



Norwegian  
Meteorological  
Institute

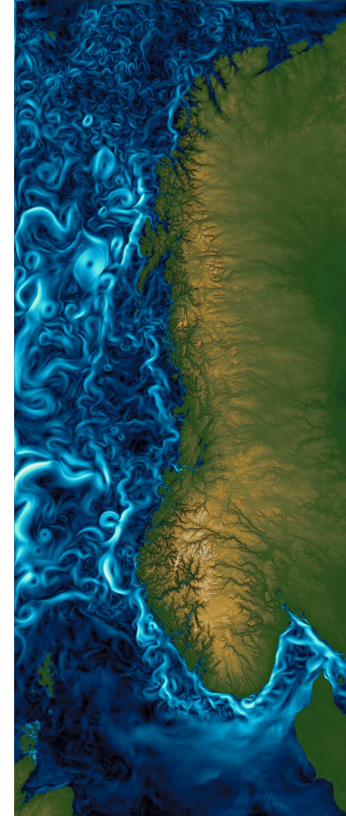


# Developing data-driven ocean models for the Norwegian coast and fjords using graph neural networks

Ina K. B. Kullmann, Mateusz Matuszak, Johannes Röhrs, Thomas Nils Nipen,  
Ivar Ambjørn Seierstad, Kai H. Christensen, Ann Kristin Sperrevik

18.06.2025

# How can we provide better ocean forecasts along the Norwegian coast and fjords?



# How can we provide better ocean forecasts along the Norwegian coast and fjords?

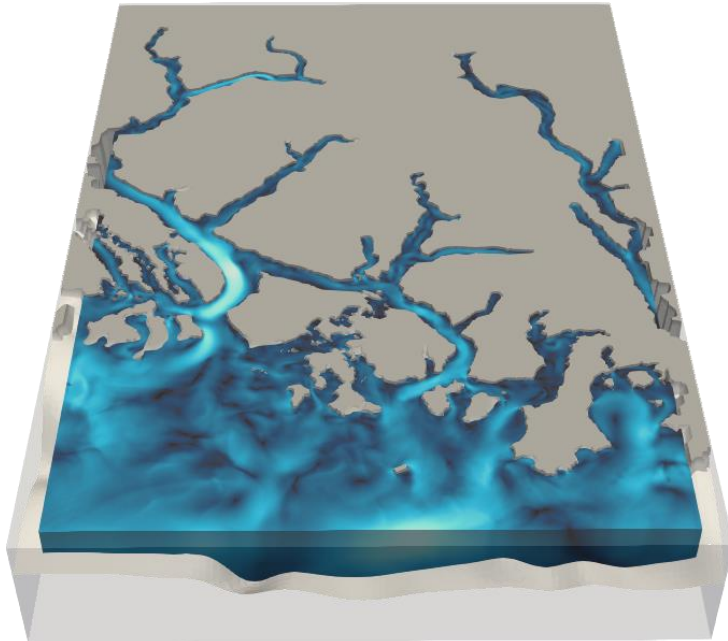
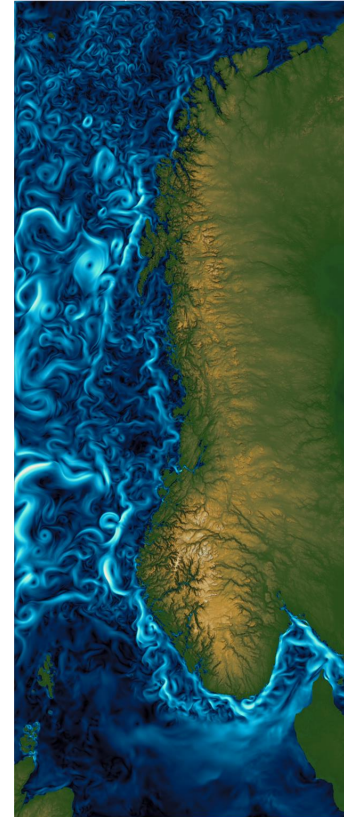
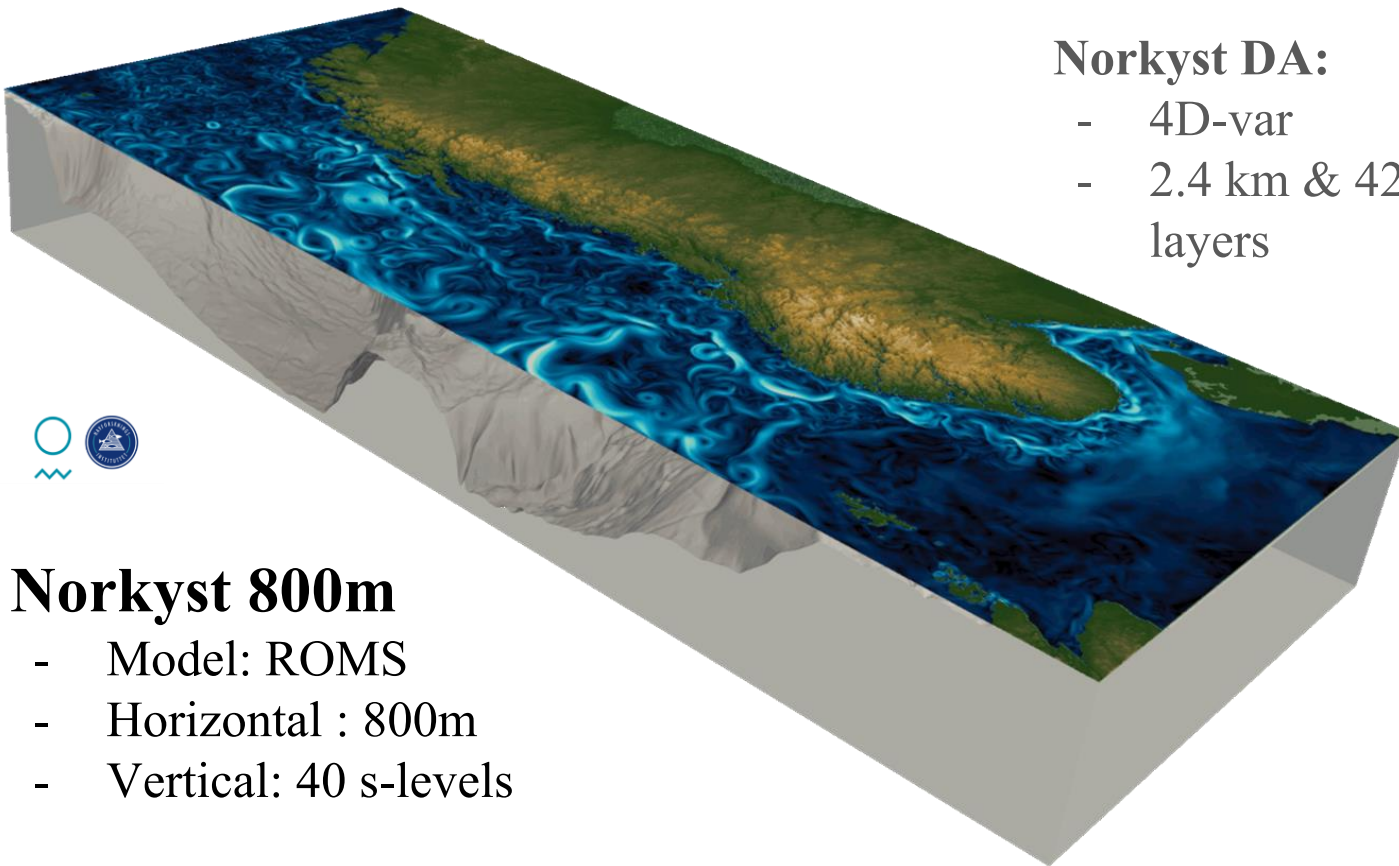


Fig: Yvonne Gusdal



Norway has a complex & long coastline

# Operational models



## Norkyst DA:

- 4D-var
- 2.4 km & 42 layers

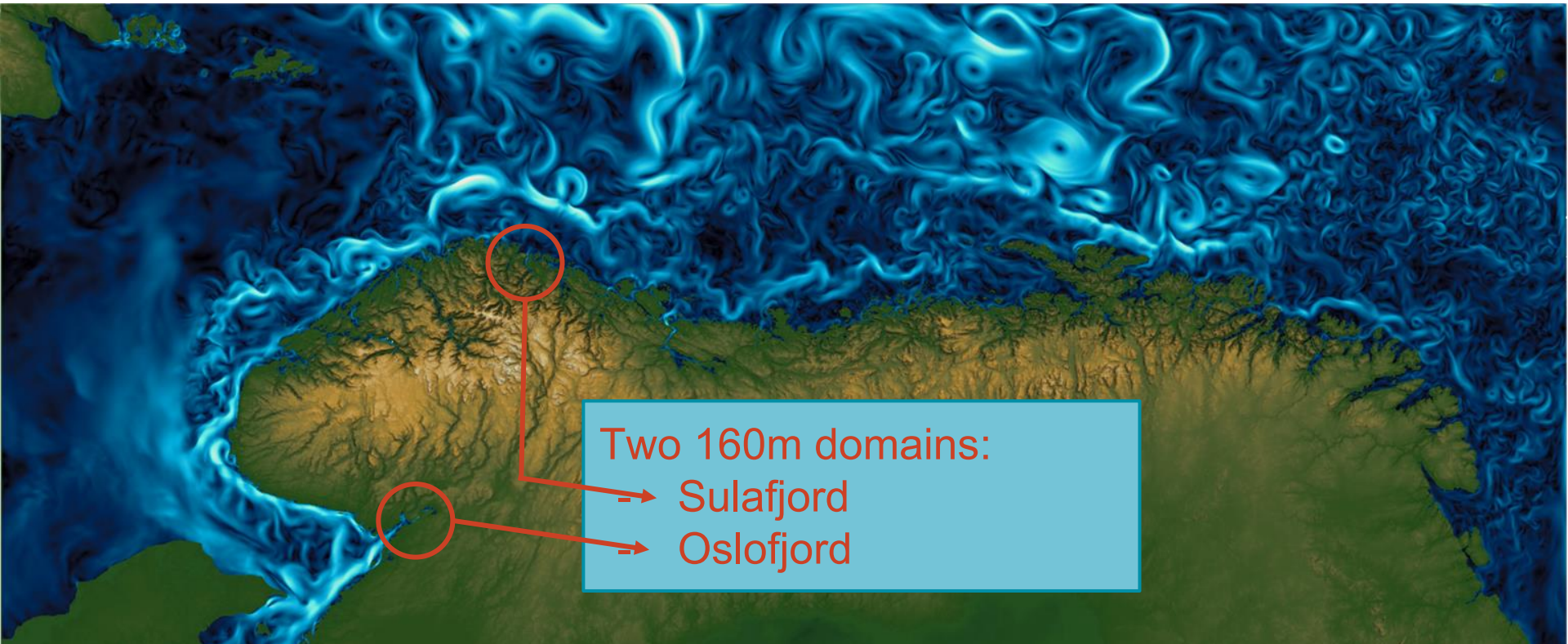


## Norkyst 800m

- Model: ROMS
- Horizontal : 800m
- Vertical: 40 s-levels

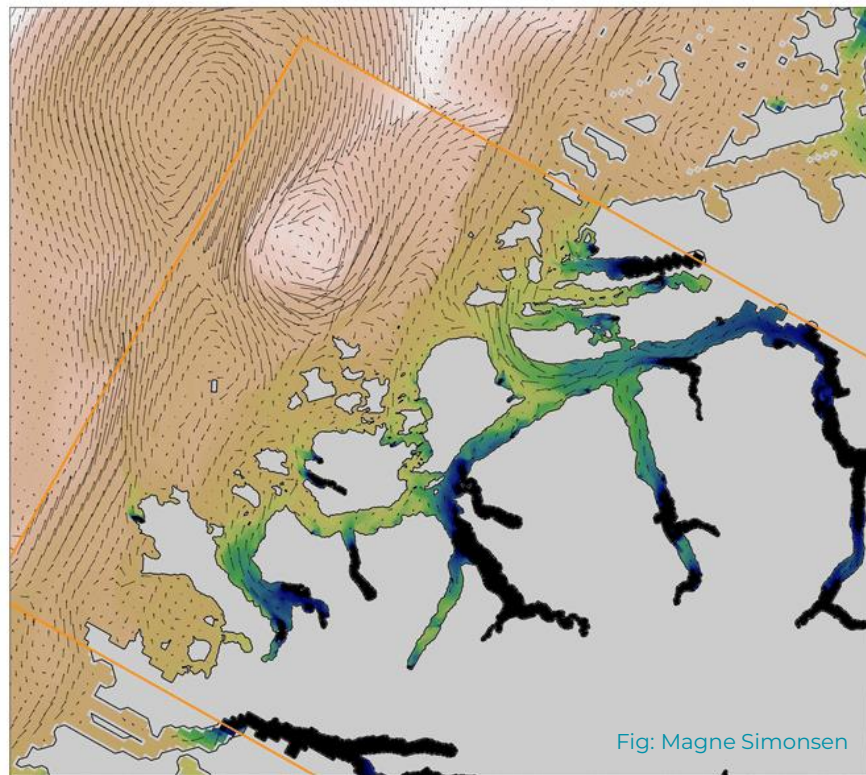
# Operational models

**Norkyst 800m**



# Norkyst 160m Sulafjord

Current speed at 0.0m depth  
2018-10-01 00:00

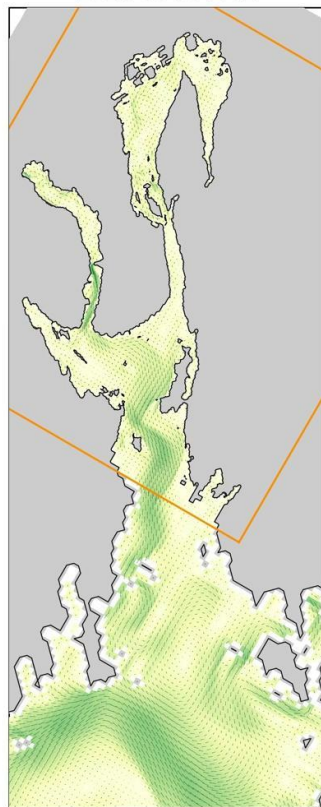


Hyper-resolution fjord models

- two-way nested
- interaction between coastal current & near-shore circulation

# Norkyst 160m Oslofjord

Currents at 1.0m depth  
2024-05-03T00Z



salt at 1.0m depth  
2024-05-03T00Z

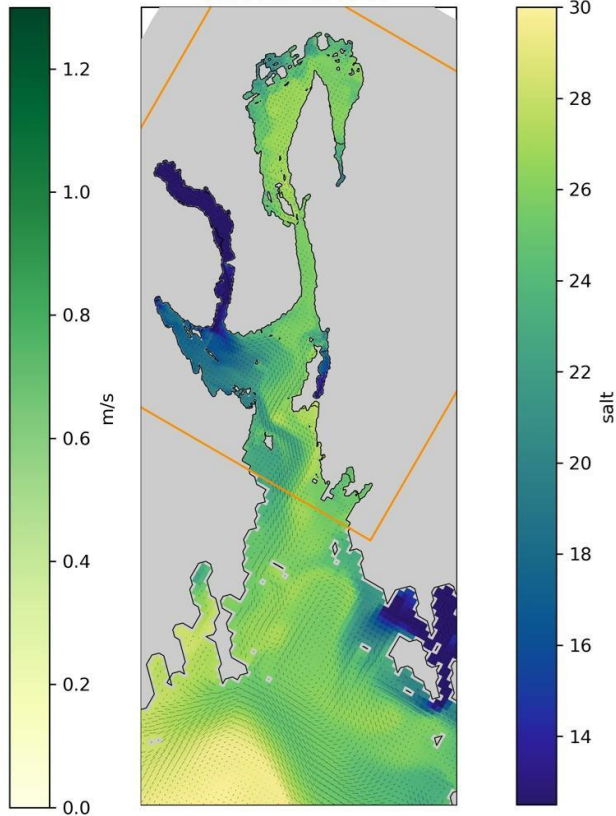
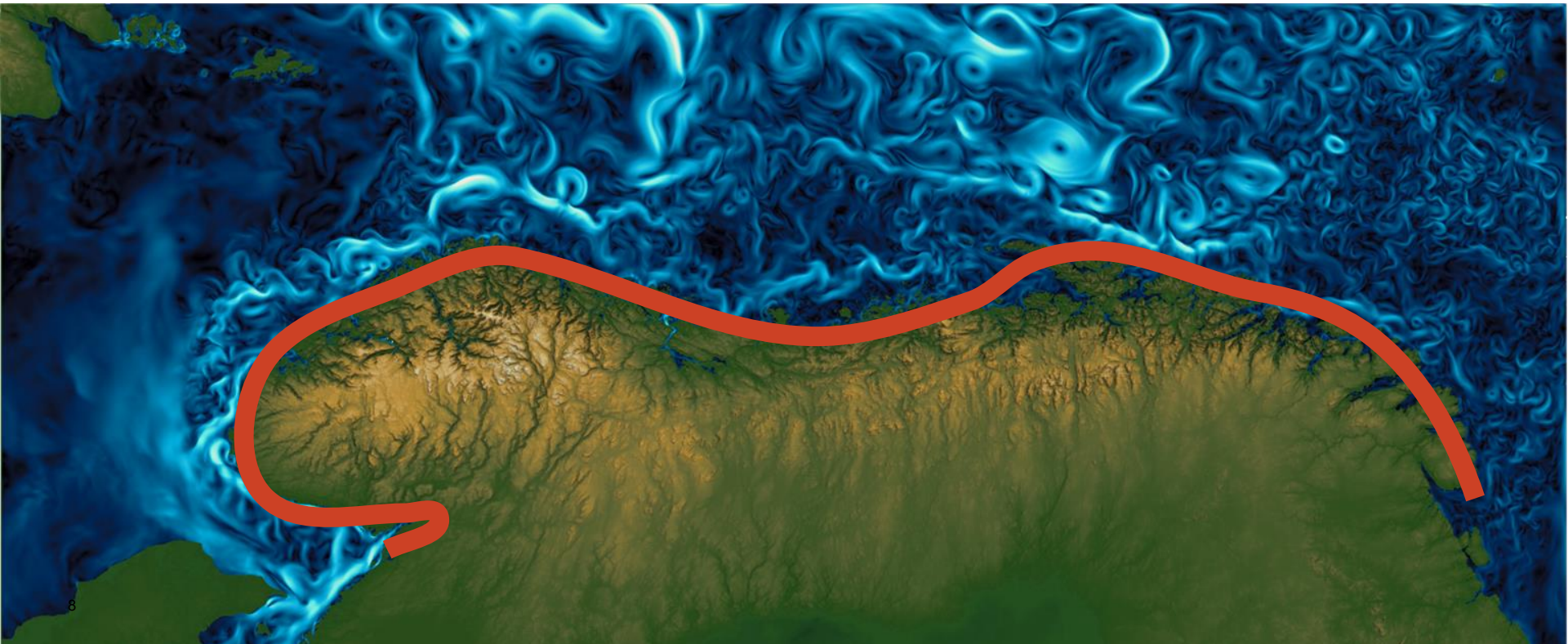


Fig: Magne Simonsen

# Main Goal:

## 160m resolution for the entire Norwegian coast



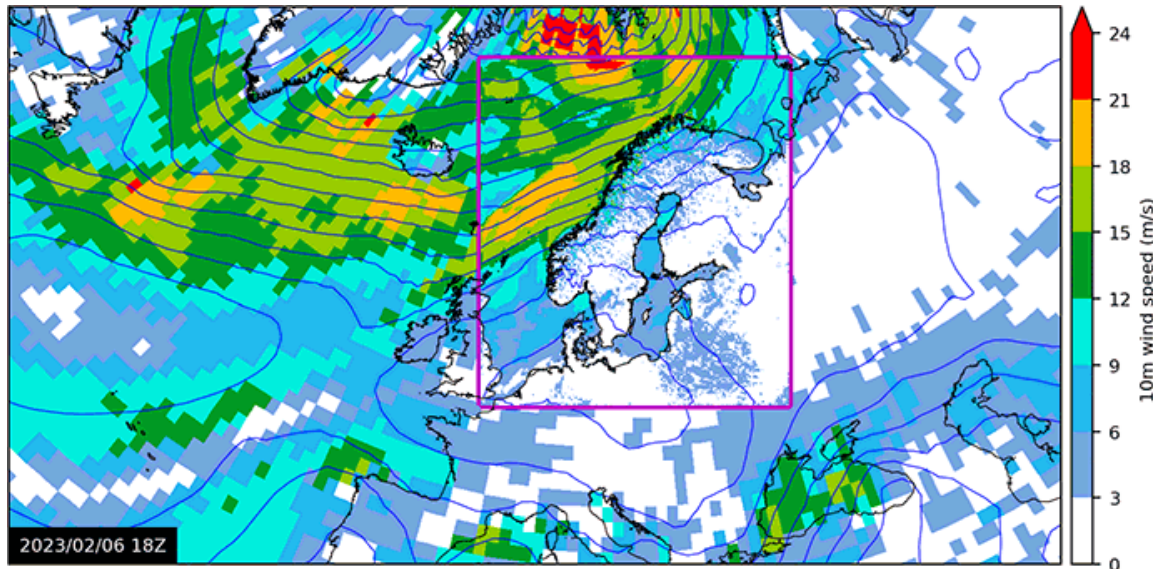
# How?

## Use the same setup as newest ML weather models!



**anemoi**

The ML framework  
developed by ECMWF  
and others.  
Used to create the AIFS  
and

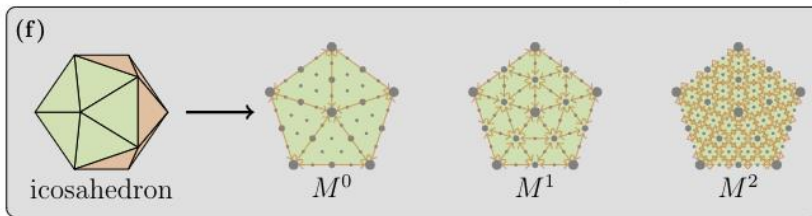
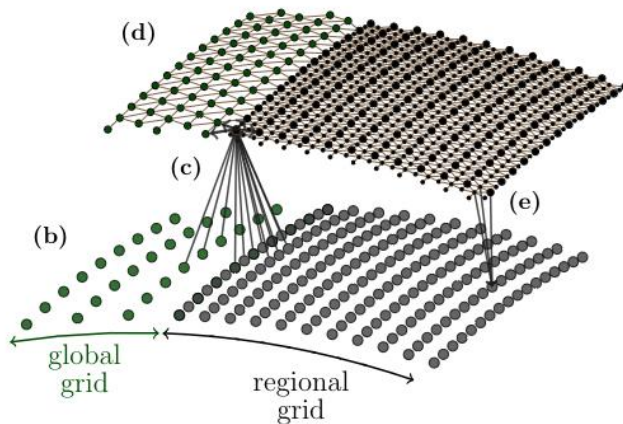
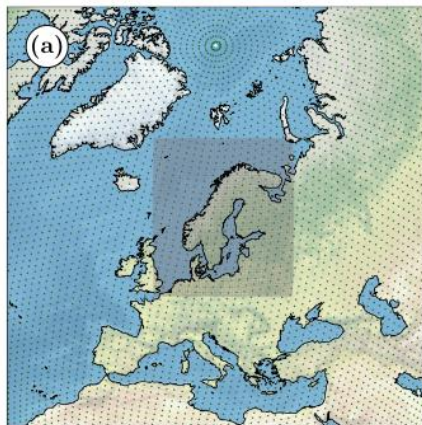


MET Norway's "Bris" weather model

Nipen et al. 2024 arXiv:2409.02891v1

<https://www.ecmwf.int/en/about/media-centre/aifs-blog/2024/data-driven-regional-modelling>

# MET Norways Bris weather model



- Graph Neural Network
- Processor, encoder, decoder architecture
- Icosahedron graph structure
- Global & regional domain

# Developing a data-driven ocean model using Anemoi

## The plan:

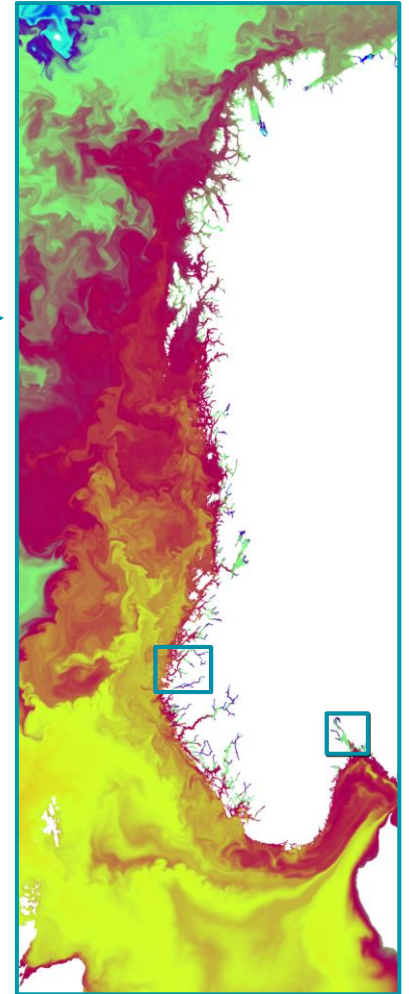
1. Simplifications
  - a. only surface layer
  - b. predict only 5 variables:  
u,v, temp, salt, surface elevation
  - c. 3h time steps
2. Train on 800m data
3. Train on 160m data
4. Increase complexity (3D, ...) and repeat step 2. & 3.

# Training Data

- Norkyst 800m hindcast ~12 years

Phase 1

Norkyst models



# Training Data

- Norkyst 800m hindcast ~12 years
- Norkyst 160m operational ~1 year 3 months

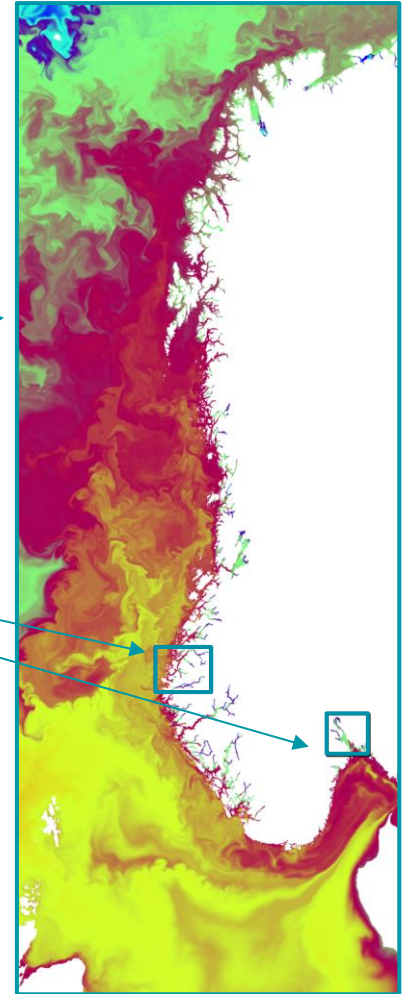
Phase 1

Phase 2  
(fine-tuning)

- hourly data, 40 vertical layers

→ 1st approach: only surface and 3h time step!

Norkyst models



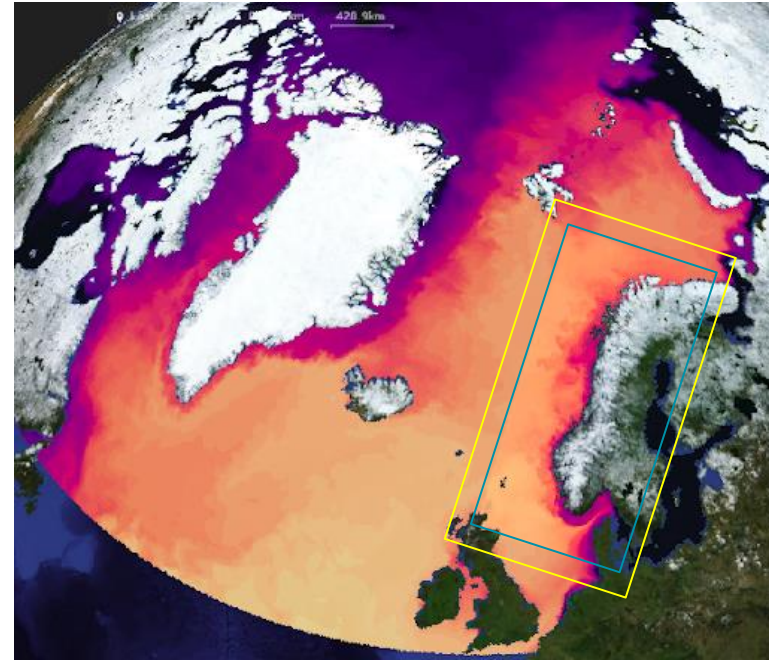
# Boundary conditions

Two approaches:

1. Shrink Norkyst domain and use the outermost grid cells as boundary conditions for the ML model.
2. Take values from Topaz (6km) and add them around the Norkyst (800m) domain.

How many km around model domain?

→To be tested...



# Forcing variables

<b>Available Forcings</b>
River runoff
Atm. forcing
Topography
Boundary (previous slide)

Wind (u,v)
Atm. temp
Solar rad.
Rel. humidity
Cloud cover
Precipitation
Air pressure

# Graph node positions (Oslo region)

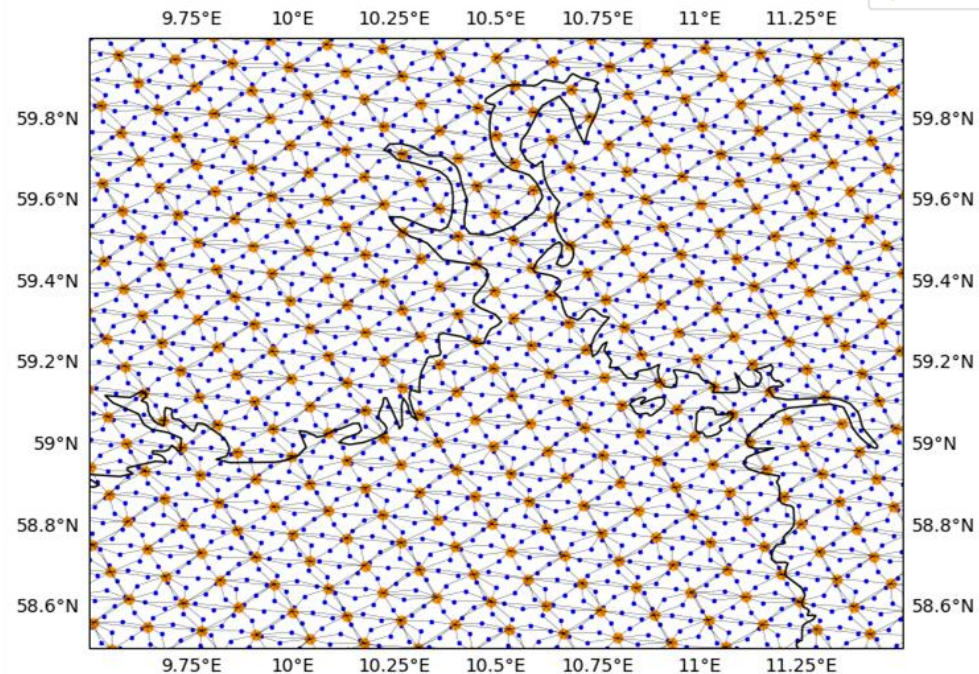
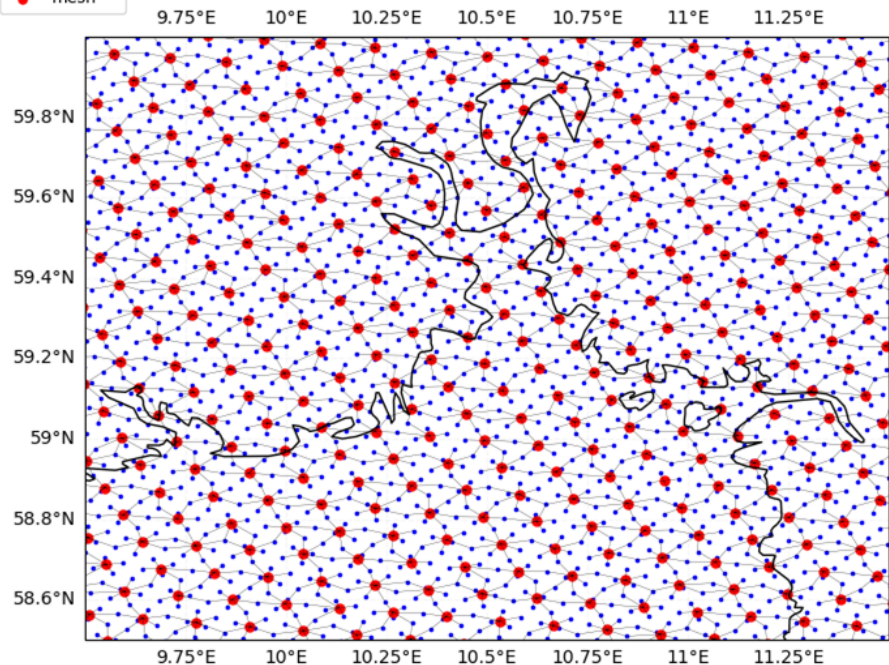
Processor

Encoder

Decoder

— encoder  
• grid  
• mesh

— decoder  
• grid  
• mesh



# Preliminary results (800m surface)

```
$$ ssh -i .ssh/id_rsa_lumi kullmann@lumi.csc.fi
```

This LUMI login node is not available right now.  
Connection closed by 193.167.209.166 port 22



Results unavailable...

# Future plans

Find the best configuration of:

- model hyper parameters
  - - learning rate
  - - batch size
  - - graph node positions, weights, refinement level, ...
- size of boundary forcing domain
- time step size vs. roll-out length
- adding depth layers
- test performance 160m domains
- ...

# Collaborators & Acknowledgements



## From MET Norway

Mateusz Matuszak, main co-author together with Ina Kullmann.

Thomas Nipen and the rest of the MET Norway Bris team: Håvard Homleid Haugen, Magnus Sikora Ingstad, Aram Farhad Salihi, Ivar Seierstad, Paulina Tedesco.

Others from the Department of Ocean and Ice including Johannes Röhrs, Kai H. Christensen, Ann Kristin Sperrevik, Yvonne Gusdal and Magne Simonsen.

## From ECMWF

Matthew Chantry and the rest of the AIFS team & others in the Anemoi-community.

**Thank you for your time!**

# Backup Slides

# Config setup

## LUMI:

- 3.1 M grid points (ocean & land)
- Graph resolution / refinement level 10 (would like to test higher)
- Hardware: 16 nodes, 8 GPUs per node, model parallel 4
- Model: Graph transformer
  - Processor, encoder, decoder: defaults
  - Trainable parameters 0
  - 1024 channels (for now)
- Training: defaults
  - learning rate:  $6.25e-5$ , min:  $3e-7$

# ML forcings in Anemoi

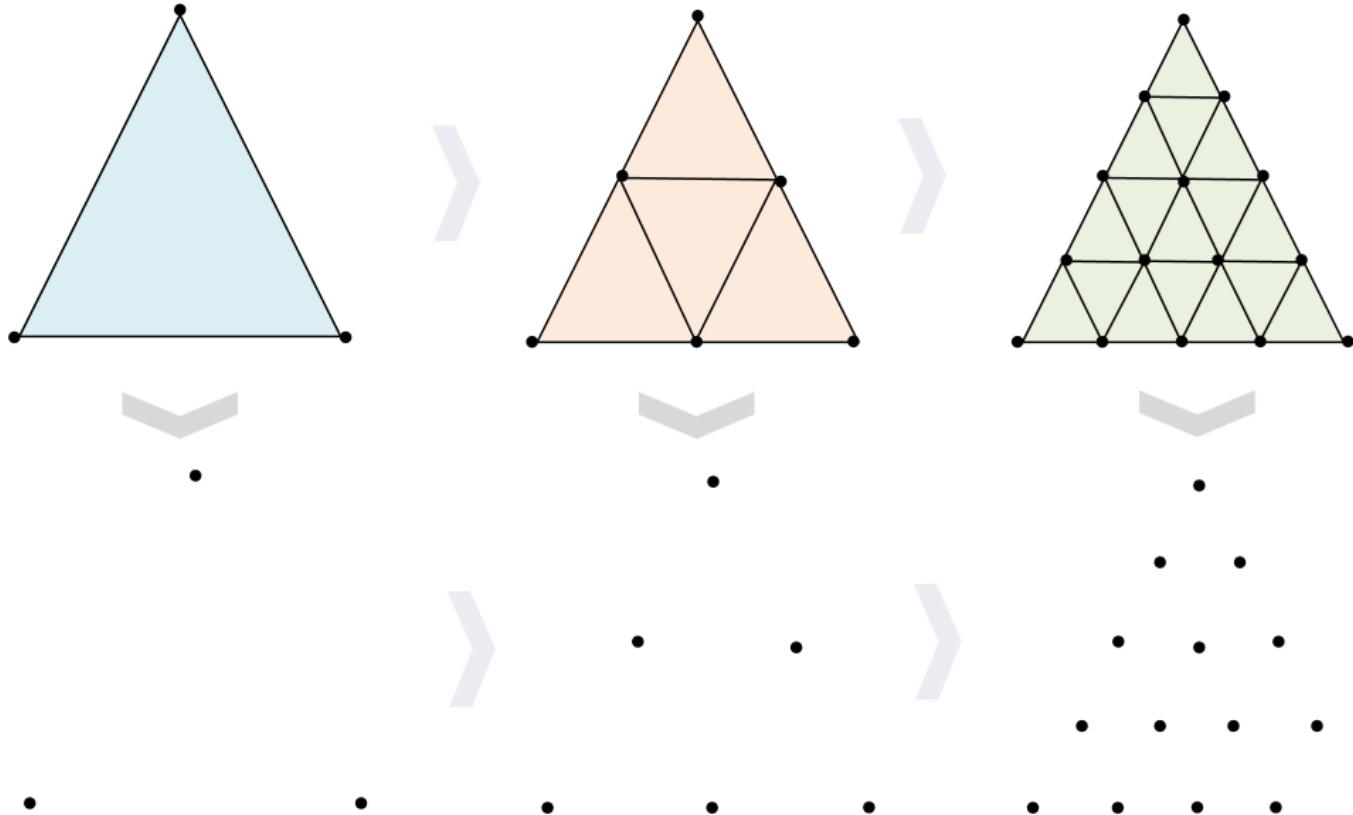
cos latitude
cos longitude
sin latitude
sin longitude
cos julian day
cos local time
sin julian day
sin local time
insolation

## Extra training data

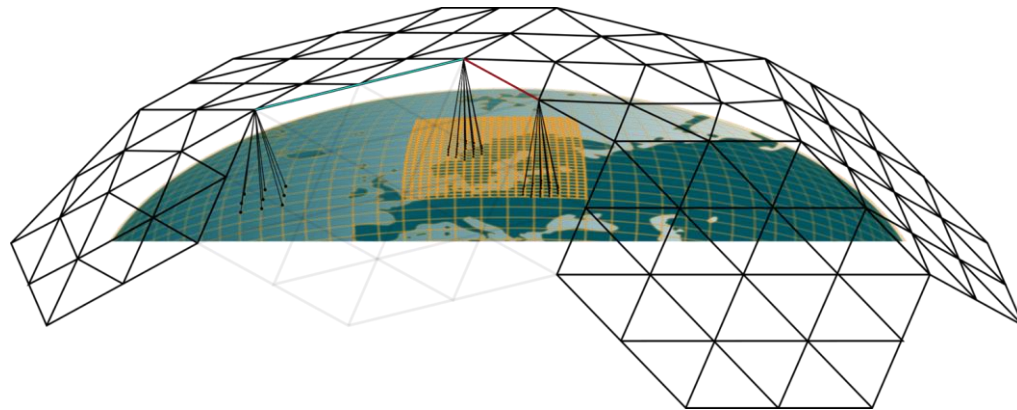
- Topaz system, northern globe.
- Norkyst v3 800m operational archive ~1 year
- Norkyst v2 ~10 years.
- NorFjord - 160m data for the whole Norwegian coast ~7 years.

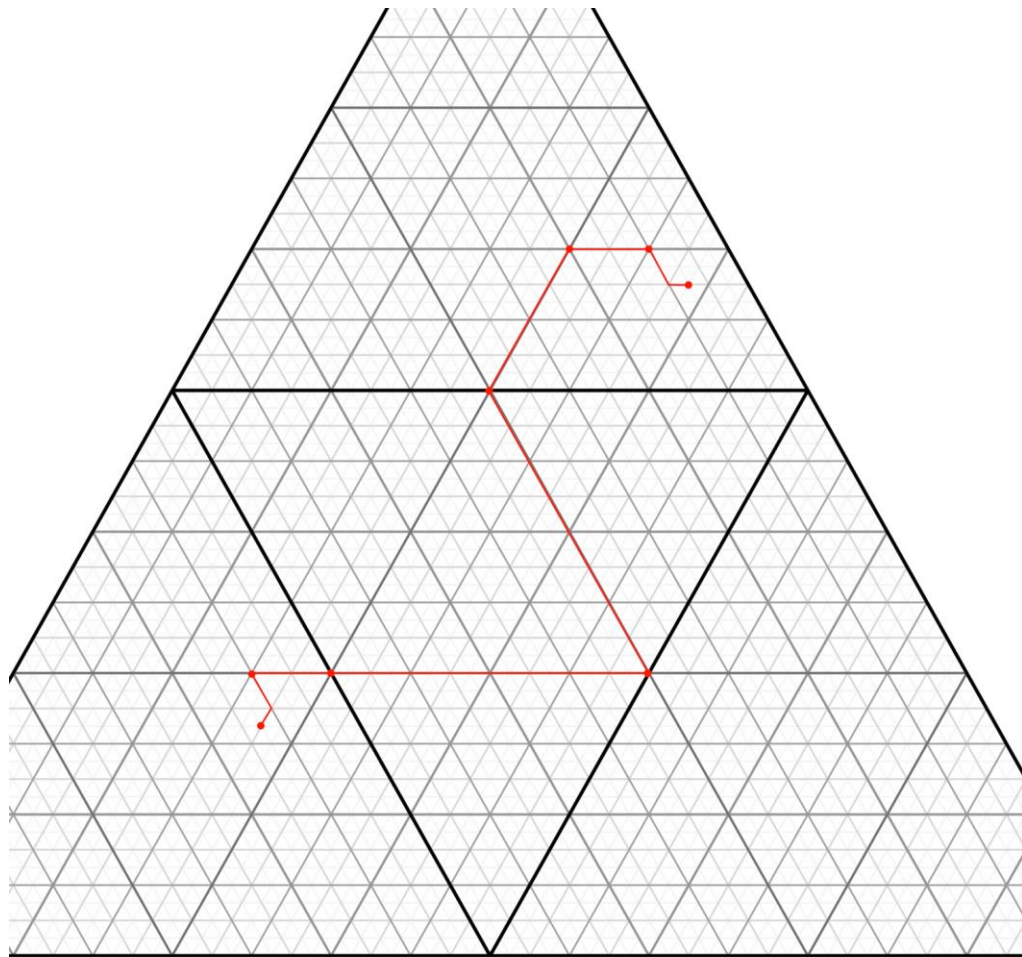
As of now we don't think there will be a need for global training data.

# Triangular refined Icosahedron



[https://anemoui.readthedocs.io/projects/graphs/en/latest/graphs/node\\_coordinates/tri\\_refined\\_icosahedron.html](https://anemoui.readthedocs.io/projects/graphs/en/latest/graphs/node_coordinates/tri_refined_icosahedron.html)





# TOPAZ

6 km  
50 vertical layers

Models:  
HYCOM + CICE + ECOSMO

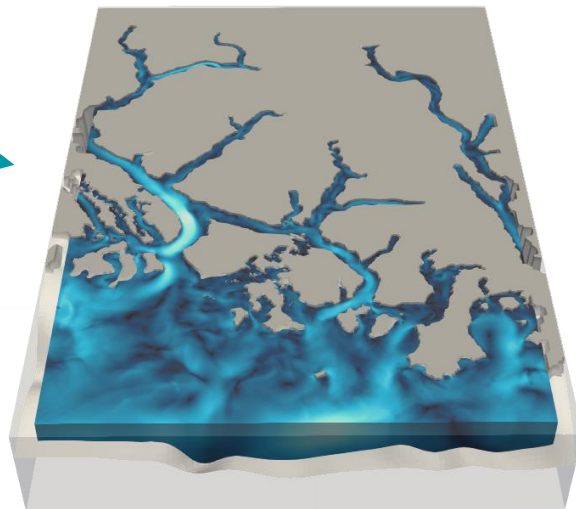
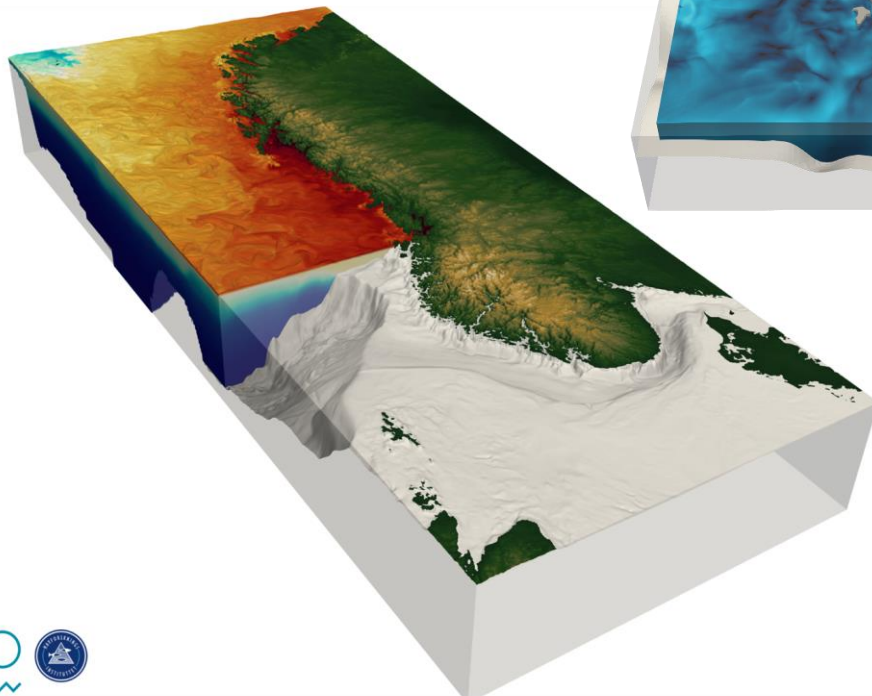
DA:  
EnKF (100 members)

## Norkyst:

Horizontal : 800 m / 160m  
Vertical: 40 s-levels

Model:  
ROMS

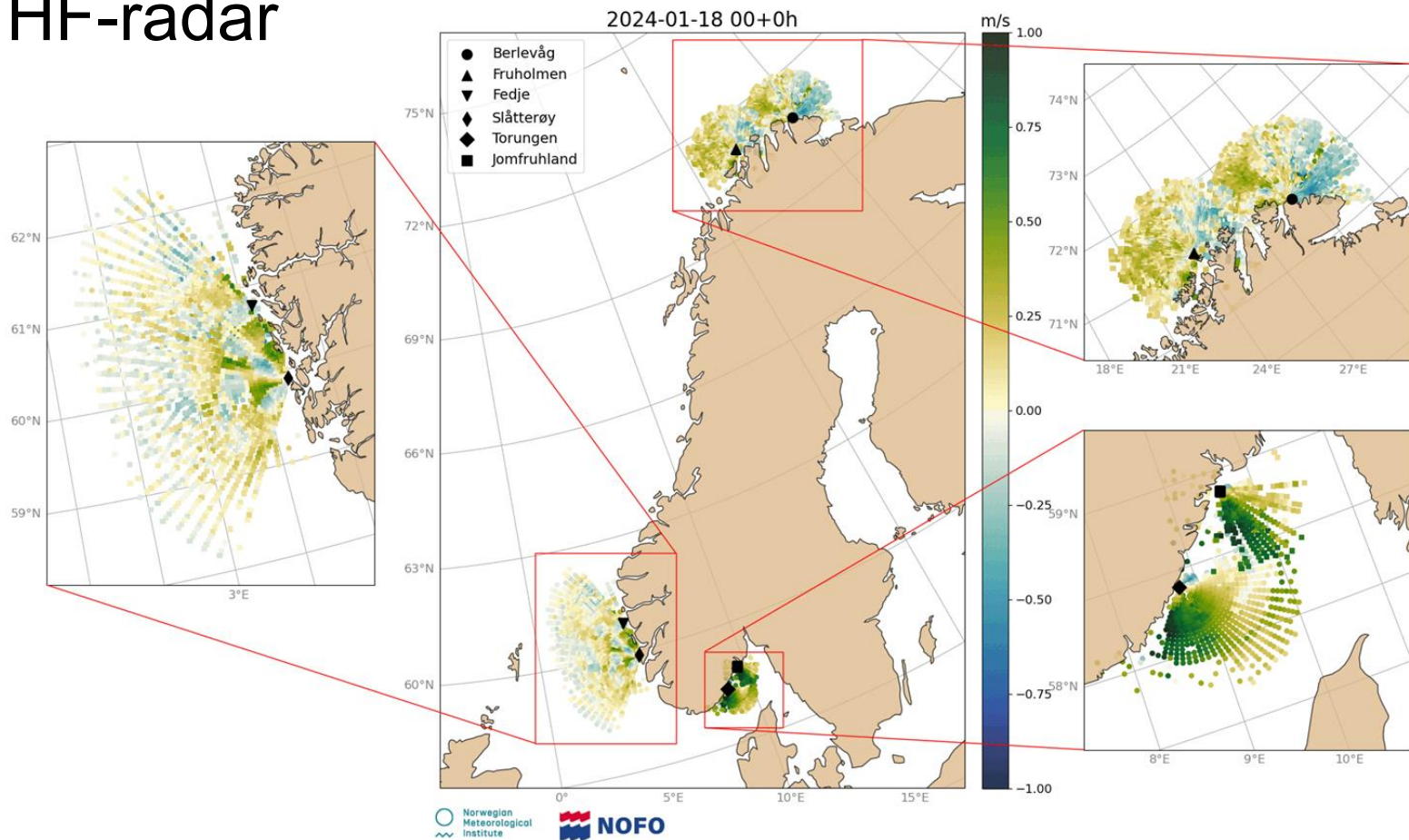
DA:  
4D-var



# Additional observations

What we would like to address in the future

# Coastal HF-radar



Can we use this  
information in our data-  
driven model?

ocean color from MODIS  
[https://earthobservatory.n  
asa.gov/](https://earthobservatory.nasa.gov/)

