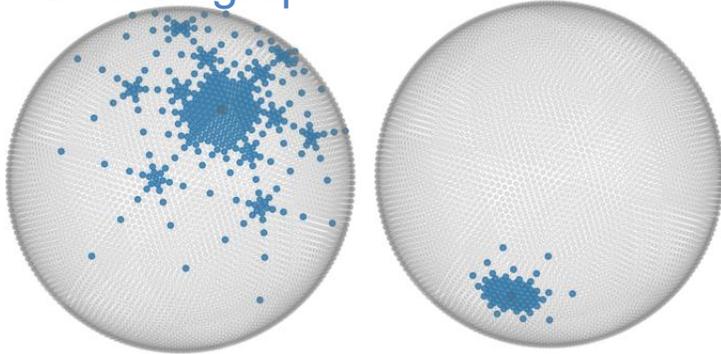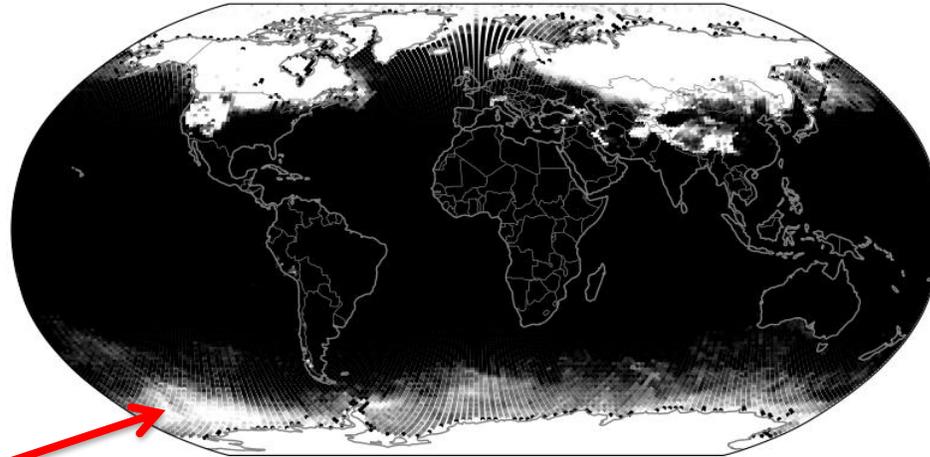
Single mesh graph
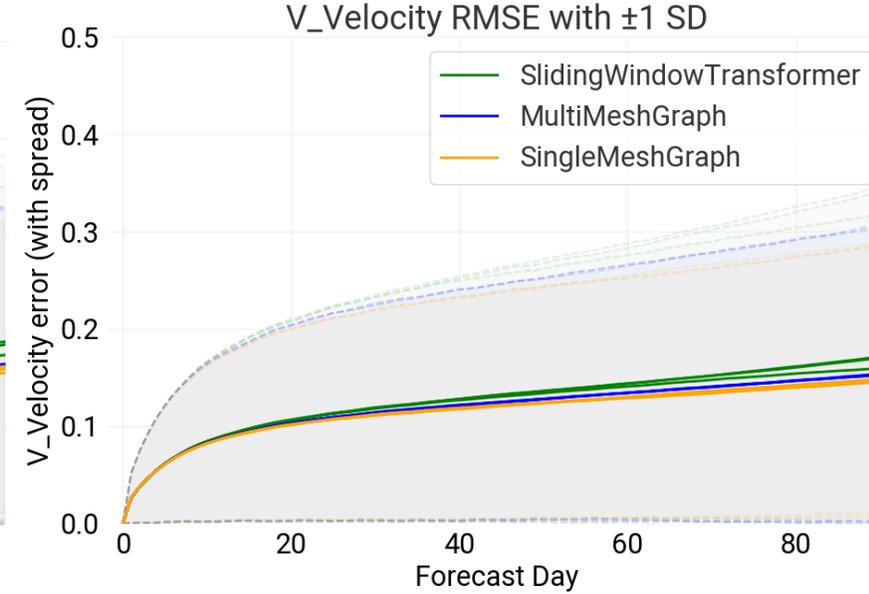
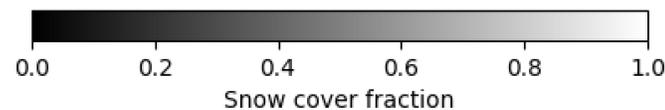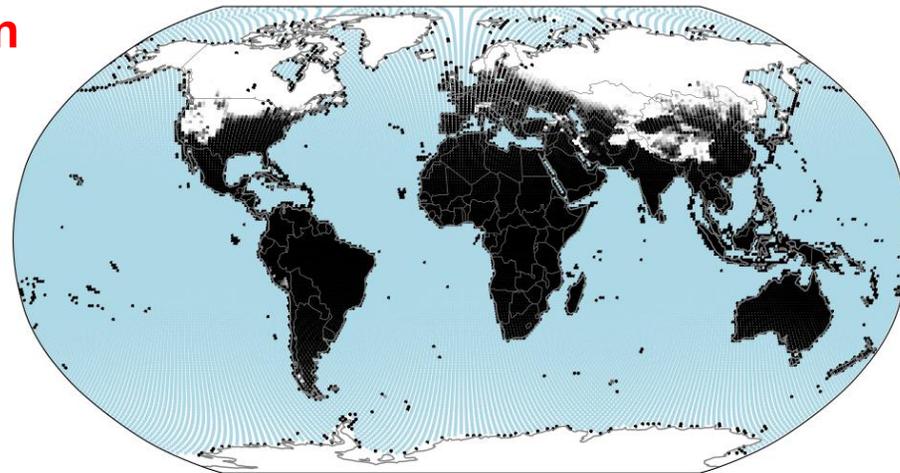# Architecture, and dynamic forcings – land

Dealing with missing values

Without masking NaNs

Snow cover over the ocean

Masking NaNs (lightblue)

Snow cover fraction

- In training: NaNs are replaced by values and **masked in the loss function with zeros**.

- Nans can replaced with constant values (i.e. zero), or statistics (i.e. the variable mean)

- At inference: the model **'predicts'** values **in the masked areas** – these are **removed** in a **post-processing** step.

- Ability to handle **moving NaNs** during training and inference – relevant for ML wave modelling.
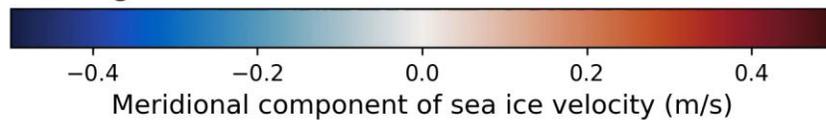
Ensuring physical consistency

Meridional component of sea ice velocity
2022-01-03 00:00
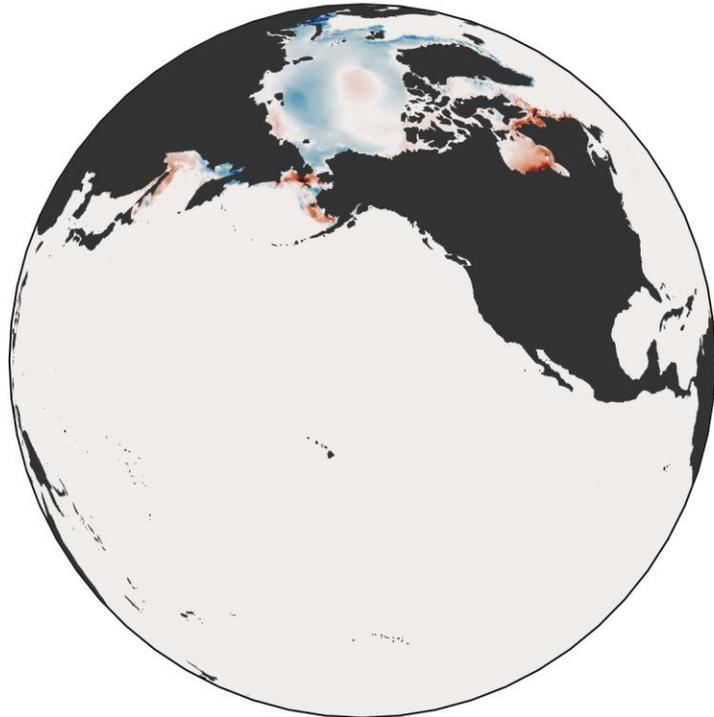
Without sea ice concentration conditioning
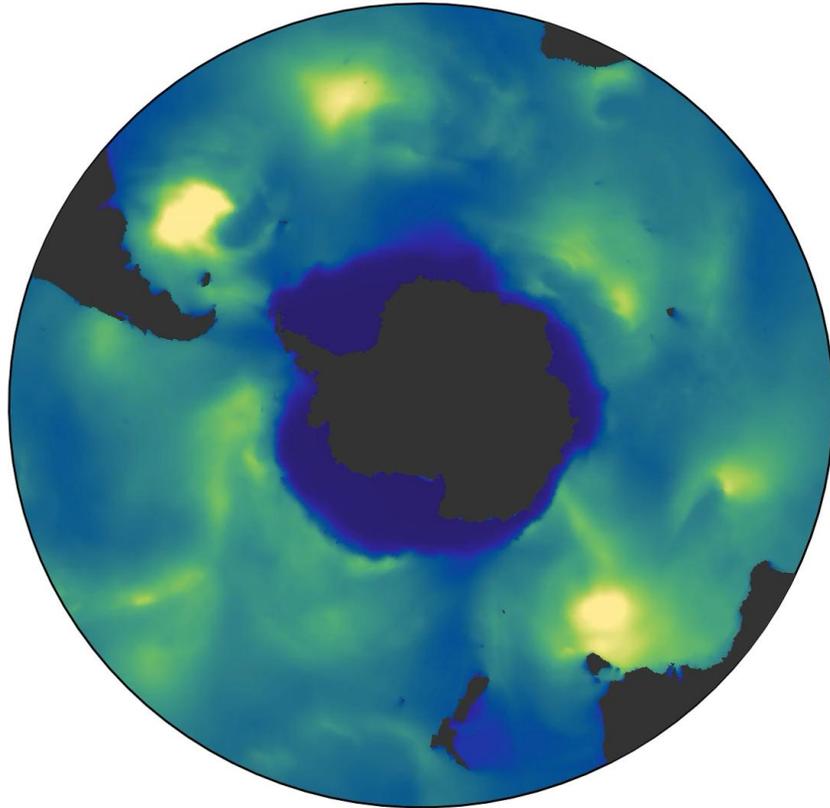
With sea ice concentration conditioning

Meridional component of sea ice velocity (m/s)
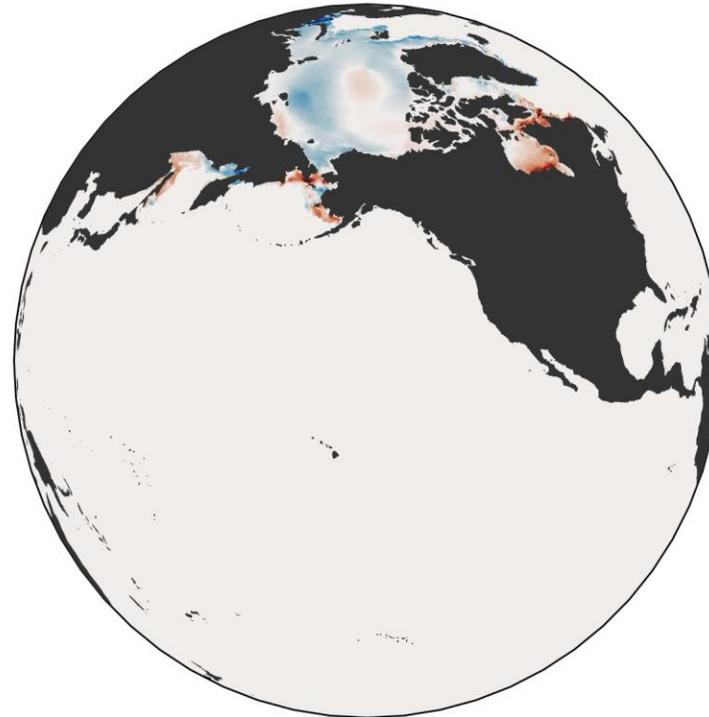
In fact, a hierarchy of sea ice tracers is nothing new but rather used and enforced in all numerical sea ice models

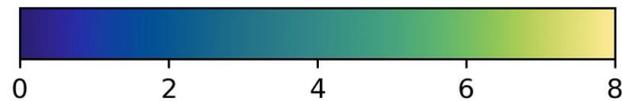e.g., **melt pond tracer dependency in Icepack**
(CICE Consortium)

Significant wave height
## 2023-06-10 00:00

Data-driven model
AIFS

Significant wave height (m)

Implicit modelling:
Sea ice representation through waves

- Under ice, the significant wave height is dramatically reduced

- The ML model learns an implicit representation of sea-ice – the region of reduced SWH realistically varies over time

- We see this *implicit representation* throughout Earth system models

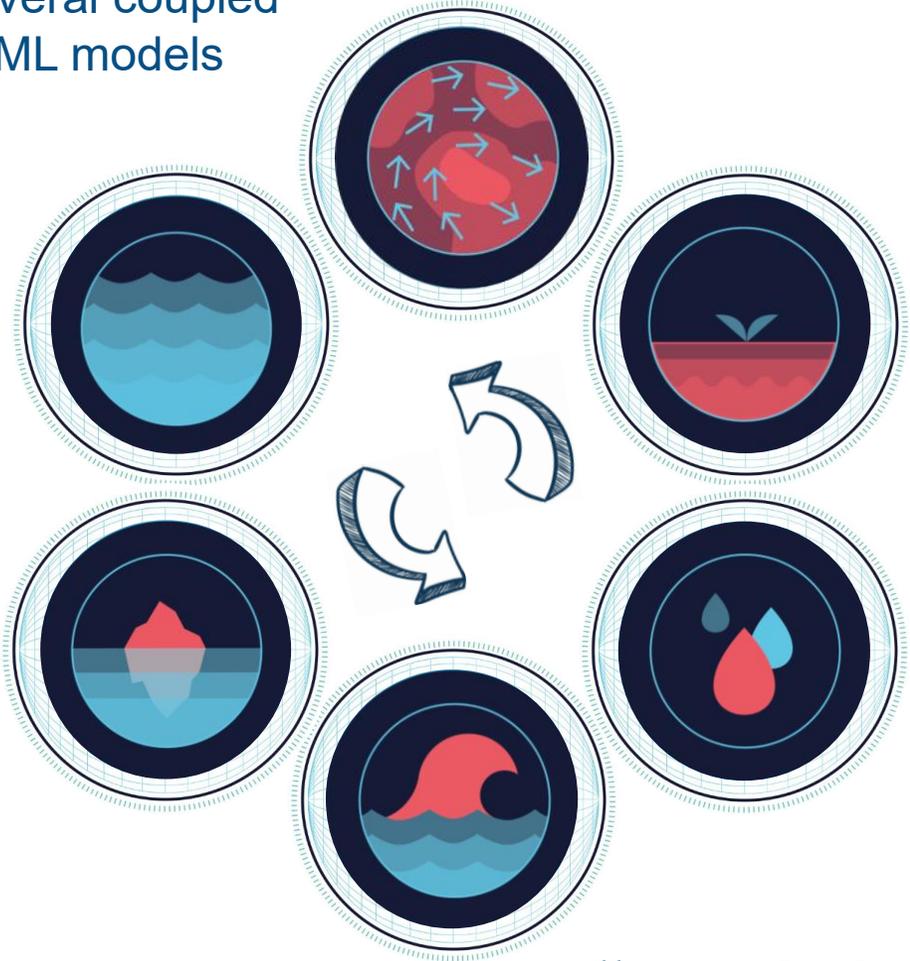- The models may not be coupled, but the datasets are

Investigating approaches for multi-component ML models
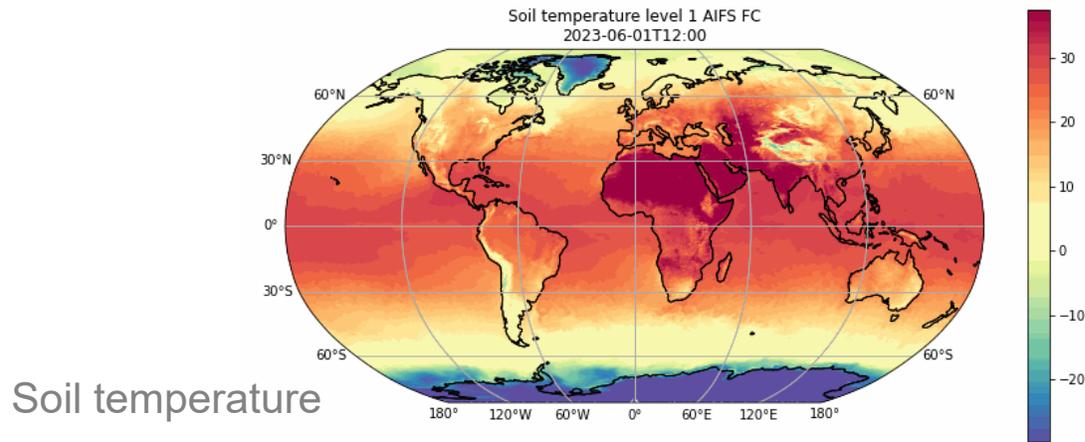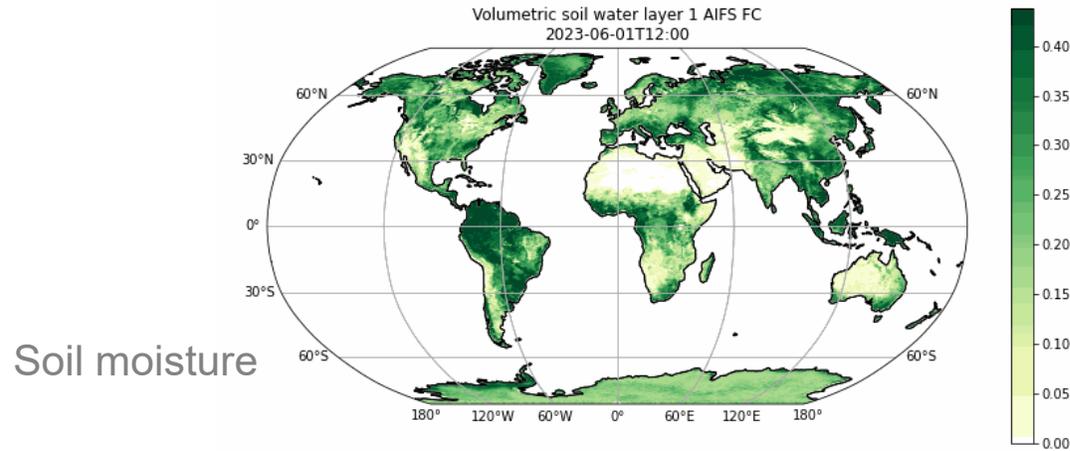
One jointly
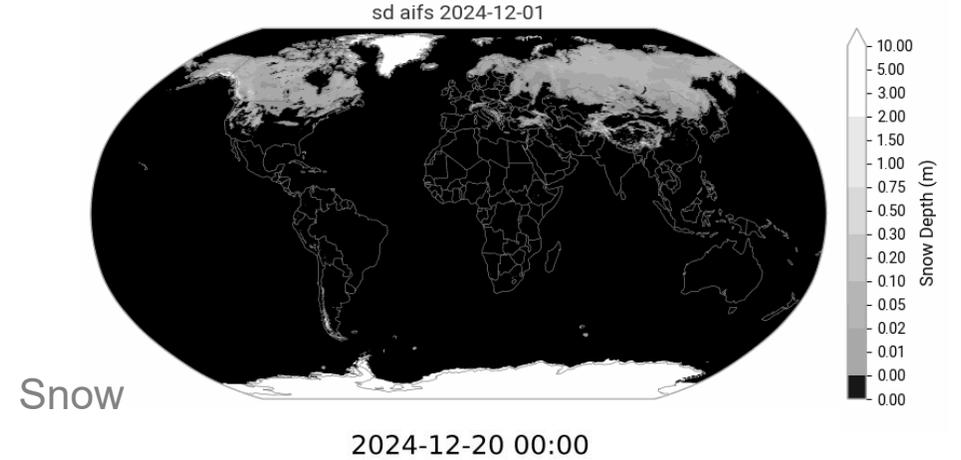learnt ML model
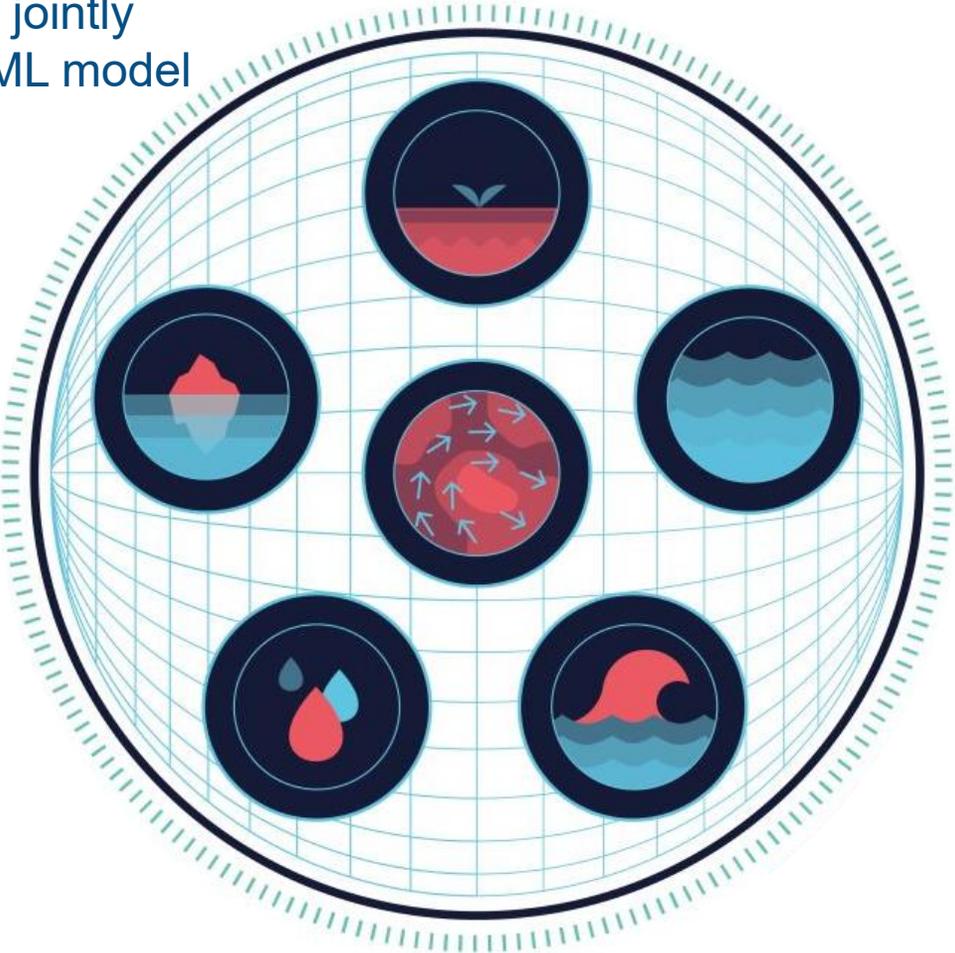
"joint model"

Several coupled
ML models

"coupled model"

The AIFS is already a joint model

AIFS Single v1:

AIFS single v2:

Soil moisture

Soil temperature

Snow

Waves

How do we build a ML coupled model? ATMOSPHERE + OCEAN

Our Goal:
Building an ML modelling system that mimics the coupling
strategy of traditional numerical models

Coupled fields from **atmosphere** to **ocean**:

10u	**10v**	**2t**	**2d**
ssrd	**strd**	**tp**	**msl**

Coupled fields from **ocean** to **atmosphere**:

sst
siconc

Well defined coupling interface
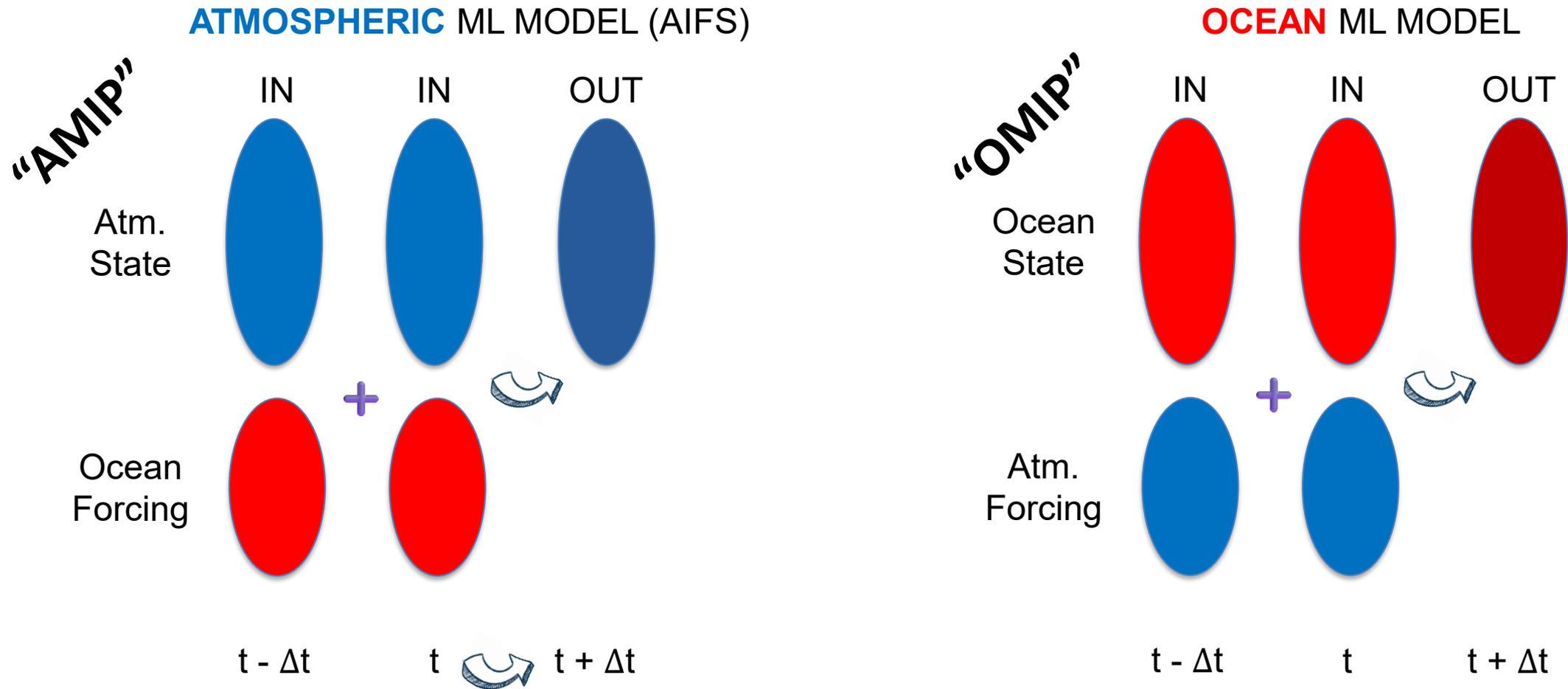
Using compatible reanalyses
(ERA5 and ORAS6)

State variables and not
fluxes are coupled

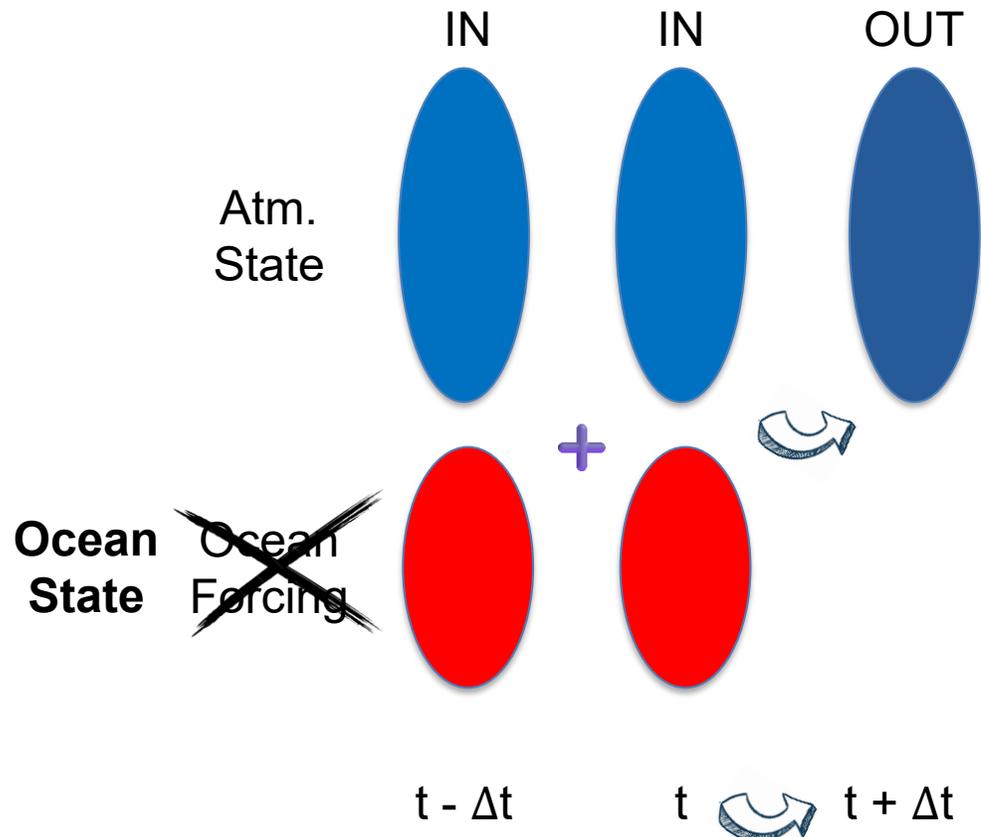The bulk formulation is
learned from data

How do we build a ML coupled model? ATMOSPHERE + OCEAN

During training
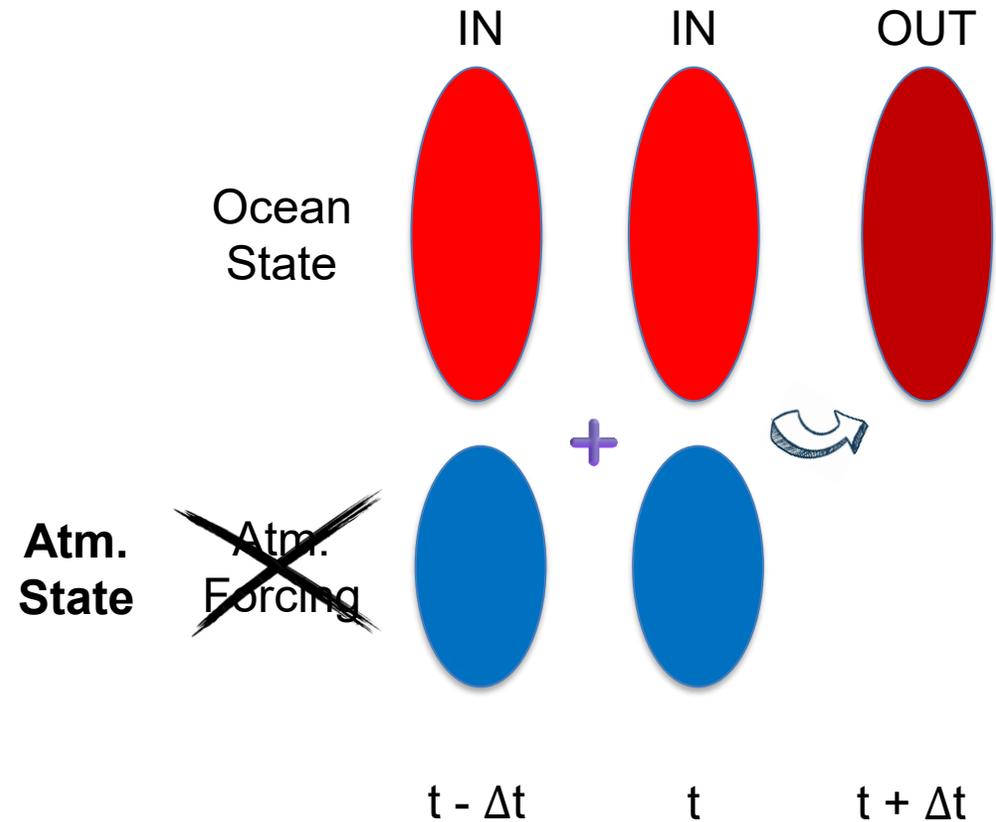
ATMOSPHERIC ML MODEL (AIFS)

OCEAN ML MODEL

No joint finetuning training possible at this stage
with our software (anemoi)

How do we build a ML coupled model? ATMOSPHERE + OCEAN

During inference

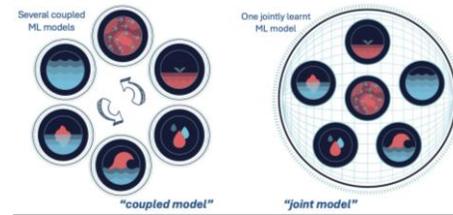

ATMOSPHERIC ML MODEL (AIFS)

IN   IN   OUT

Atm. State

Ocean State ~~Ocean Forcing~~

t - Δt   t   t + Δt

OCEAN ML MODEL

IN   IN   OUT

Ocean State
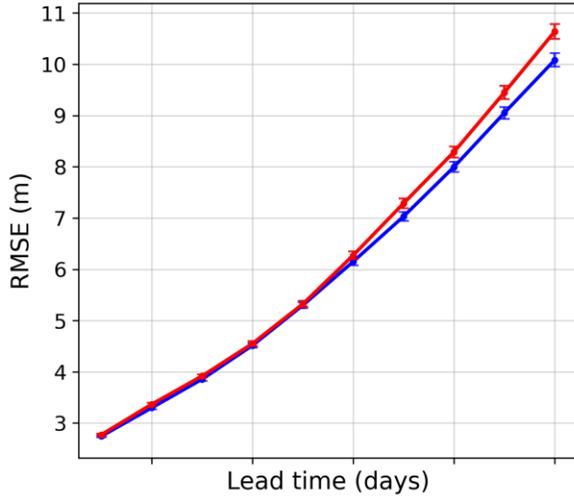
Atm. State ~~Atm. Forcing~~

t - Δt   t   t + Δt

By choice, we do not impose a driving model that leads the coupling
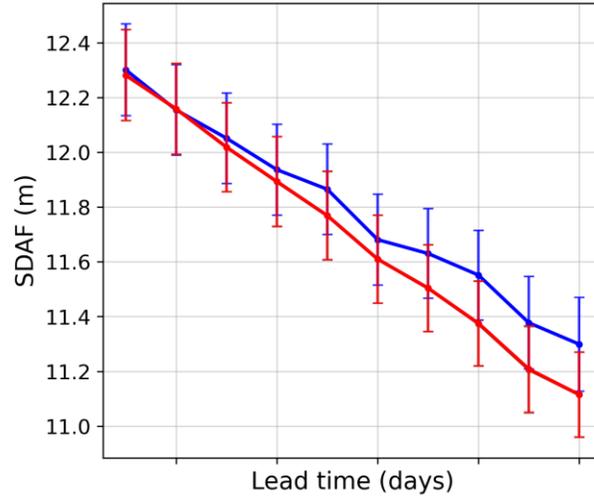
Comparing **coupled** and **joint** model configurations
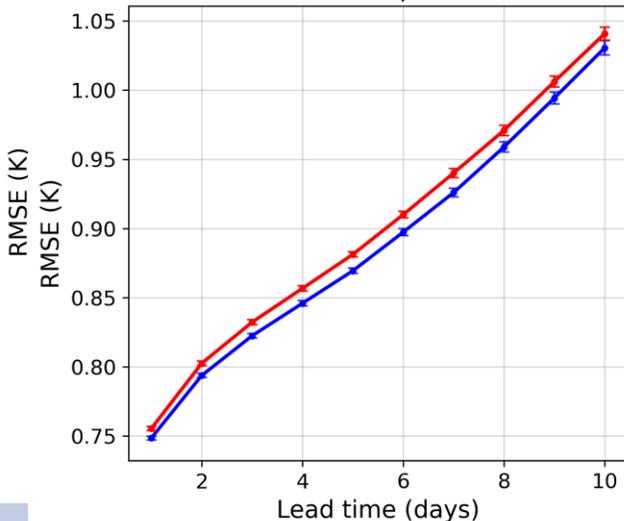
ATMOSPHERIC Perspective

Both approaches are viable and present overall similar behaviors.

2t representation benefits from a joint training, particularly in the tropics, where knowledge about the ocean circulation is more relevant.

Joint model shows slightly better scores and more active forecasts.

Similarity between coupled and forced (AMIP) atmospheric model suggests that the atmosphere uses its intrinsic ocean representation rather than the forcing information.
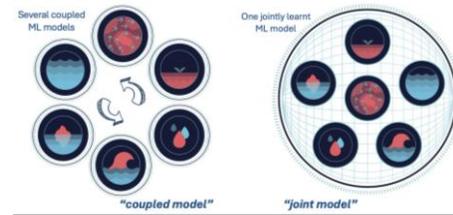
Experiment Configuration :

— Atmosphere (786) + ocean (512), **coupled model** — 174 forecasts
— Atmosphere (786) **forced model** — 174 forecasts
— Atmosphere + ocean (1024), **joint model** — 174 forecasts

Coupling timestep : 6h (also the forecast timestep)

Verification year : 2023

Comparing **coupled** and **joint** model configurations

OCEAN Perspective

Coupled model substantially better approach for ocean fields. Joint training leads to degradation for all variables.

Forced ocean model (OMIP) shows best performances, indicating that at these timescales the atmosphere has a driving role on the surface ocean.

SSH bias develops similarly for all configurations.

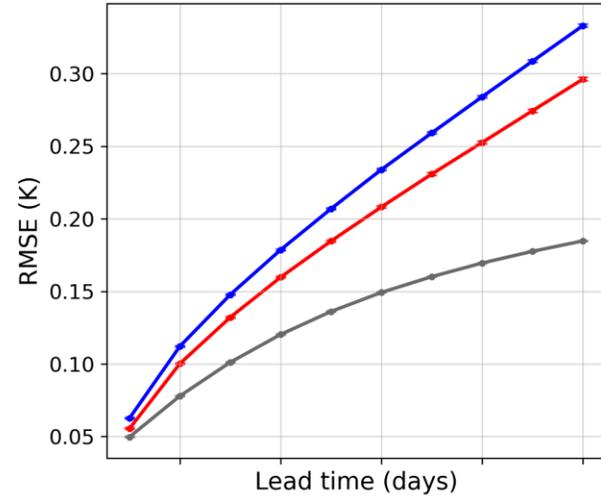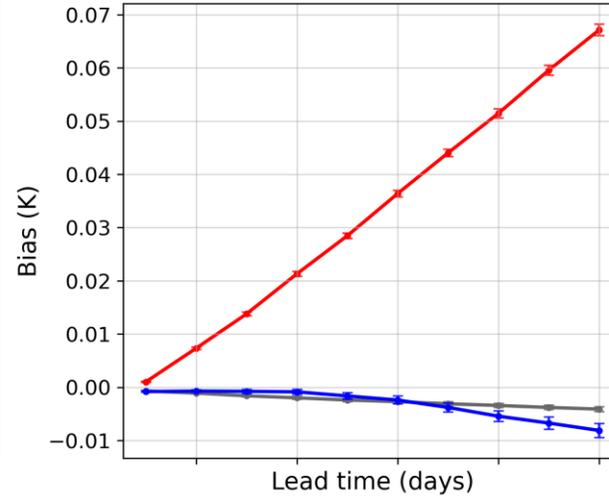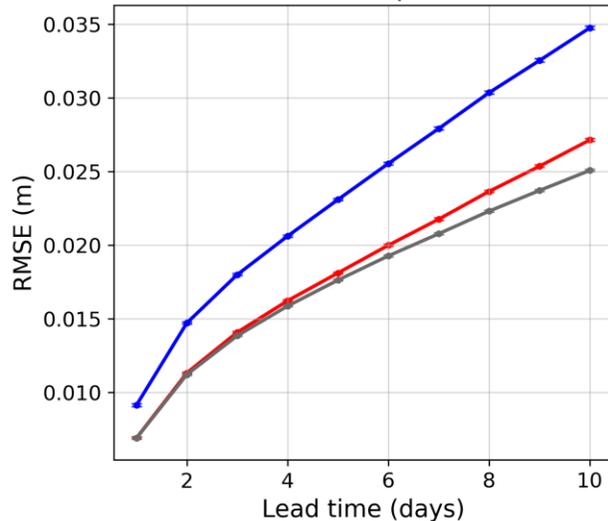Sea ice scores are less affected by differences in coupling strategy.

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
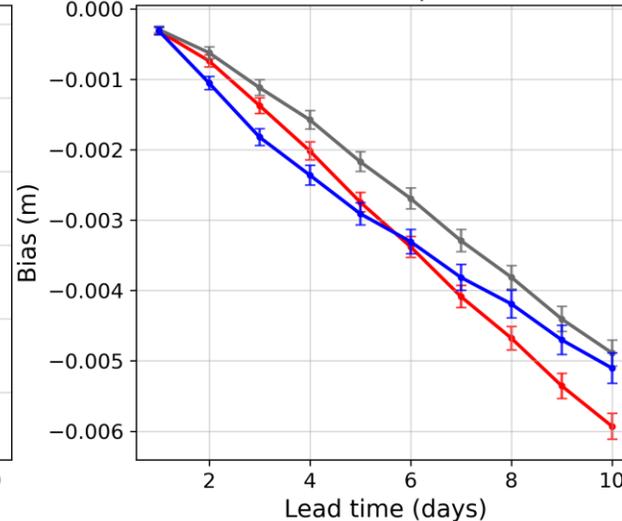
Highlights from ML component models
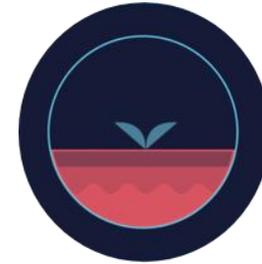
ECMWF — EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
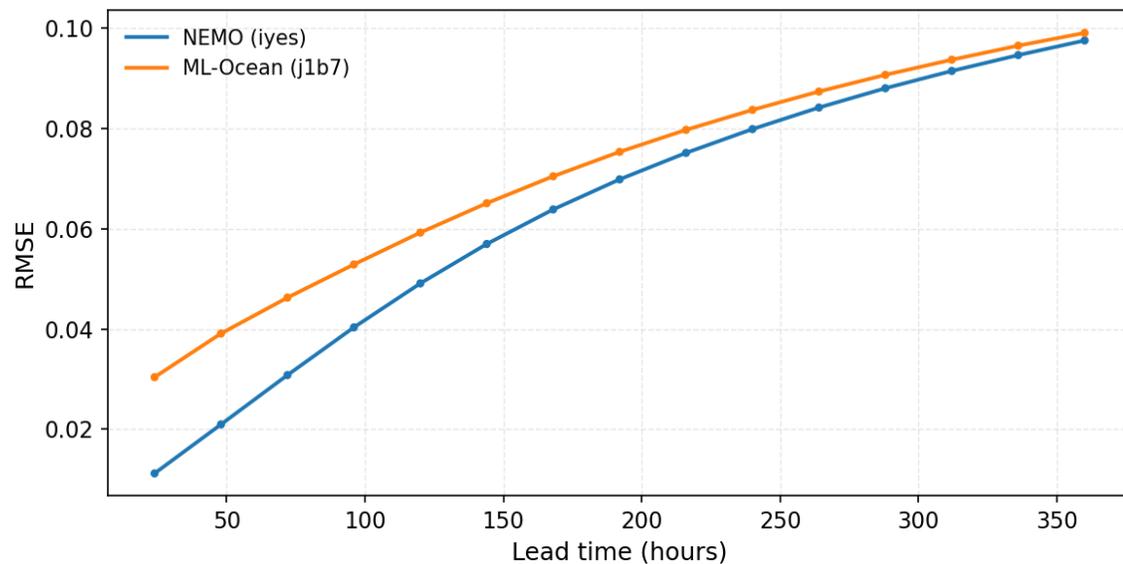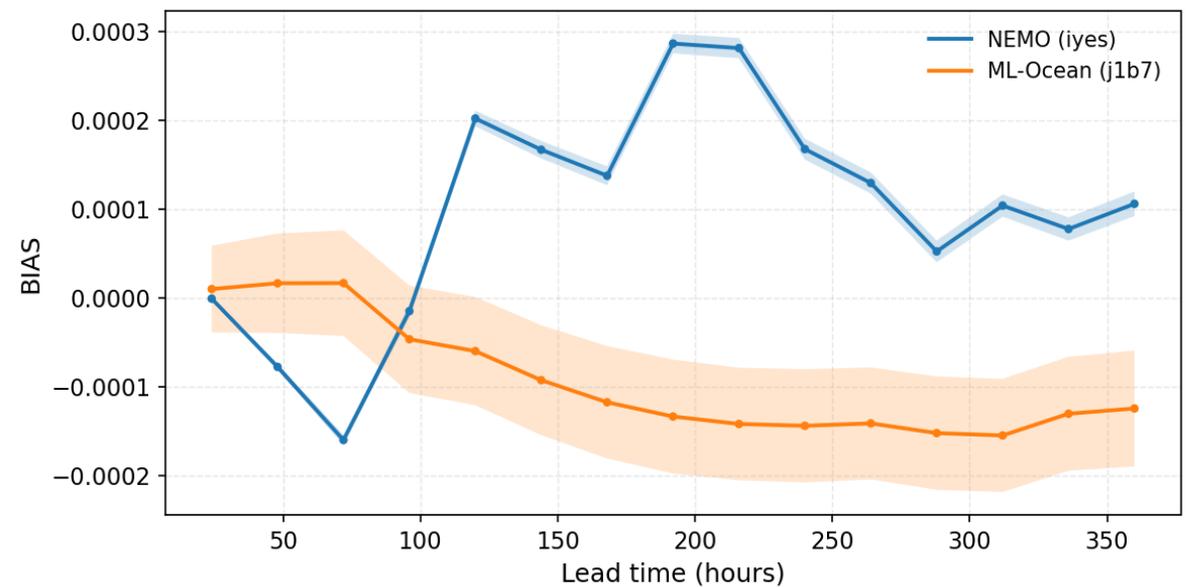
ML-Ocean model comparable to physics based NEMO
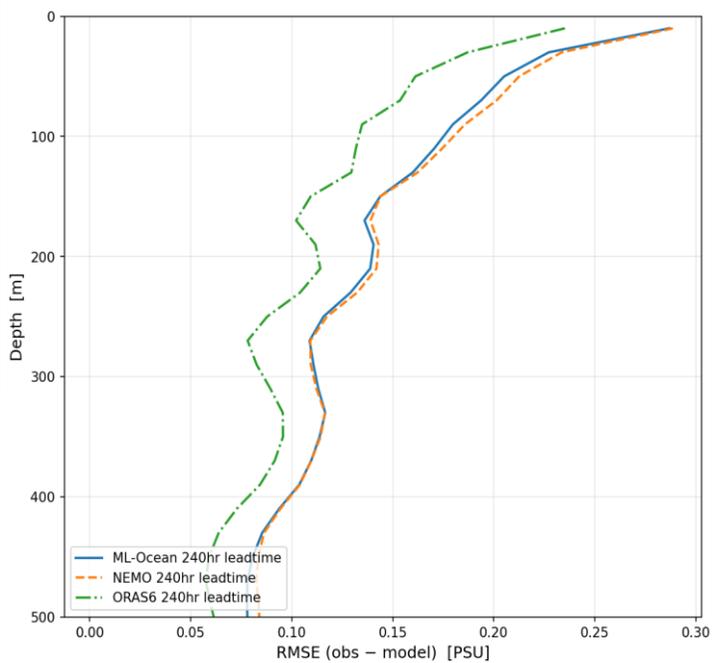
ML-Ocean model comparable to physics based NEMO
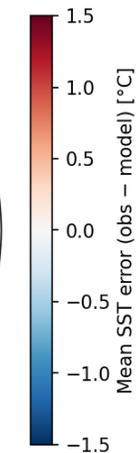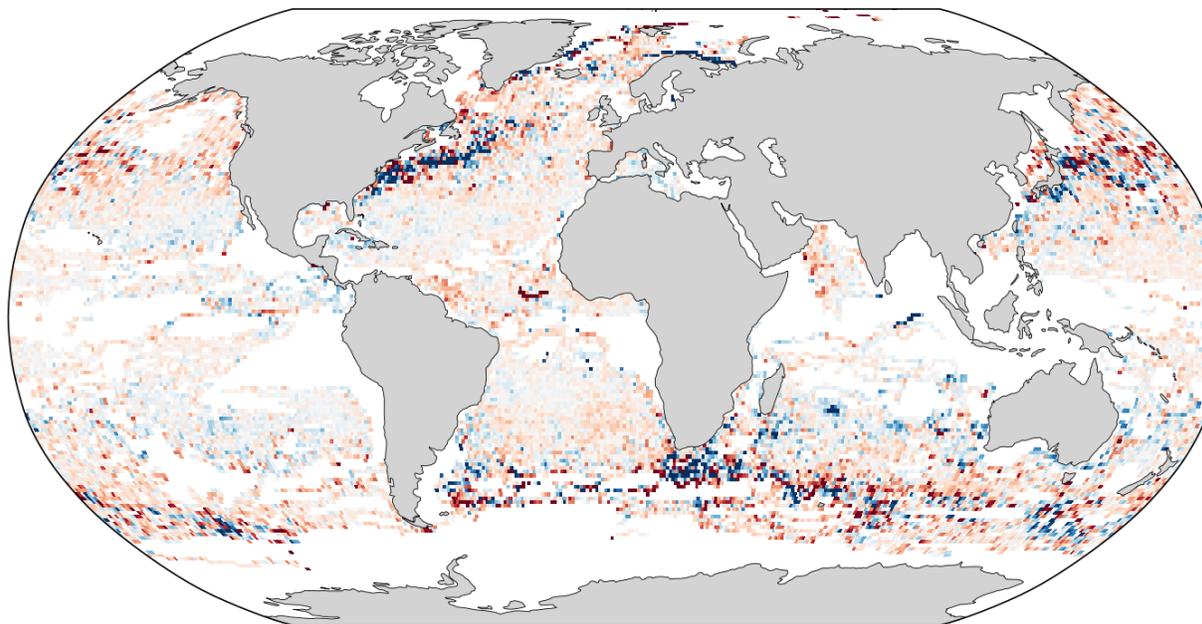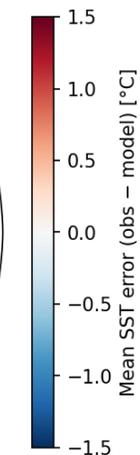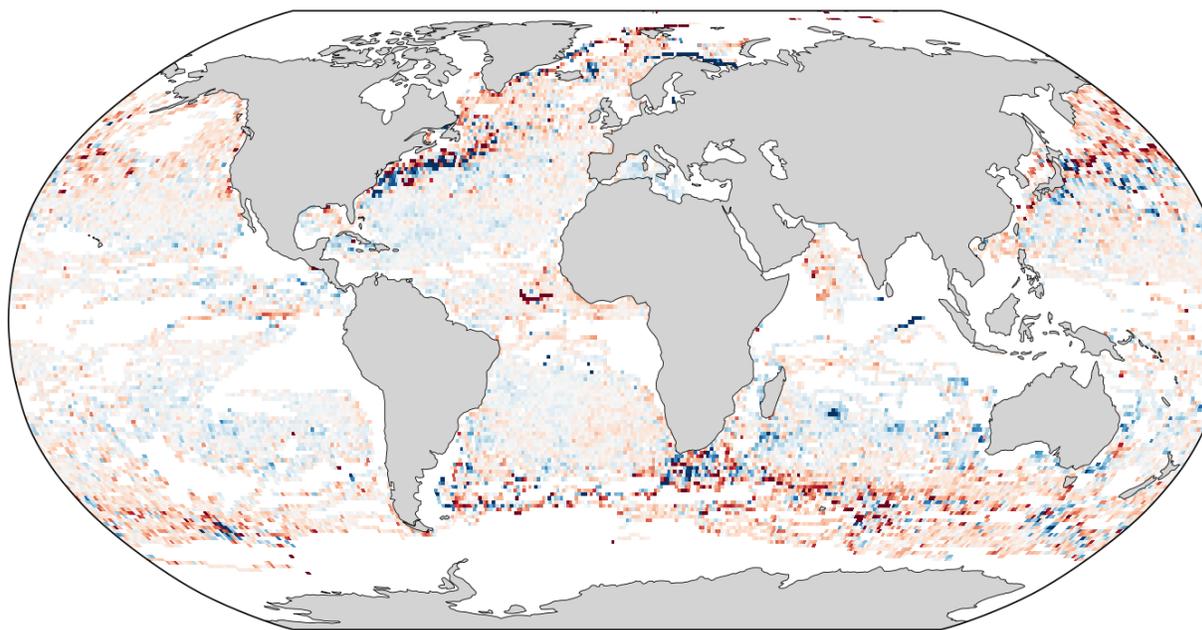
Data-driven wave forecasts reduce error by 10%

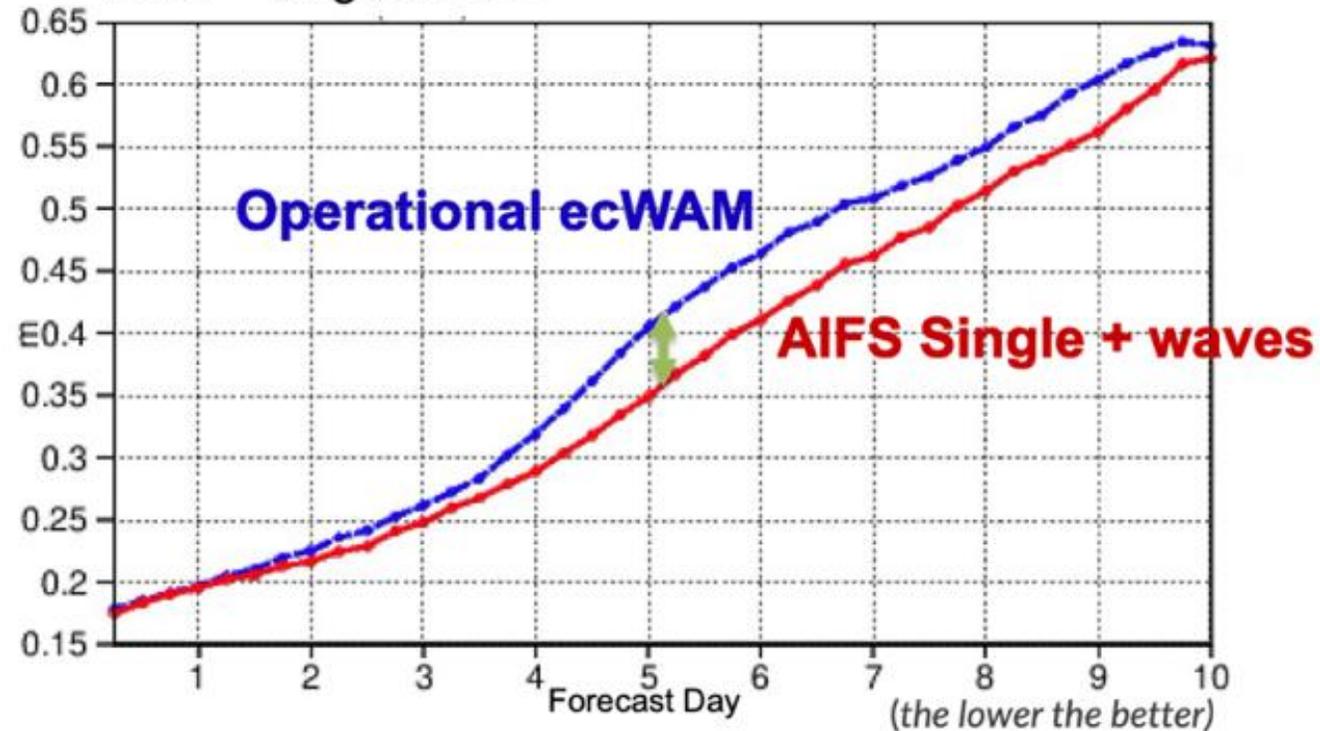Northern Hemisphere Significant Wave Height

Standard deviation of forecast error
June – August 2023

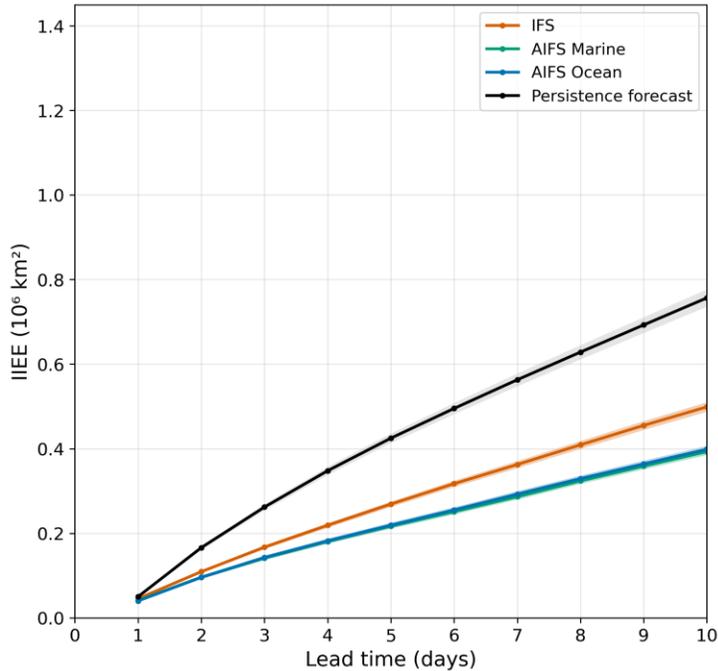Operational ecWAM

AIFS Single + waves

Forecast Day

(the lower the better)

Significant wave height forecast errors are reduced by approximately 10% for medium-range forecasts against operational wave model.
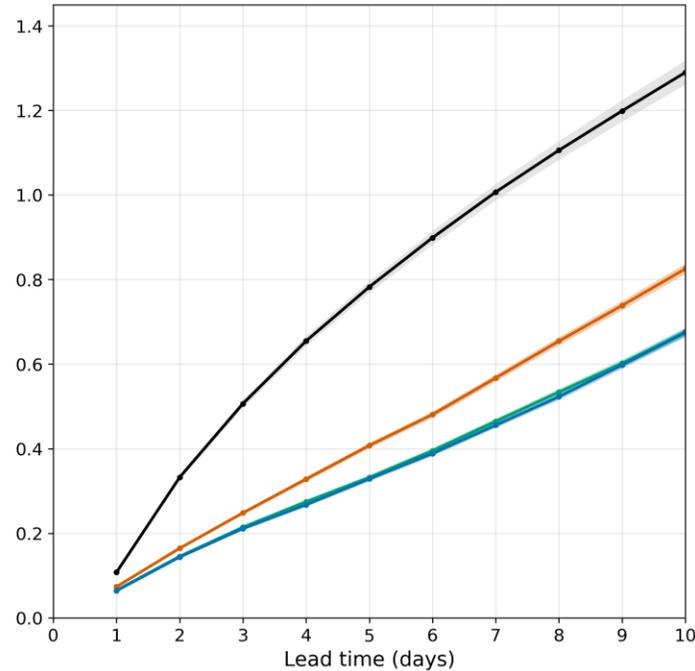
ECMWF EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS

Advancing sea ice edge position predictions at the medium range
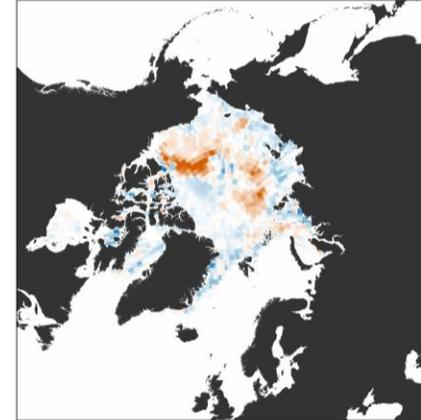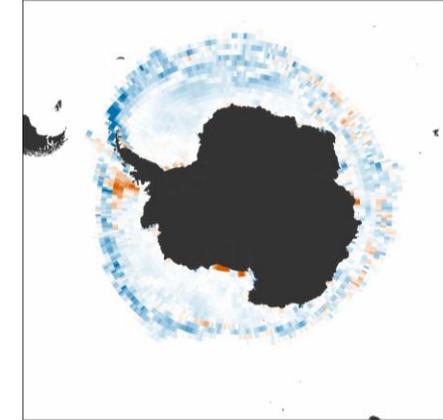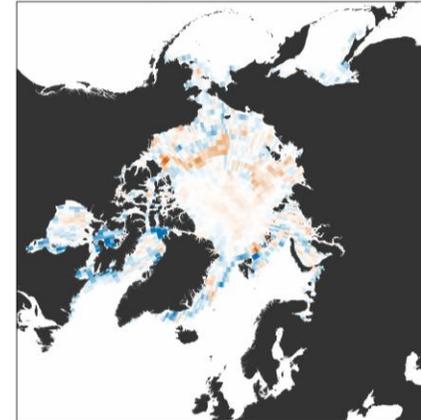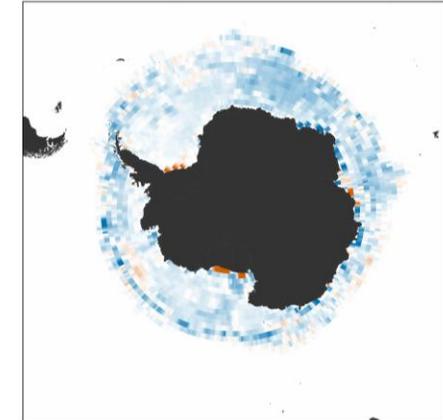
Interaction between sea ice, ocean, and waves:
Jointly learning the coupled climate system

Antarctic Bellingshausen & Weddell Seas · AIFS Marine
Initialized on 2023-08-26 00 UTC

EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS
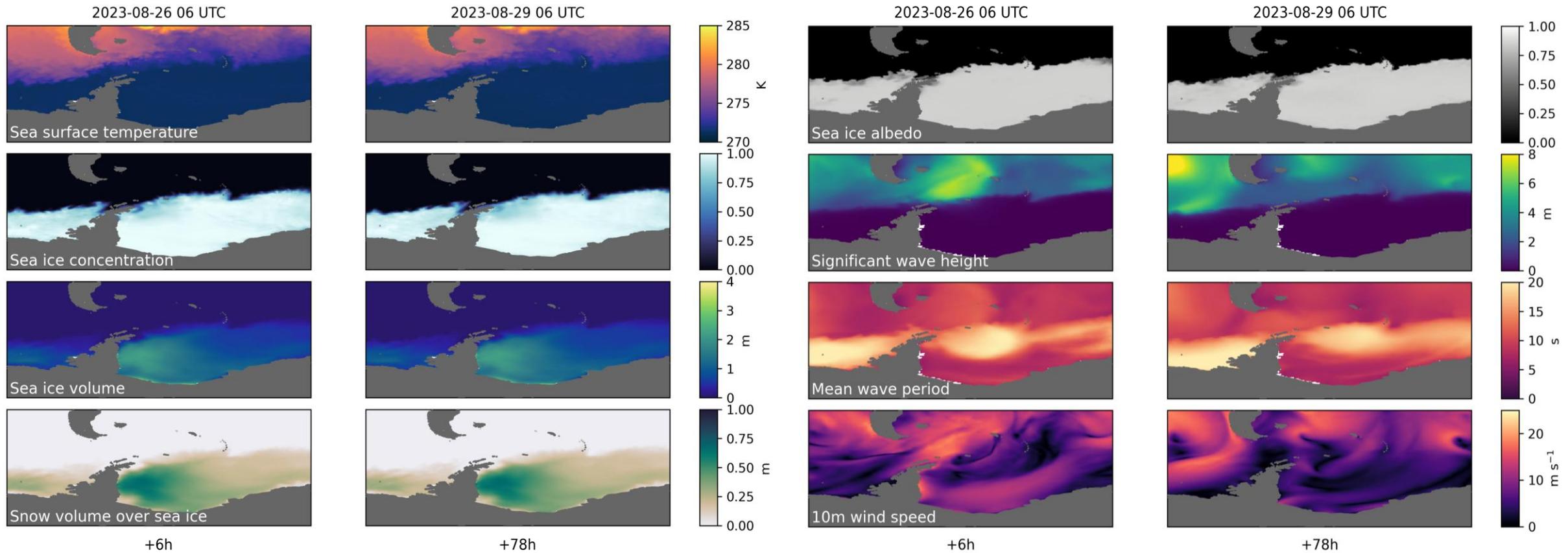
Land emulation and forecasting

aiLand: stand-alone emulator

➢ Emulator of ECMWF's land surface model

Land variables in the AIFS

➢ Snow in the AIFS is less biased than in the IFS

Soil moisture (level 1)
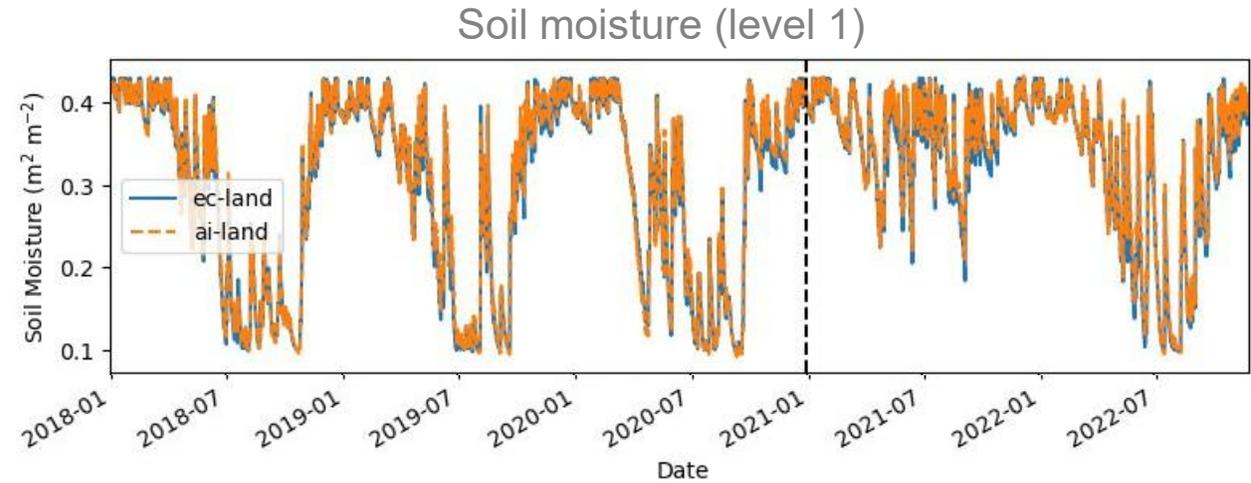
Snow cover "hits" AIFS vs IFS 10-day FC
compared to IFS snow-analysis (2024-12)

1: AIFS predicted snow when IFS & AN did not (**AIFS BAD**, IFS GOOD)

-1: IFS predicted snow when AIFS & AN did not (IFS BAD, **AIFS GOOD**)
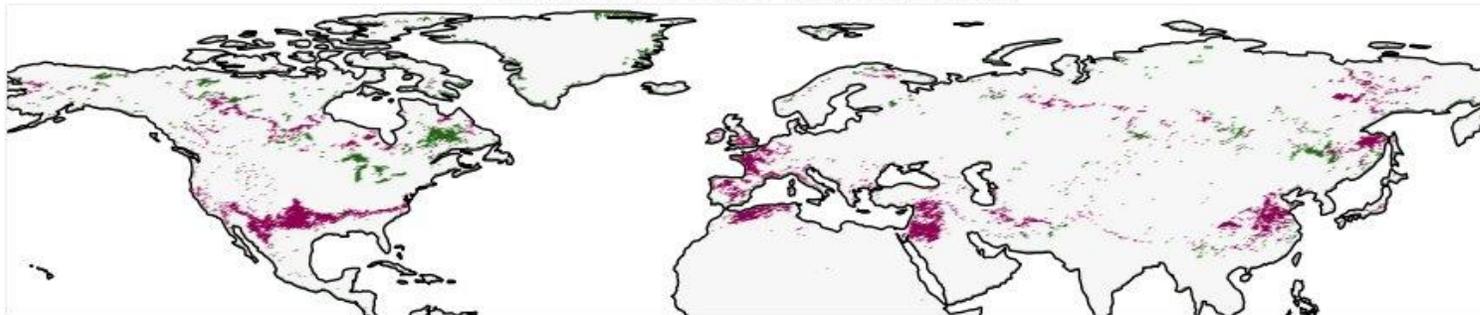
Hydrology
AIFL: Towards ECMWF's data-driven flood forecasting system
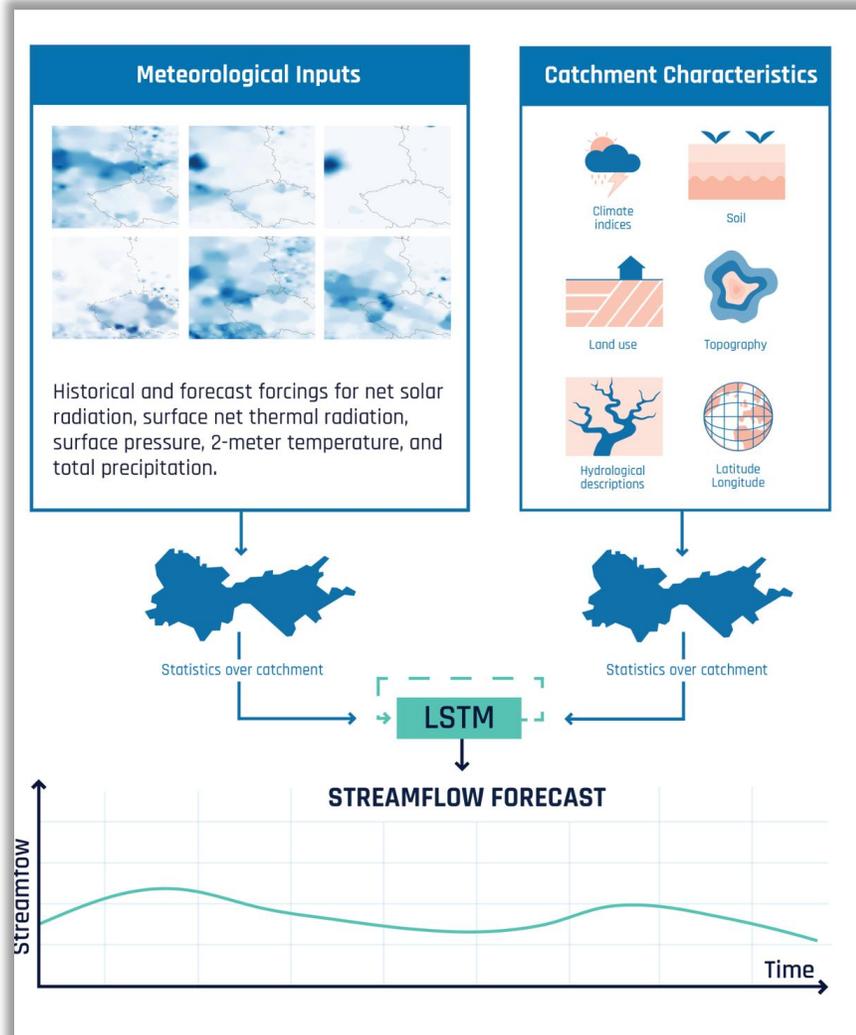
- Flooding is widespread and the most destructive natural disasters worldwide

- River discharge is a key independent diagnostic of hydrological systems

AIFL is a Global LSTM-based flood forecasting trained with up to 20,000 basins

Conclusions

- **Successful prototypes** have been developed across Earth system components.

- **Extending ML success** beyond the atmosphere requires tailored approaches to data, scales, and physical processes.

- **Progress relies on unifying expertise** across domains.

- **Technical advances** are essential to overcome challenges in adapting NWP methods to other components.

> ➤ *many of these development have already fed into **operational** models at ECMWF!!!*

Check out our blog posts!

AIFS Blog:

Representing ocean wind waves in ECMWF's AIFS

Let it snow! How machine learning will forecast snow in the AIFS

DestinE Blog – Towards an ML-based Earth System Model:

Land

Waves

Hydrology