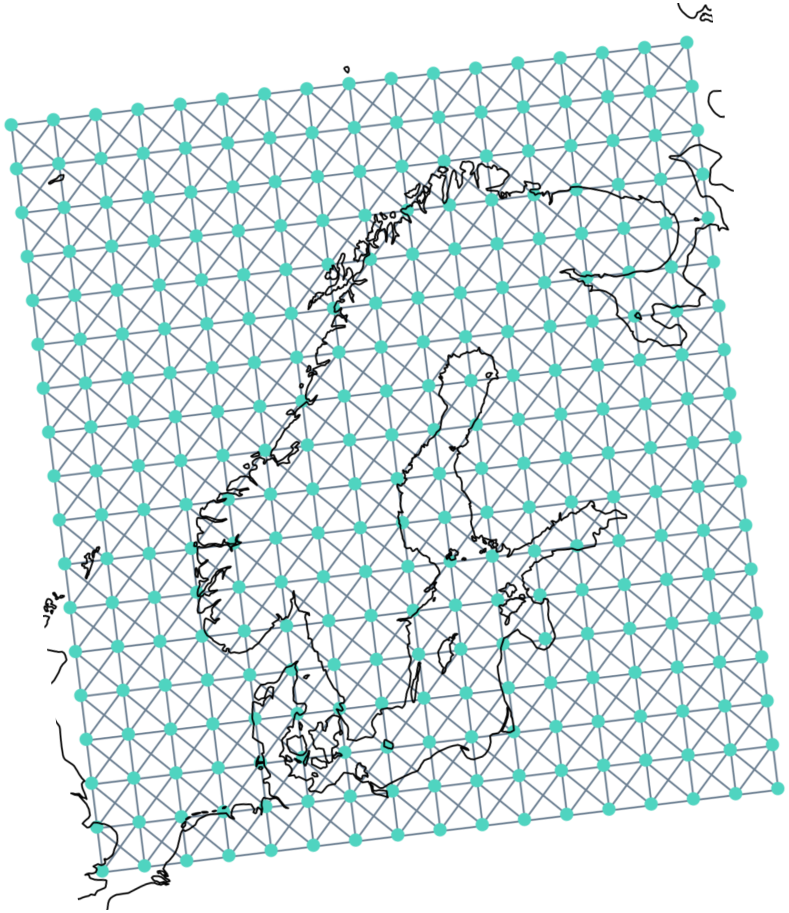
Graph-based Neural Weather Prediction for Limited Area Modeling

Seminar @ DMI, 10/10 2023

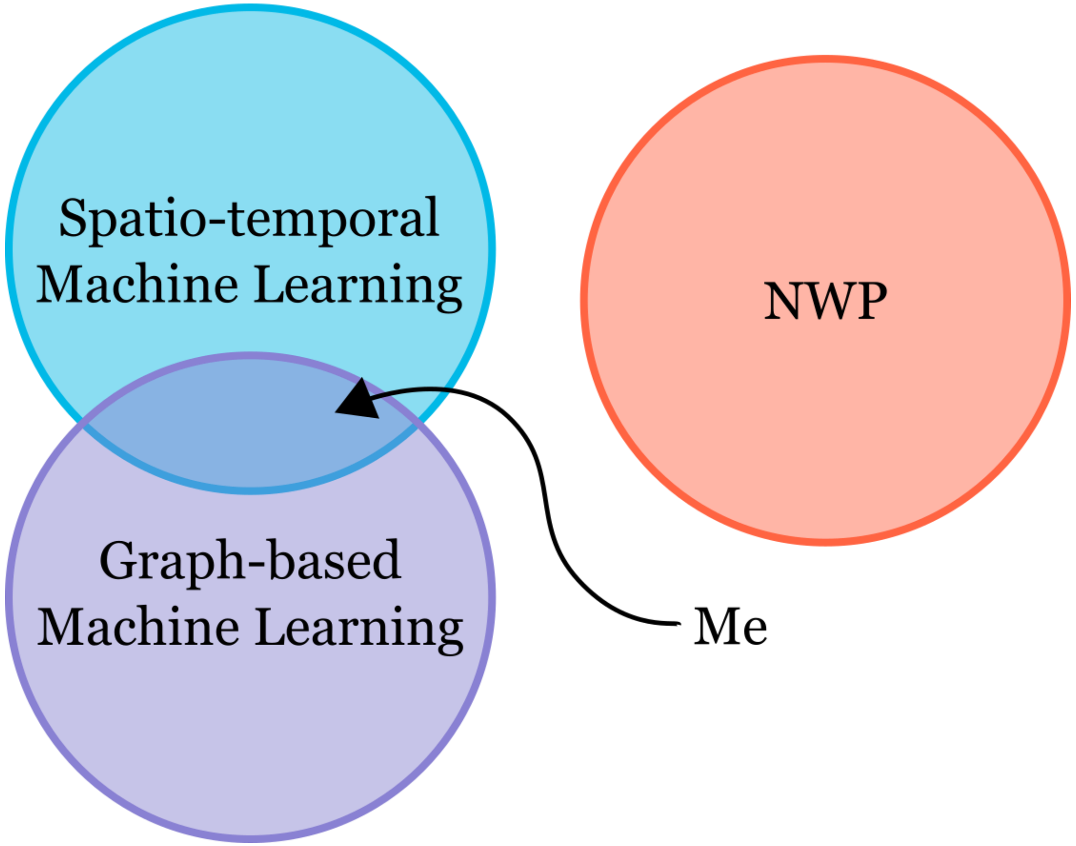
Joel Oskarsson
Division of Statistics and Machine Learning,
Department of Computer and Information Science,
Linköping University, Sweden

Joint work with: Tomas Landelius (SMHI), Fredrik Lindsten (LiU)

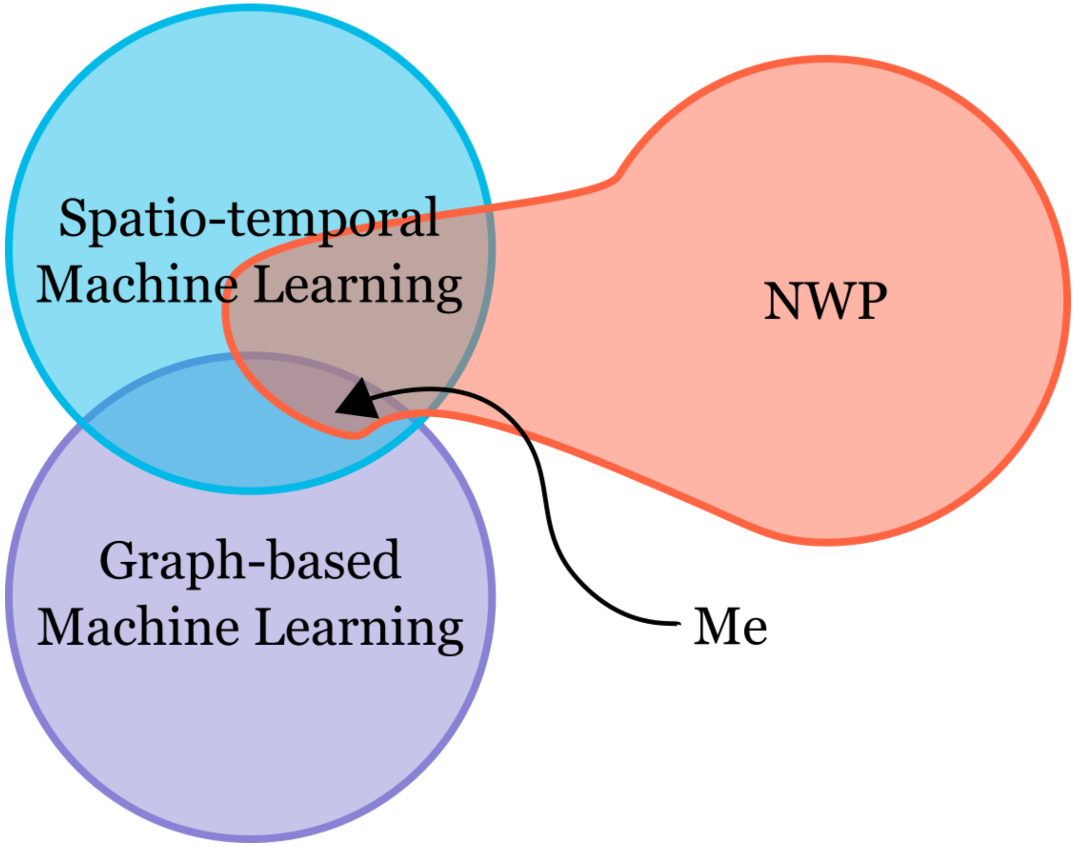


LINKÖPING UNIVERSITY

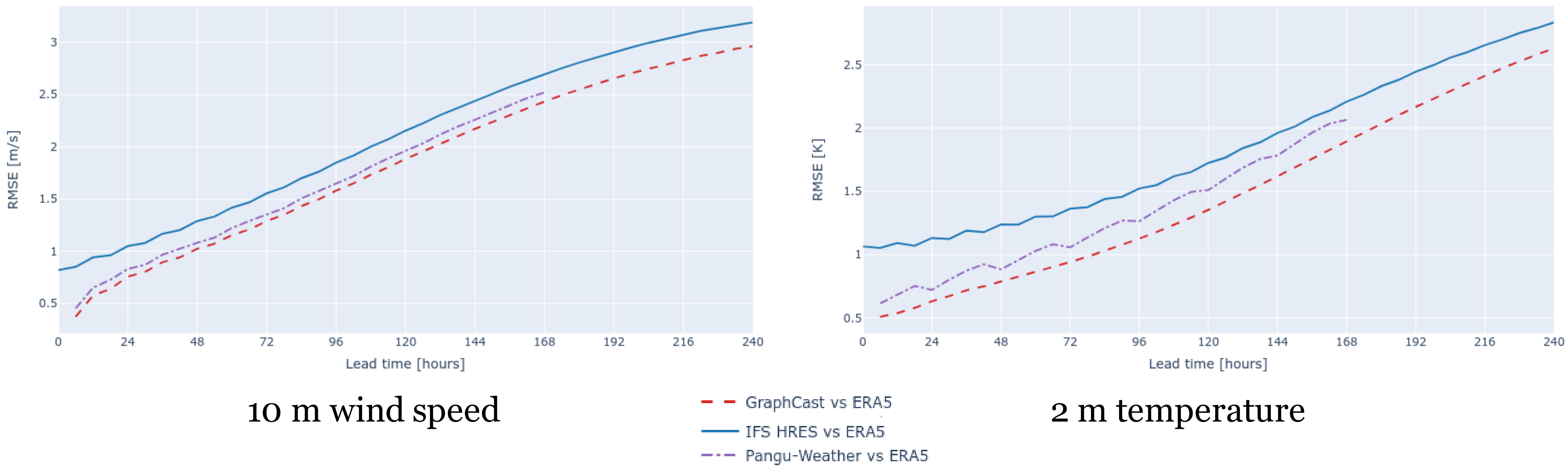
My machine learning background



My machine learning background



Machine learning for NWP



10 m wind speed

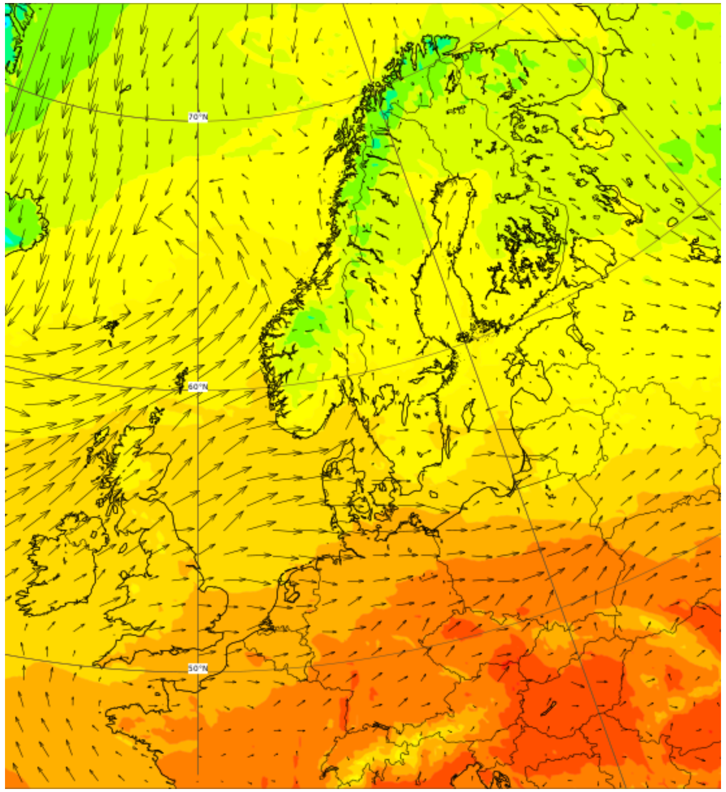
2 m temperature

"In our view, we are currently placed at an exciting moment in weather forecasting history." - ECMWF[1]

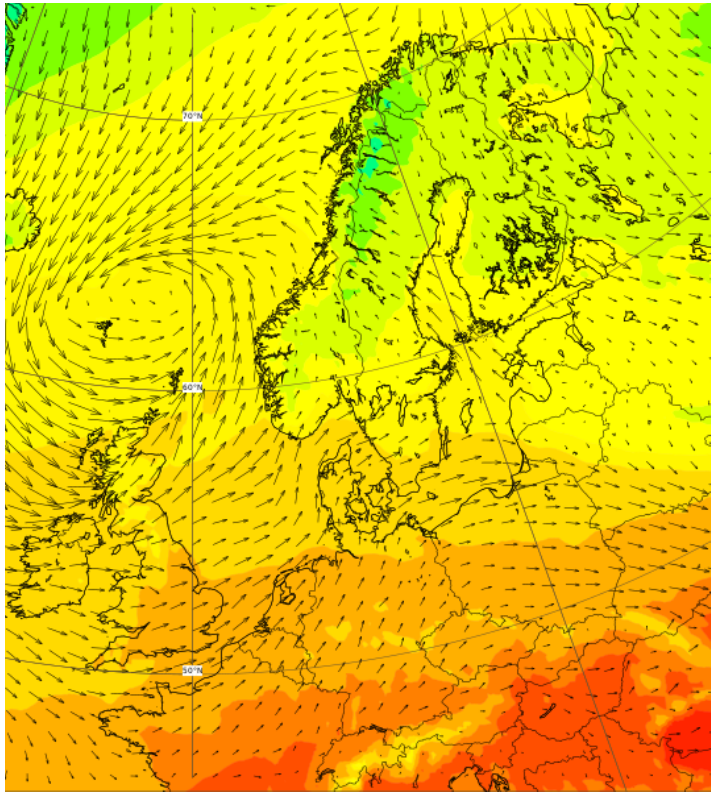
[1] https://www.ecmwf.int/en/about/media-centre/science-blog/2023/rise-machine-learning-weather-forecasting

Forecasts from ECMWF

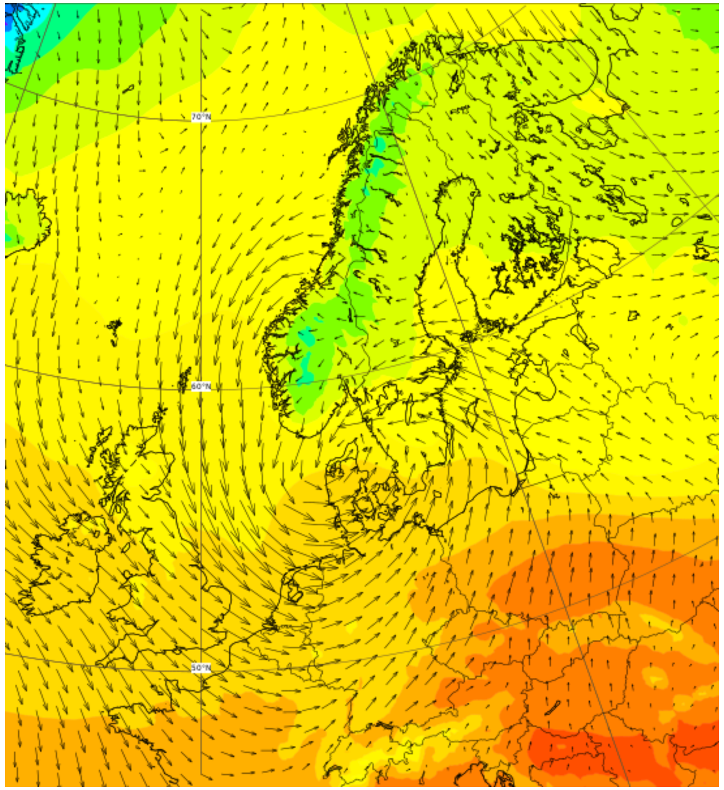
2m temperature + 10m wind, 1 week lead time



IFS



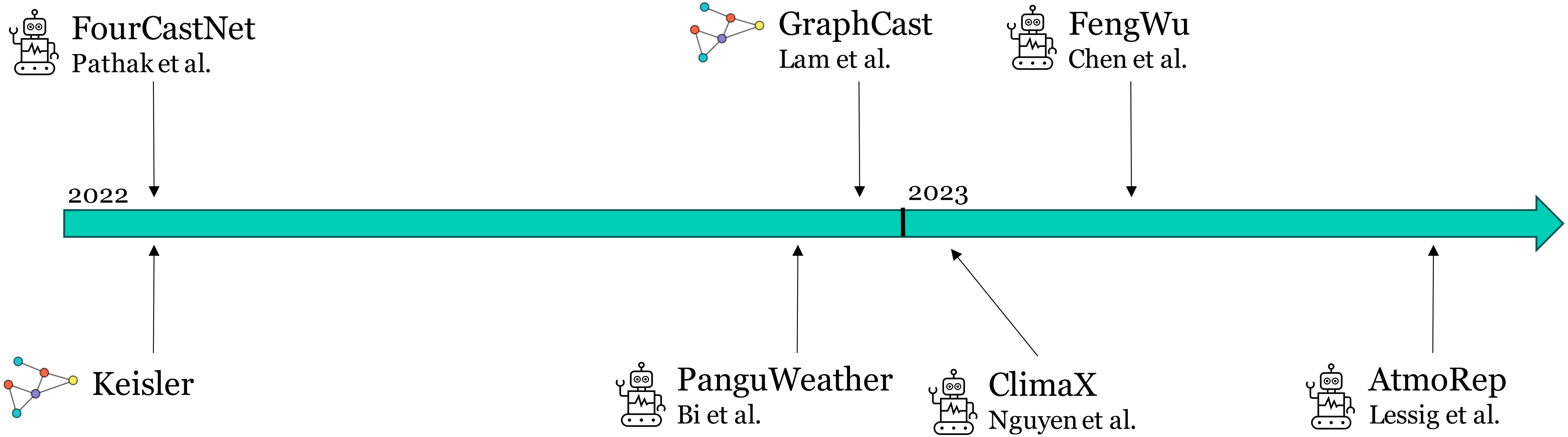
GraphCast



FourCastNet

Forecasts from:
https://charts.ecmwf.int/?facets=%7B%22Product%20type%22%3A%5B%22Experimental%3A%20Machine%20learning%20models%22%5D%7D

A timeline of global model

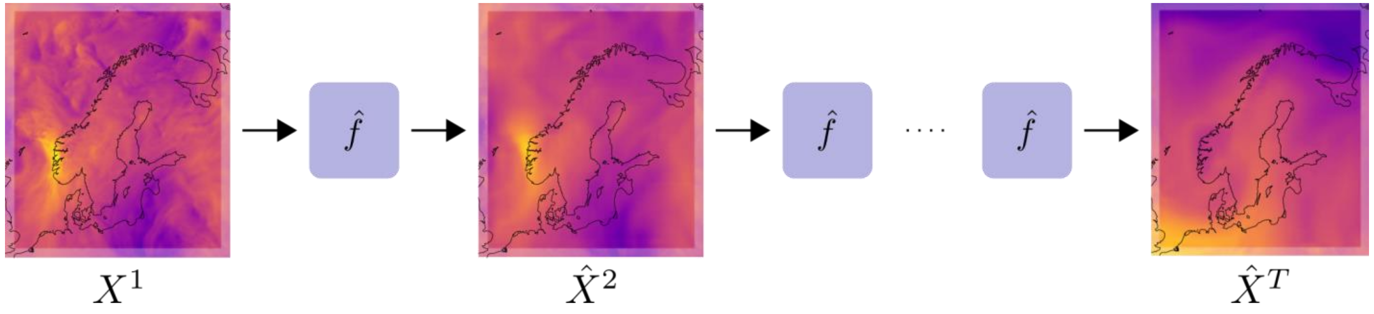


= Graph-based

= Transformer-based

Machine learning for NWP: How?

- Weather state X^t
- Dynamics model $X^t = f\left(X^{t-1}, \ldots, X^{t-p}\right)$
- Approximate with machine learning model $\hat{f} \approx f$



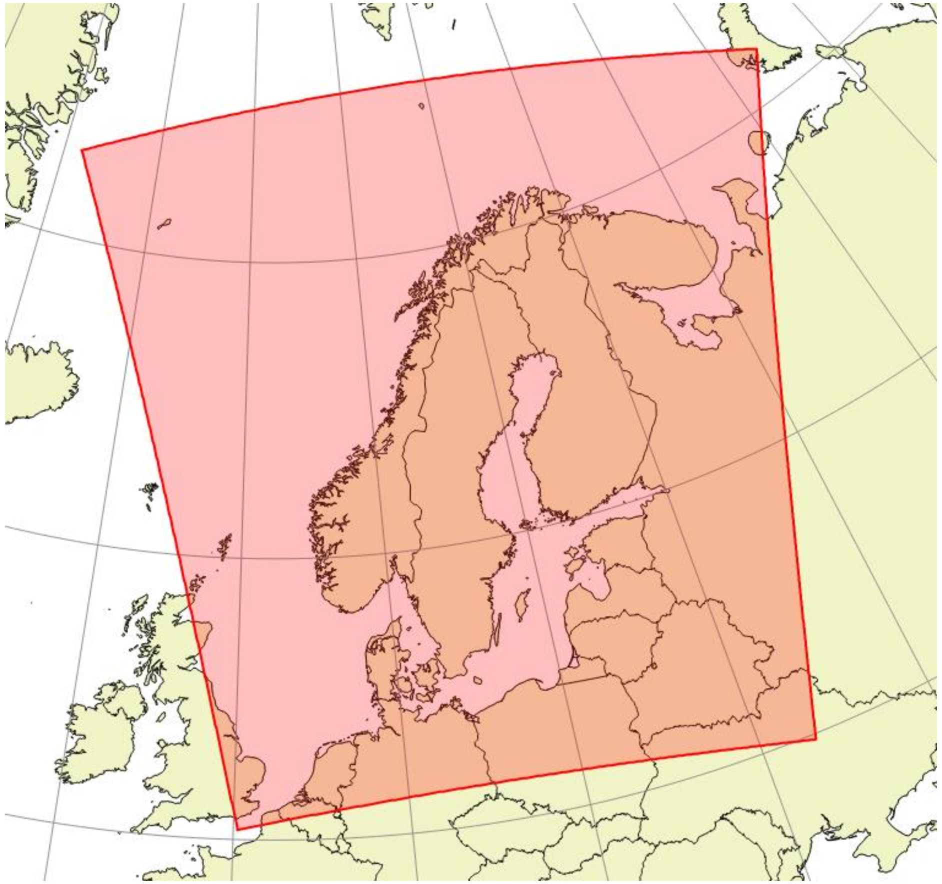
- Train on dataset of trajectories $X^1, X^2, \ldots, X^T$.
  - Forecast data: Fast surrogate model
  - Reanalysis data: Surpass existing NWP

Neural Weather Prediction for Limited Area Modeling

MetCoOp Ensemble Prediction System (MEPS)

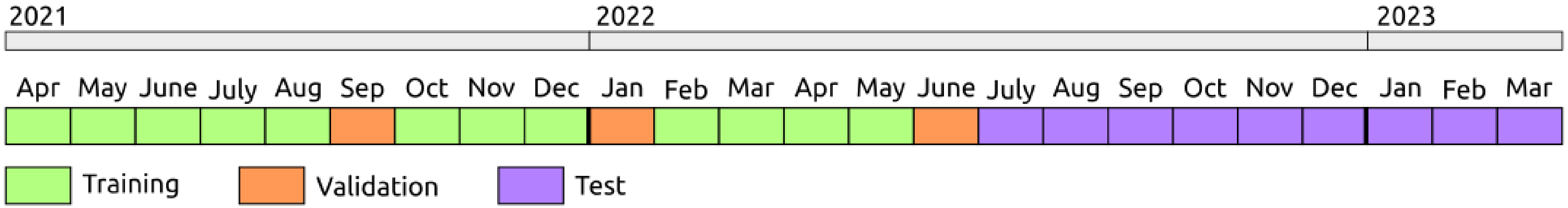
- 960×1080 (2.5 km) x 65 vertical
- Non-hydrostatic dynamics
  - HARMONIE-AROME physics
- IFS HRES and IFSENS boundaries
- 66h forecasts run hourly with 5 ensemble members

- Idea: Emulate with fast deep learning model



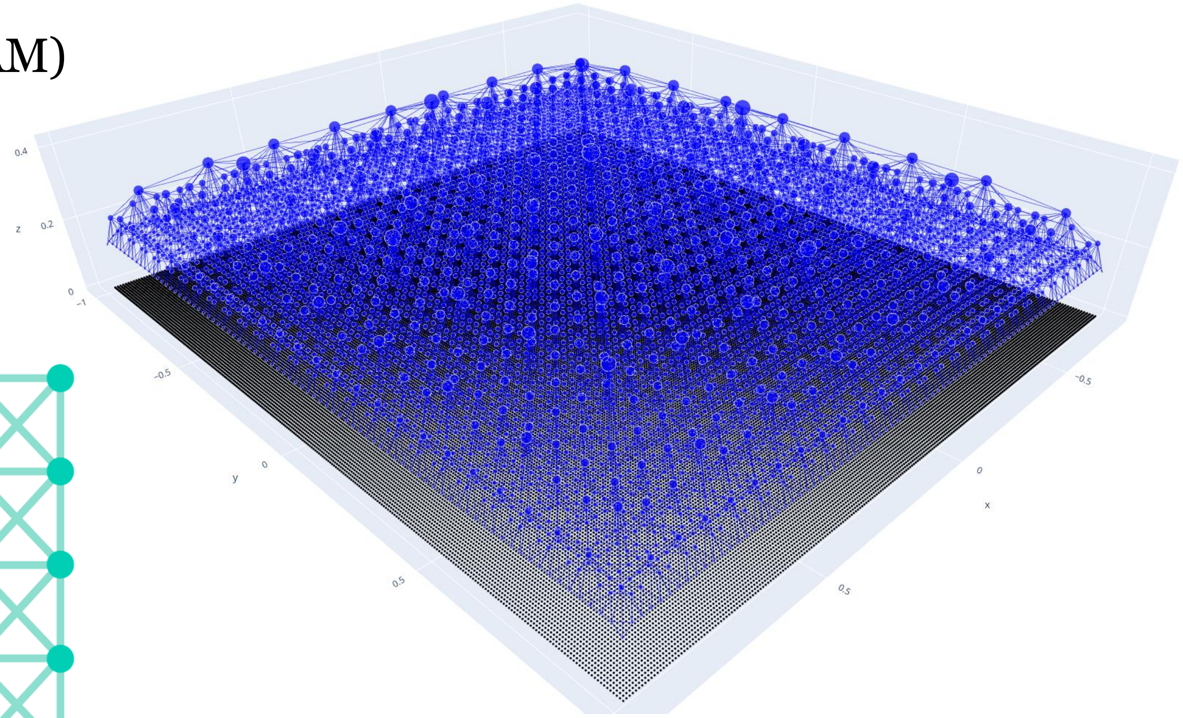
Dataset

- Subset of atmospheric variables used:
  - Pressure (surface, MSL)
  - Geopotential (500, 1000 hPa)
  - Wind (lev 65, 850 hPa)
  - Temperature (2m, lev 65, 500, 850 hPa)
  - Relative humidity (2m, lev 65)
  - Total water vapor column
  - Net short- and longwave 3h radiation
- Spatial down-sampling ×4 (10 km)
- Additional forcing inputs:
  - TOA radiation, time, land/water mask
  - Forecast as boundary forcing
- 10 forecasts per day from ~2 years
- 3h time-steps

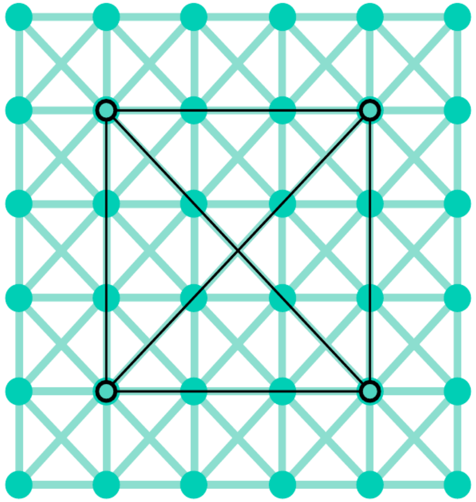


Graph-based Neural Weather Prediction

- Graph framework for Limited Area Modeling (LAM)
  - Adapting GraphCast[1] to LAM: GC-LAM



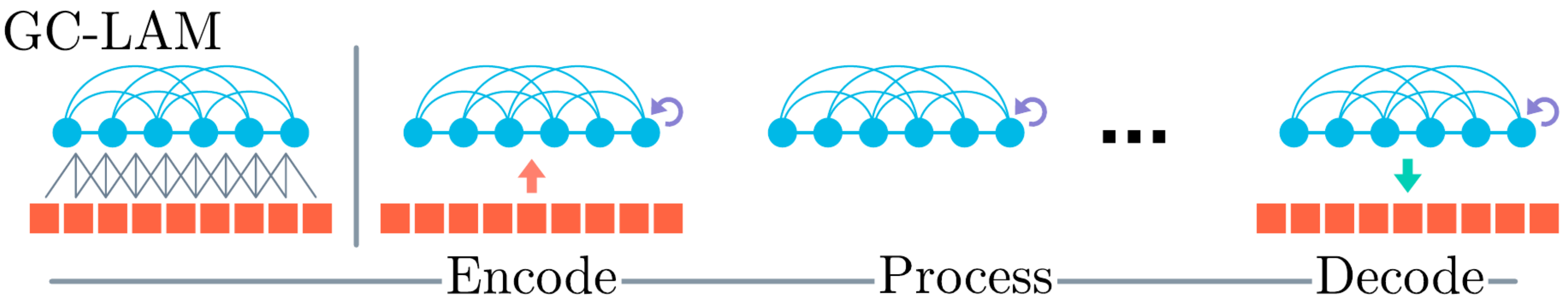
- Grid nodes (grid cells)
- Mesh graph
- Multi-scale edges



[1] R. Lam, et al. *GraphCast: Learning skillful medium-range global weather forecasting*, 2022.

Graph-based Neural Weather Prediction

- The encode-process-decode framework



- Graph Neural Networks (GNNs)

A Brief Introduction to Graph Neural Networks

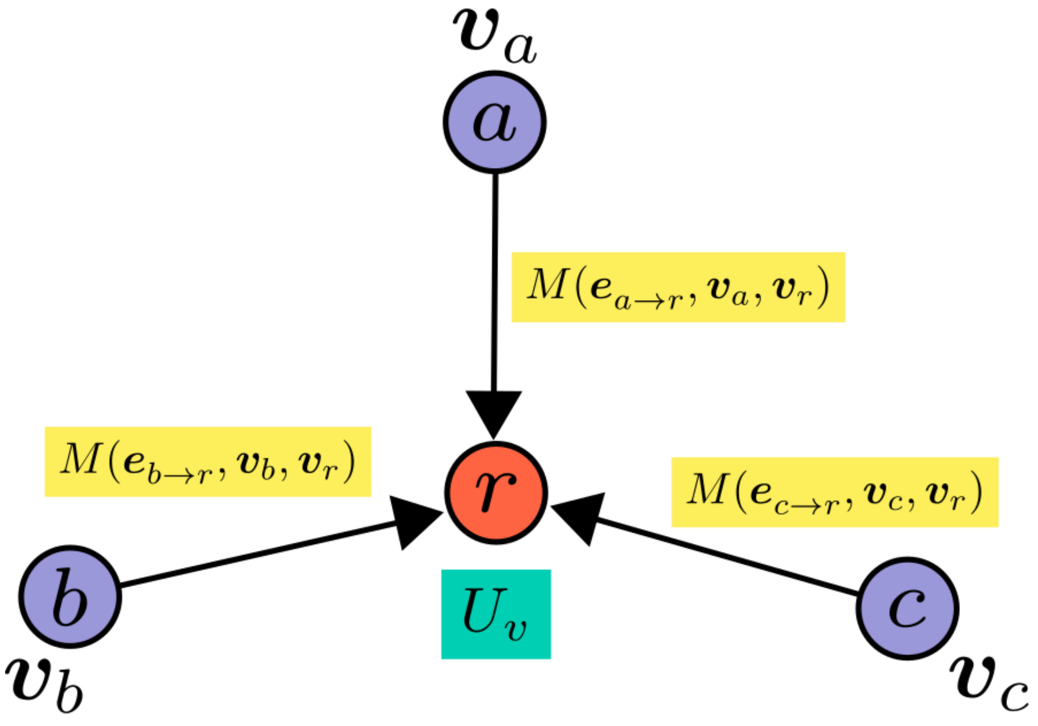
Graph Neural Networks (GNNs)

- Message Passing Neural Network[1]
- Vector representations of
  - Nodes $\boldsymbol{v}$
  - Edges $\boldsymbol{e}_{s\rightarrow r}$

$$\tilde{\boldsymbol{e}}_{s\rightarrow r} \leftarrow M(\boldsymbol{e}_{s\rightarrow r}, \boldsymbol{v}_s, \boldsymbol{v}_r)$$

$$\boldsymbol{e}_{s\rightarrow r} \leftarrow U_e(\boldsymbol{e}_{s\rightarrow r}, \tilde{\boldsymbol{e}}_{s\rightarrow r})$$

$$\boldsymbol{v}_r \leftarrow U_v\left(\boldsymbol{v}_r, \sum_{s:(s,r)\in\mathcal{E}} \tilde{\boldsymbol{e}}_{s\rightarrow r}\right)$$



$$\mathcal{E} = \{(a,r),(b,r),(c,r)\}$$

[1] J. Gilmer, et al. *Neural Message Passing for Quantum Chemistry*, 2017. ICML.
P. Battaglia, et al. *Relational inductive biases, deep learning, and graph networks*, 2018.

Interaction Networks

- Message Passing Neural Network

$$\tilde{\boldsymbol{e}}_{s\to r} \leftarrow M(\boldsymbol{e}_{s\to r}, \boldsymbol{v}_s, \boldsymbol{v}_r)$$

$$\boldsymbol{e}_{s\to r} \leftarrow U_e(\boldsymbol{e}_{s\to r}, \tilde{\boldsymbol{e}}_{s\to r})$$

$$\boldsymbol{v}_r \leftarrow U_v\left(\boldsymbol{v}_r, \sum_{s:(s,r)\in\mathcal{E}} \tilde{\boldsymbol{e}}_{s\to r}\right)$$

- Interaction Network[1]
  - MLP = Multi-Layer Perceptron

$$\tilde{\boldsymbol{e}}_{s\to r} \leftarrow \mathrm{MLP}_E([\boldsymbol{e}_{s\to r}, \boldsymbol{v}_s, \boldsymbol{v}_r])$$

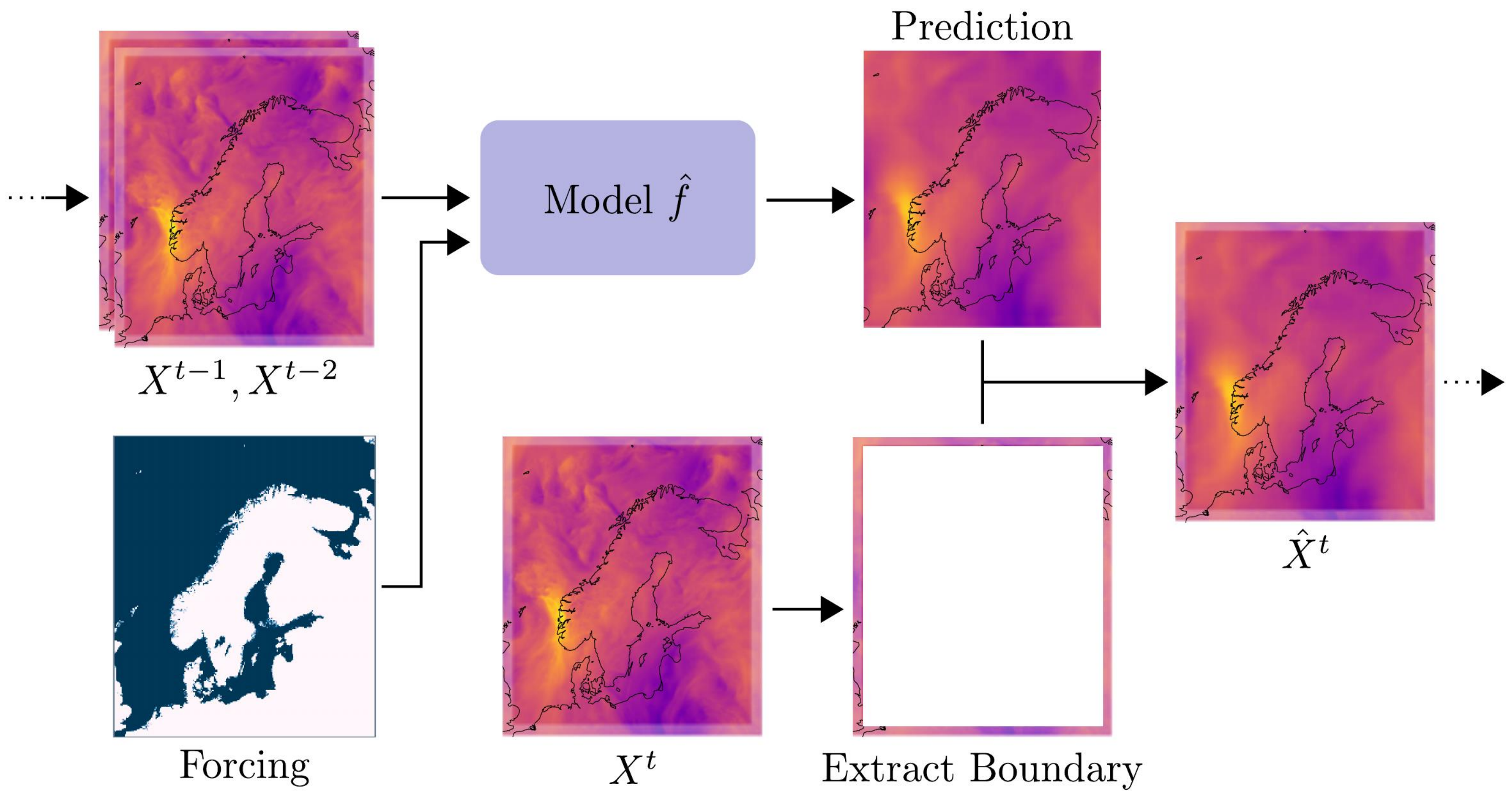
$$\boldsymbol{e}_{s\to r} \leftarrow \boldsymbol{e}_{s\to r} + \tilde{\boldsymbol{e}}_{s\to r}$$

$$\boldsymbol{v}_r \leftarrow \boldsymbol{v}_r + \mathrm{MLP}_V\left(\left[\boldsymbol{v}_r, \sum_{s:(s,r)\in\mathcal{E}} \tilde{\boldsymbol{e}}_{s\to r}\right]\right)$$

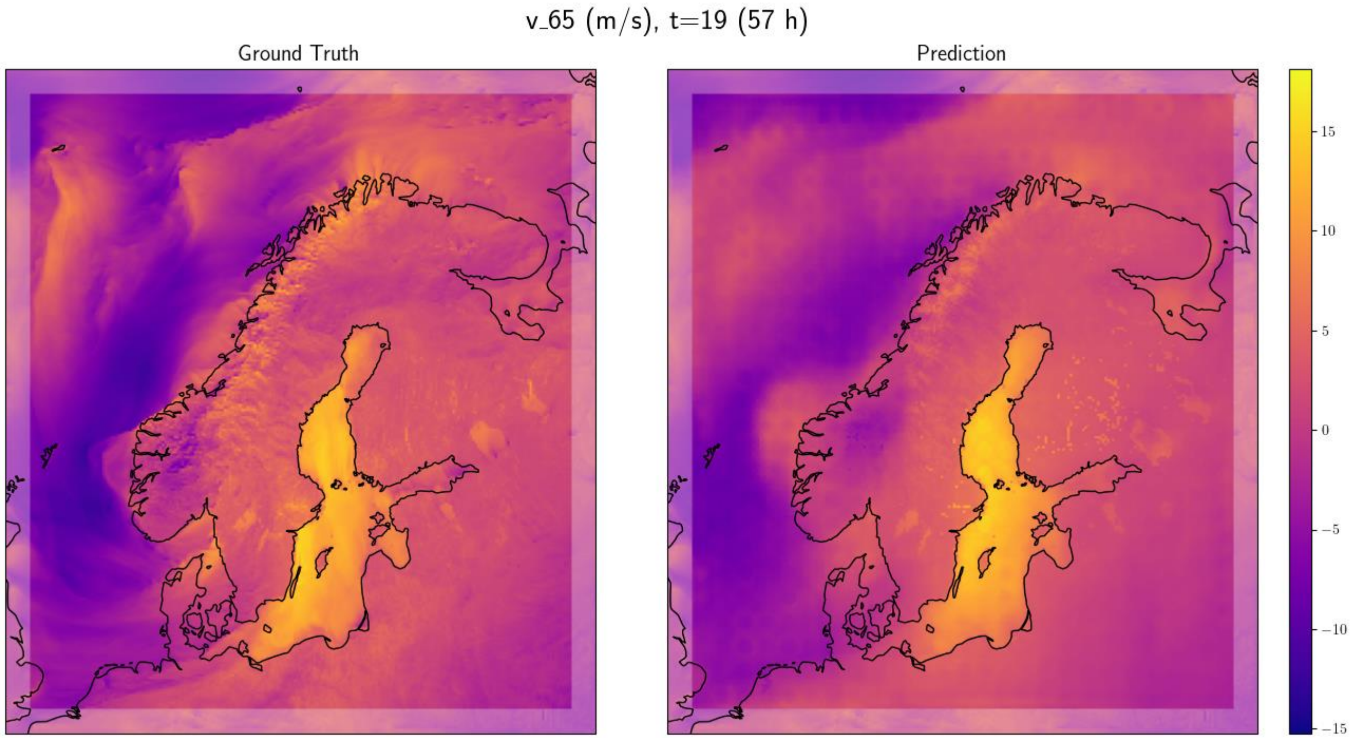
[1] P. Battaglia, et al. *Interaction Networks for Learning about Objects, Relations and Physics,* 2016. NeurIPS.

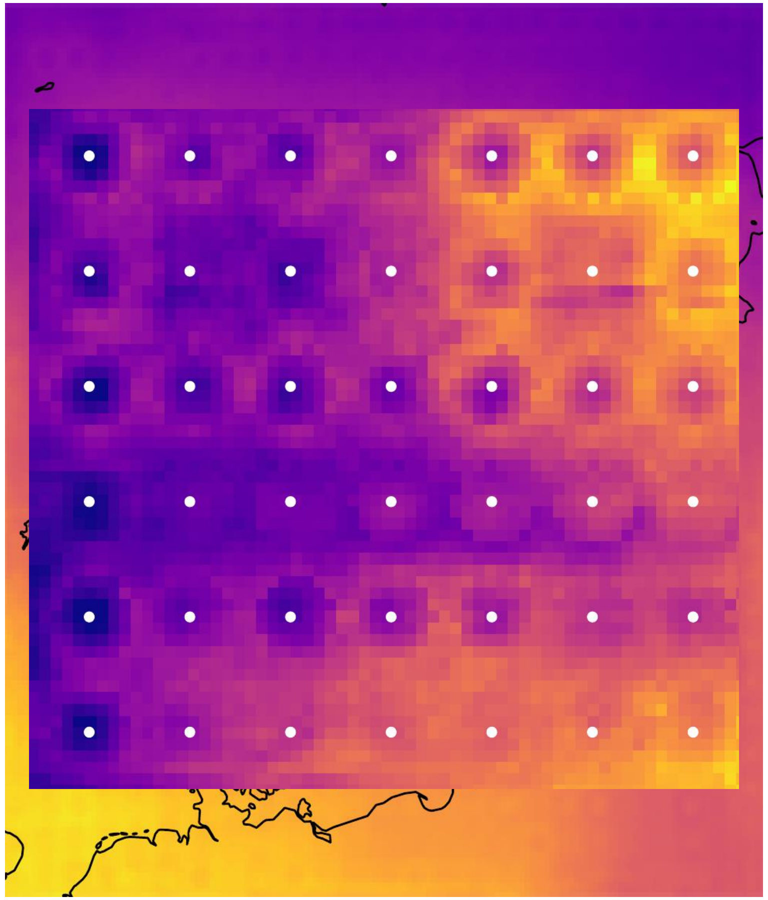
Back to our LAM models …

Boundary forcing



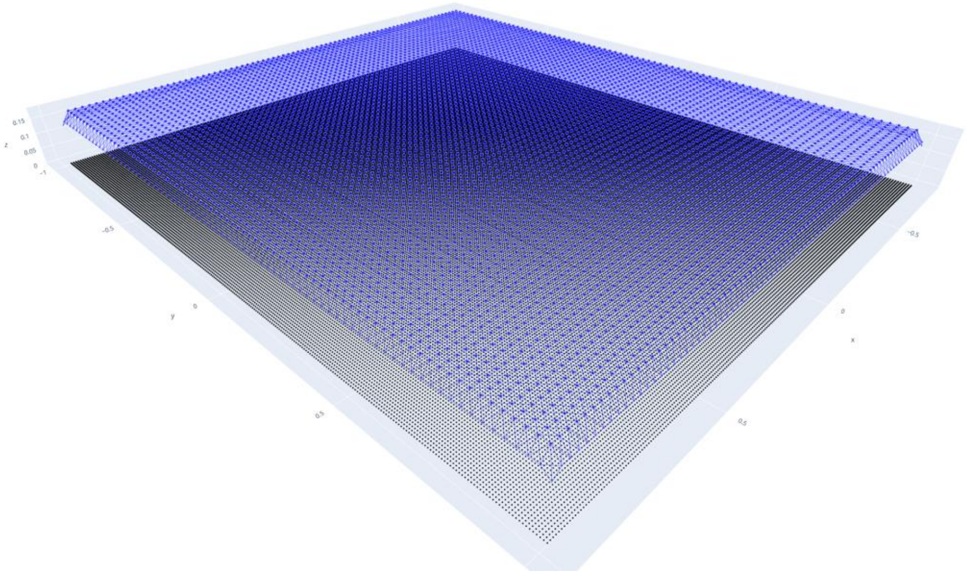
GC-LAM: First results and artefact issues





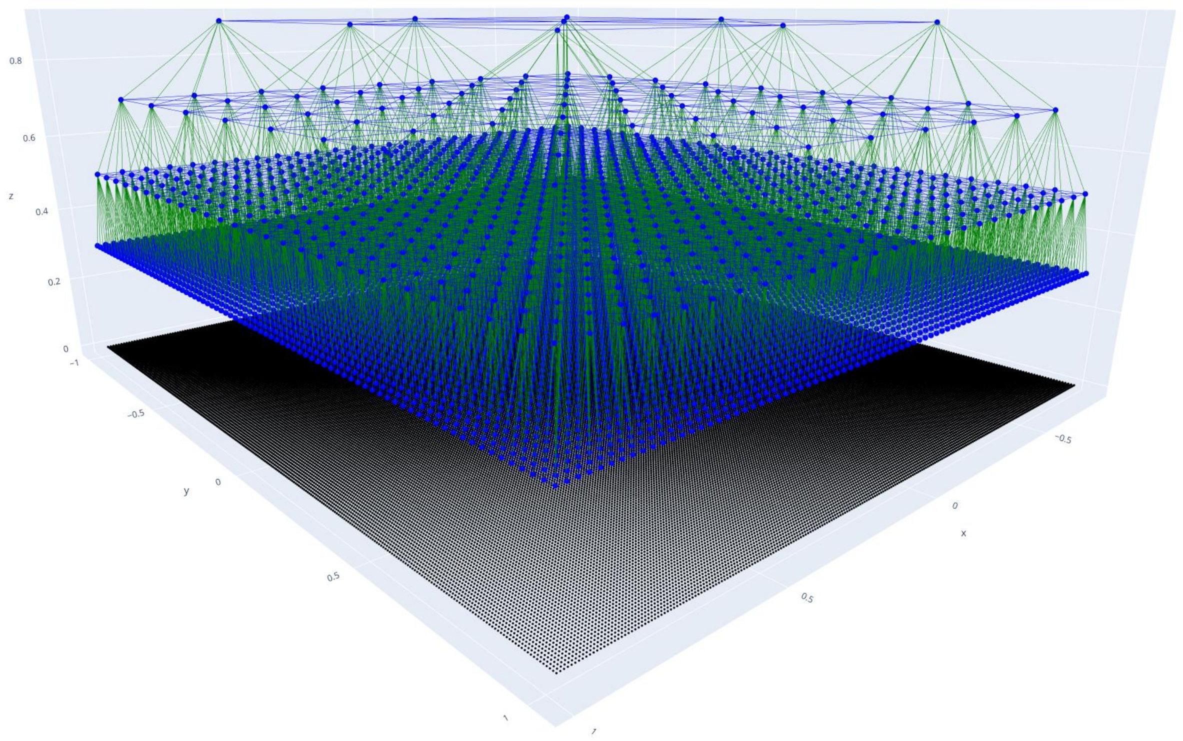
Dealing with the artefacts

- Remove multi-scale edges: 1L-LAM[1]



- Poor predictions :(

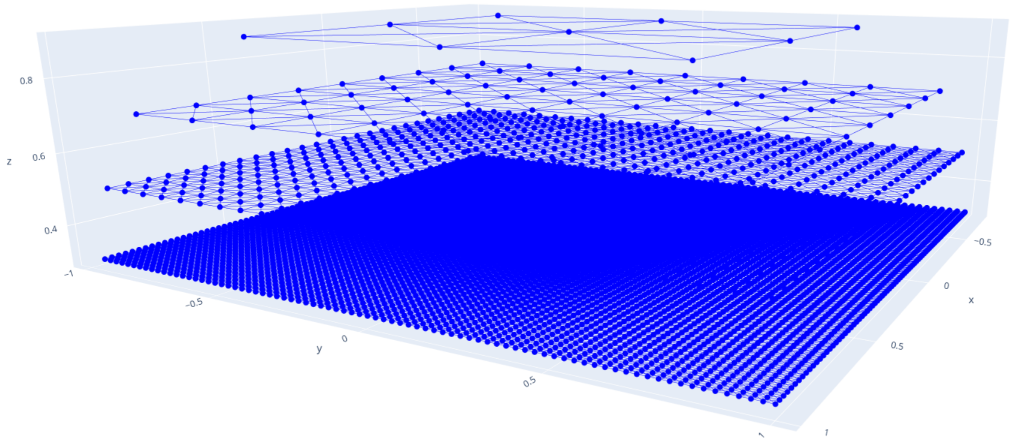
- Hierarchical graph model: Hi-LAM

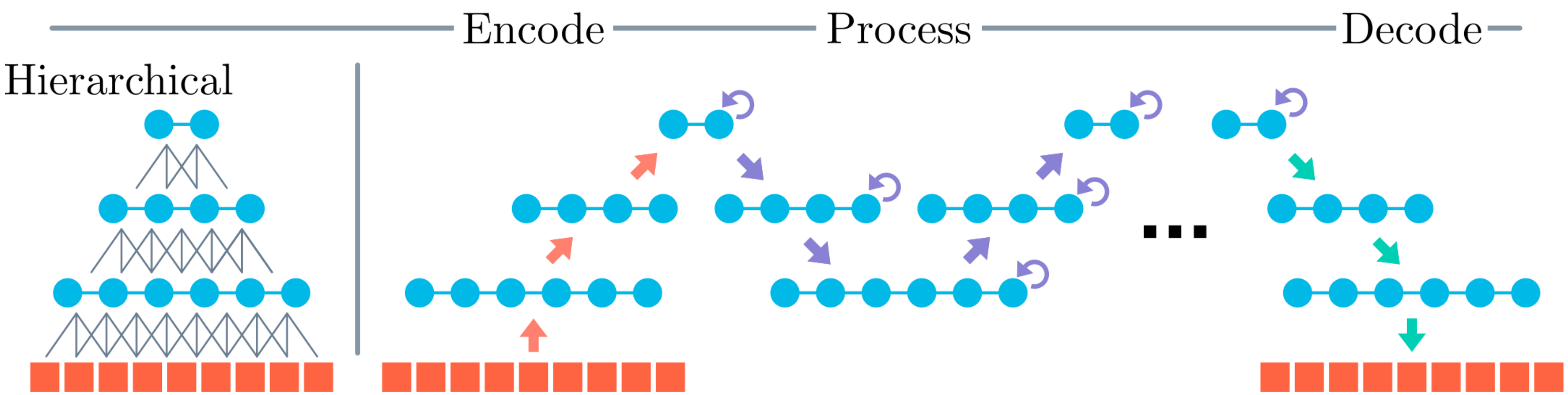


[1] A similar model to: R. Keisler. *Forecasting global weather with graph neural networks*, 2022.

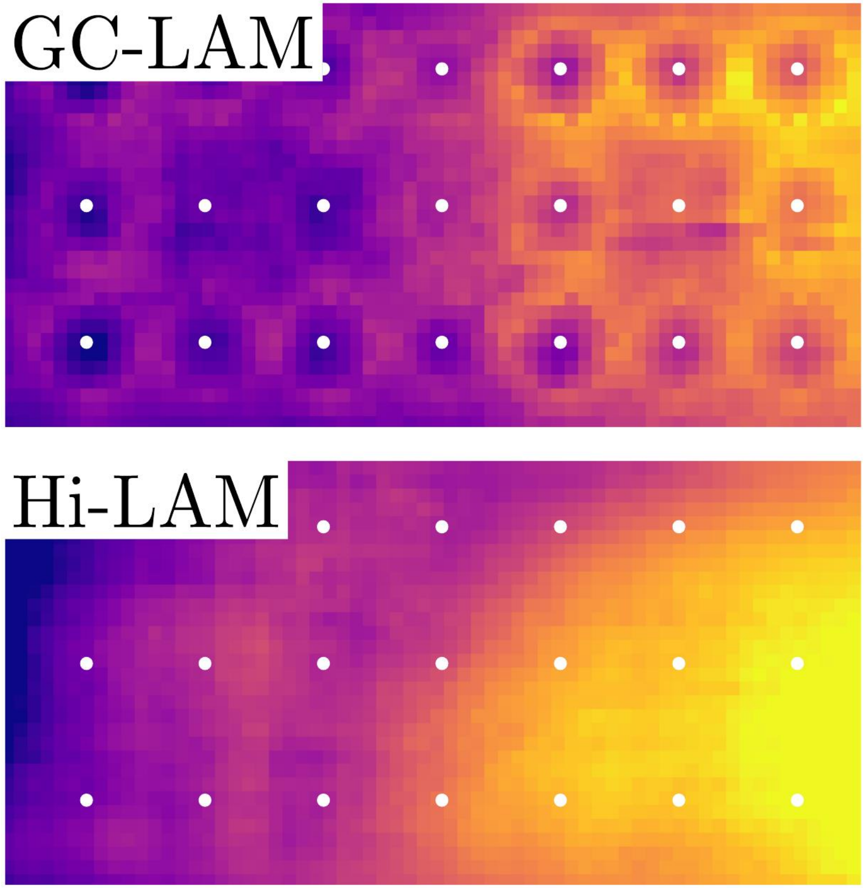
Hi-LAM: Hierarchical multi-scale graph

- 4 levels of nodes in mesh graph
  - Intra-level edges
  - Inter-level edges between adjacent levels
- Sequential GNN message passing up and down the hierarchy

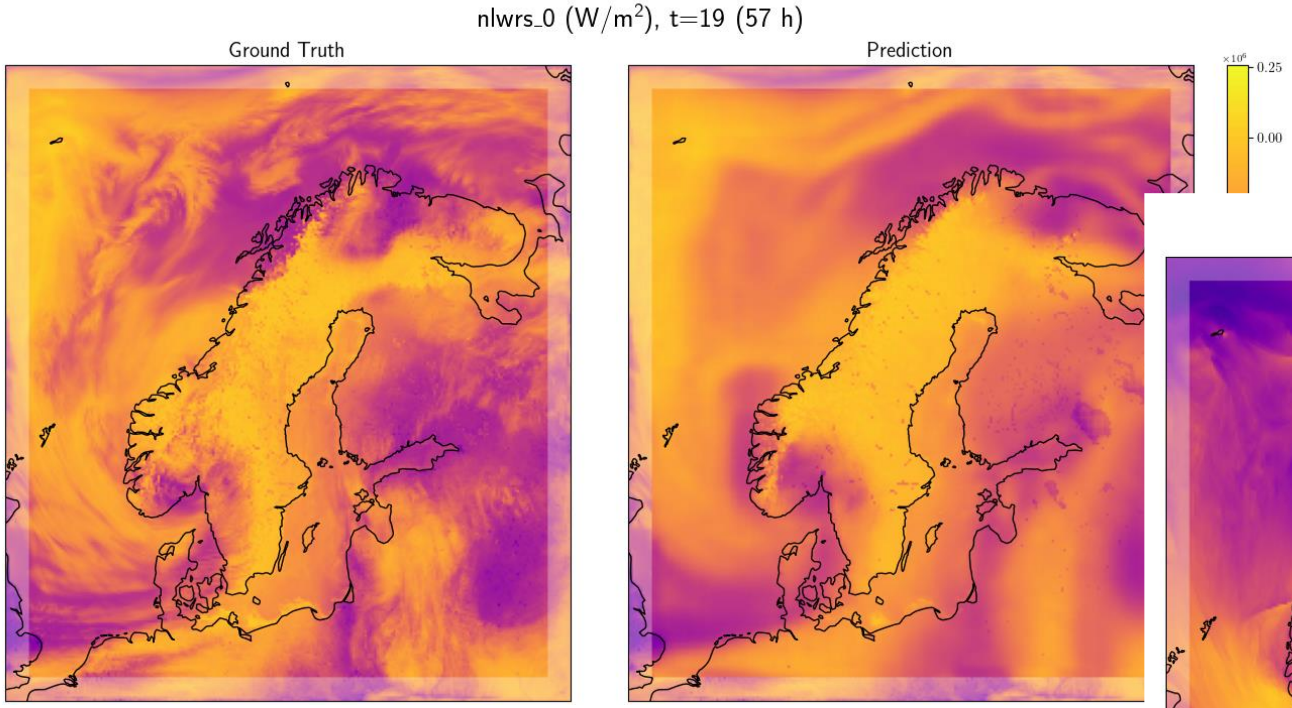




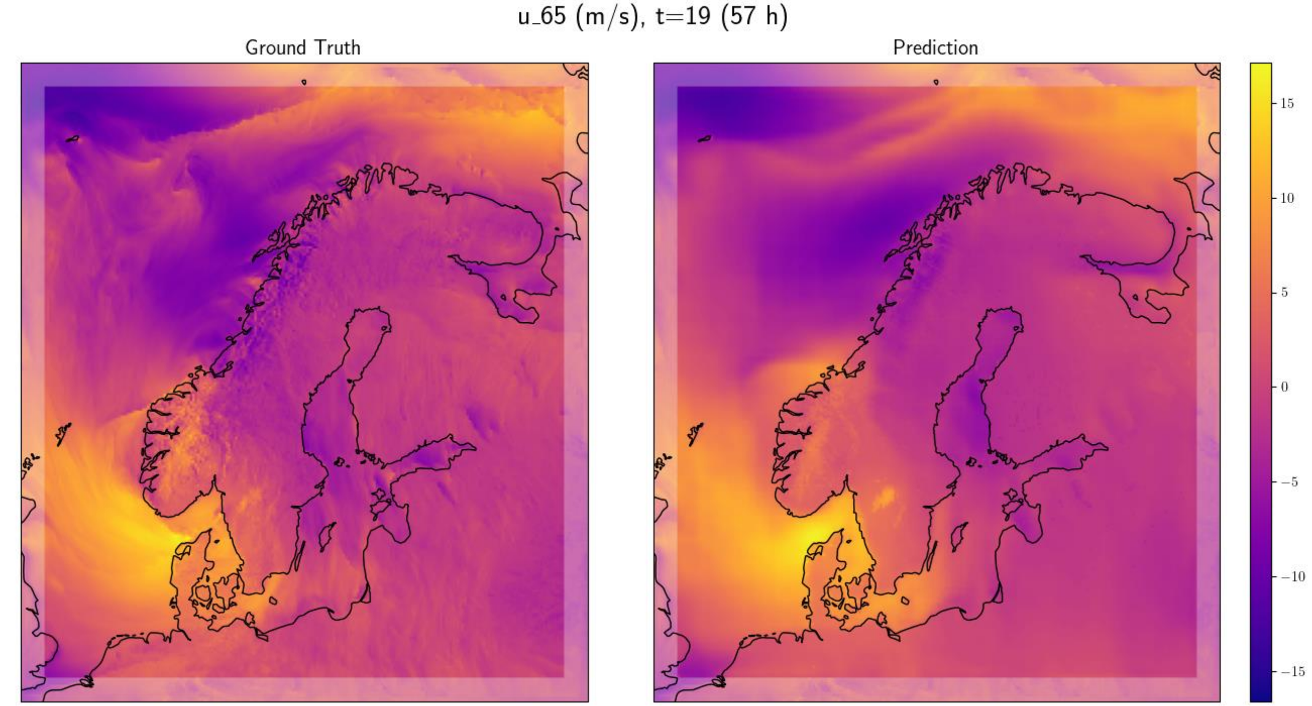
Results: Artefacts



Results: Example forecasts

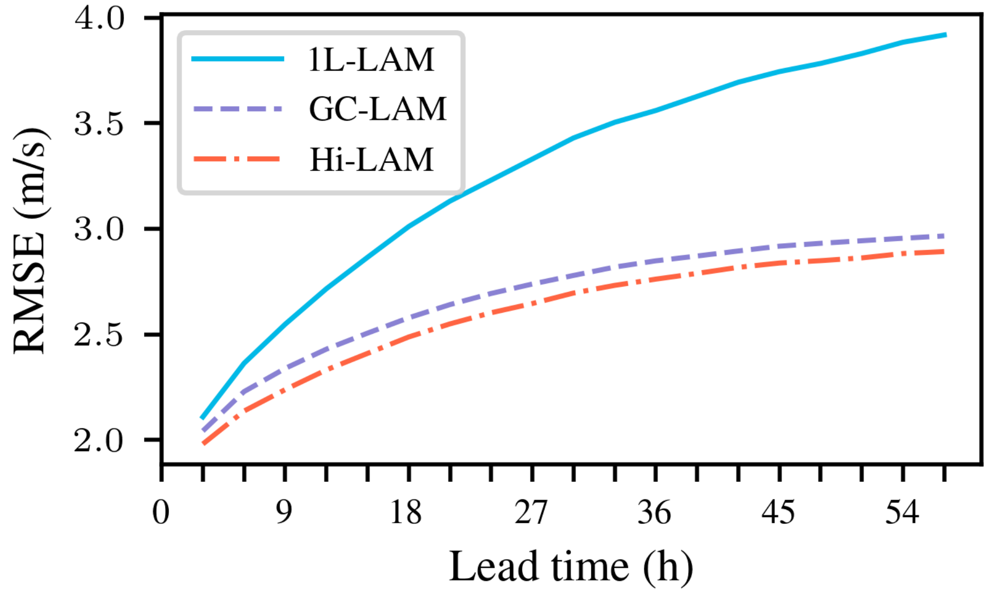


Net longwave radiation

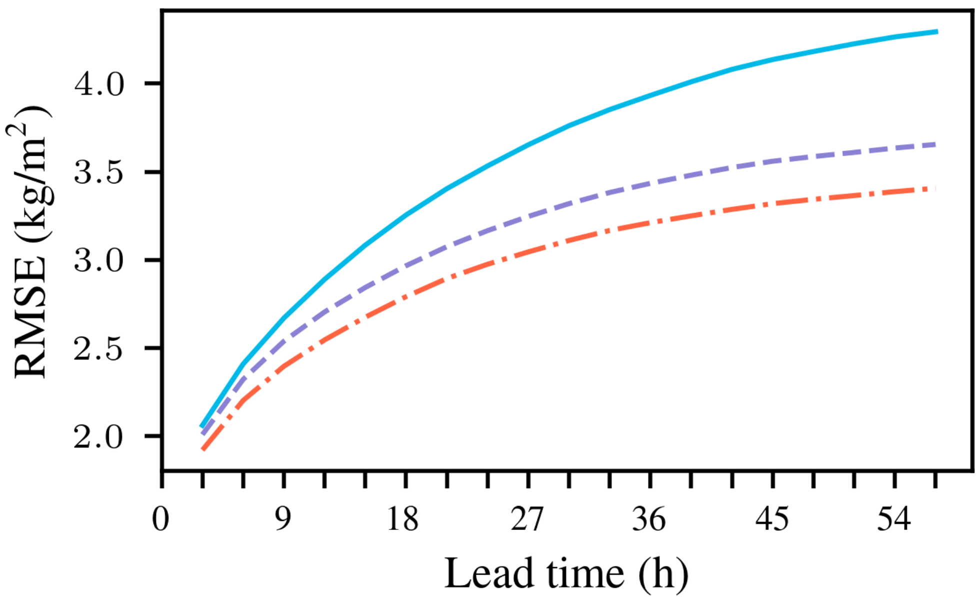


U-component of wind

Results: Errors over time



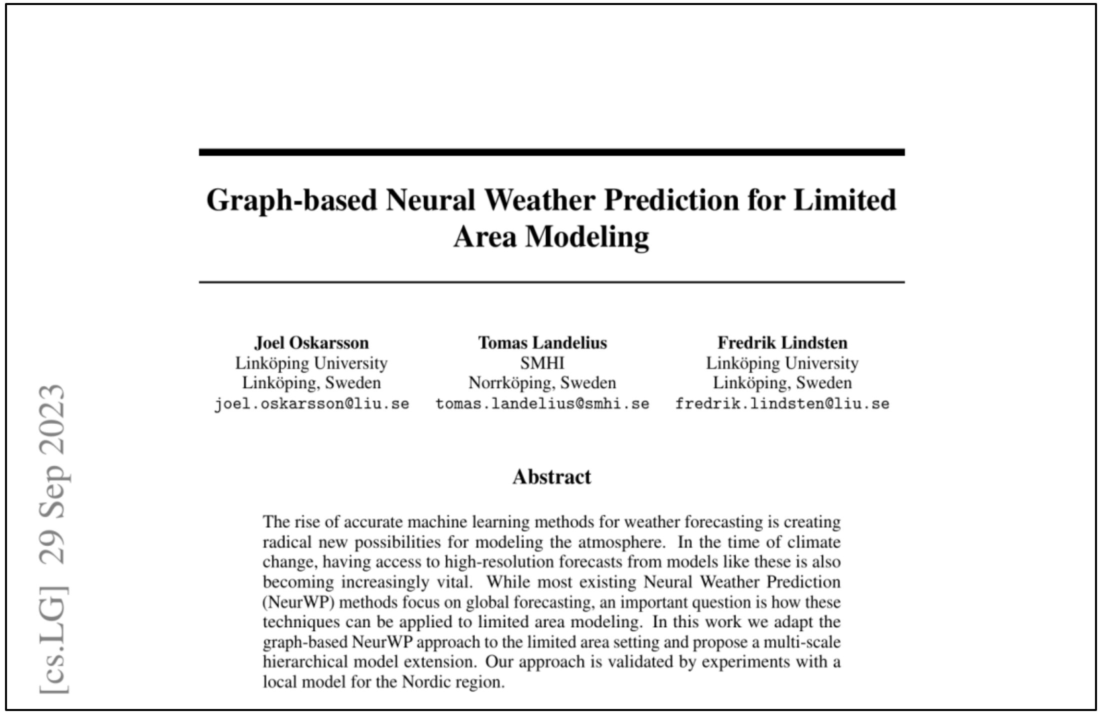
V-component of wind



Total water vapor column

Preprint available

- More details on our models, dataset and results
- https://arxiv.org/abs/2309.17370



[cs.LG] 29 Sep 2023

Graph-based Neural Weather Prediction for Limited Area Modeling

Joel Oskarsson
Linköping University
Linköping, Sweden
joel.oskarsson@liu.se

Tomas Landelius
SMHI
Norrköping, Sweden
tomas.landelius@smhi.se

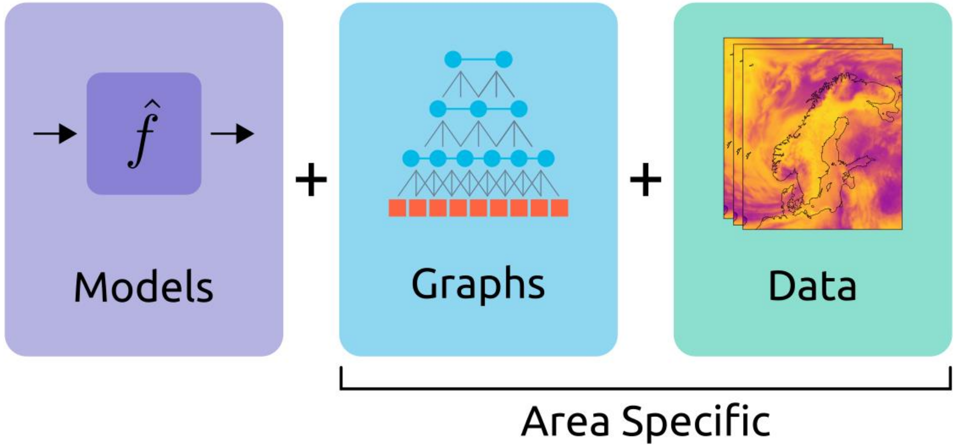
Fredrik Lindsten
Linköping University
Linköping, Sweden
fredrik.lindsten@liu.se

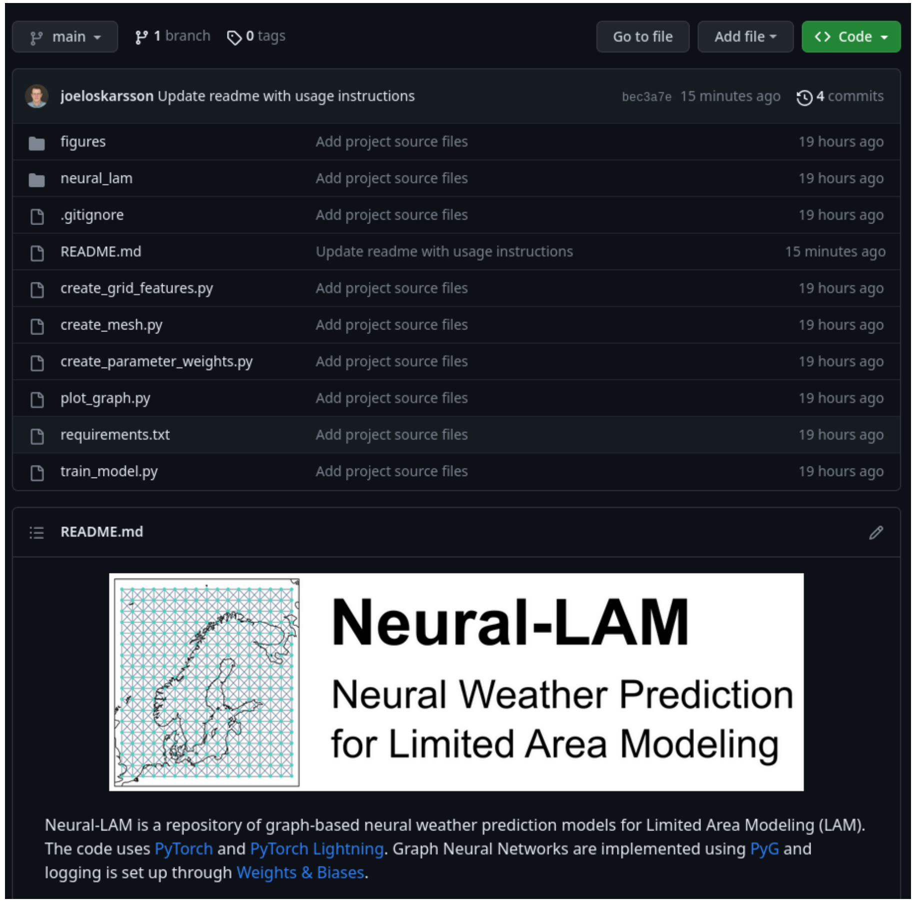
Abstract

The rise of accurate machine learning methods for weather forecasting is creating radical new possibilities for modeling the atmosphere. In the time of climate change, having access to high-resolution forecasts from models like these is also becoming increasingly vital. While most existing Neural Weather Prediction (NeurWP) methods focus on global forecasting, an important question is how these techniques can be applied to limited area modeling. In this work we adapt the graph-based NeurWP approach to the limited area setting and propose a multi-scale hierarchical model extension. Our approach is validated by experiments with a local model for the Nordic region.

Our implementation: Neural-LAM

- https://github.com/joeloskarsson/neural-lam
- PyTorch implementation
- Maintained and collaborative





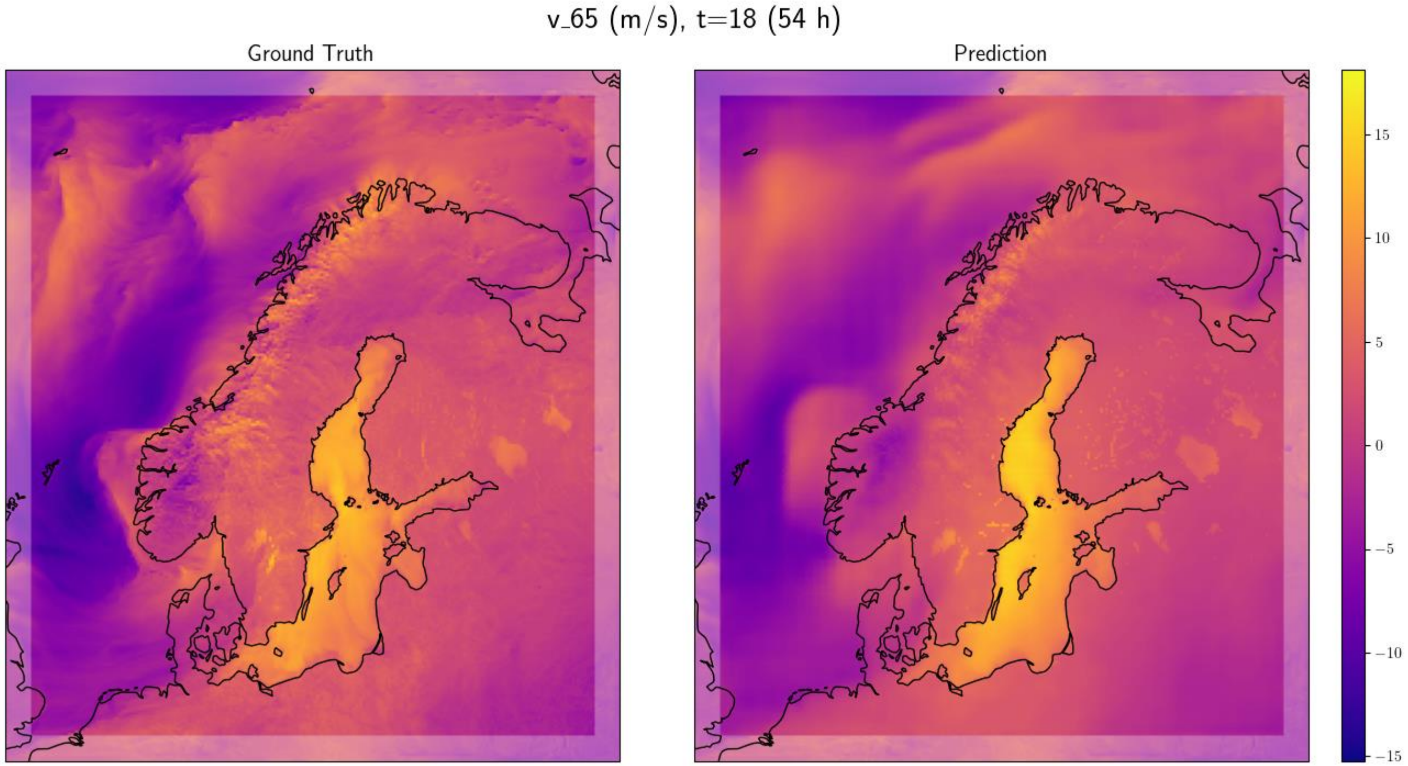
Outlook: Future Steps

The problem of over-smoothing

- Mean Squared Error loss

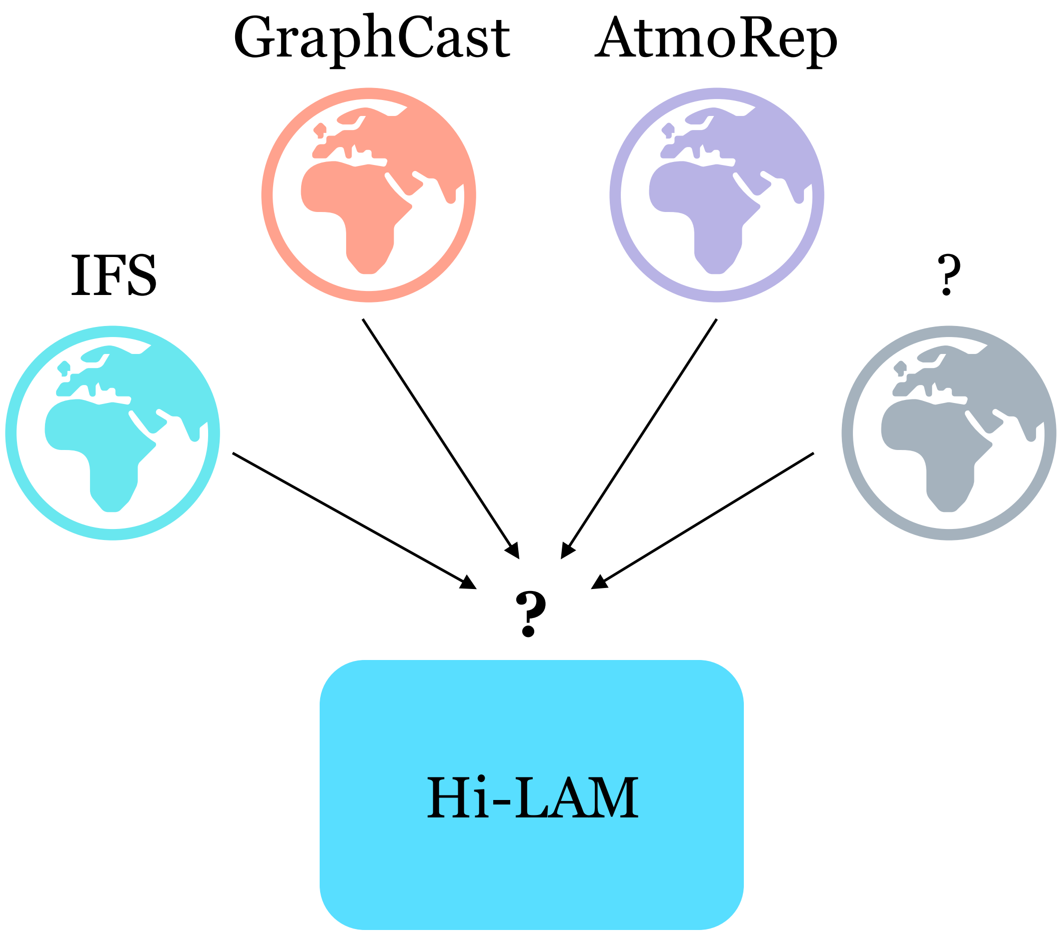
$$p\left(X^t \middle| X^{t-1}\right) = \mathcal{N}\left(X^t \middle| \hat{f}\left(X^{t-1}\right), \sigma^2 I\right)$$

- Worse at higher resolutions



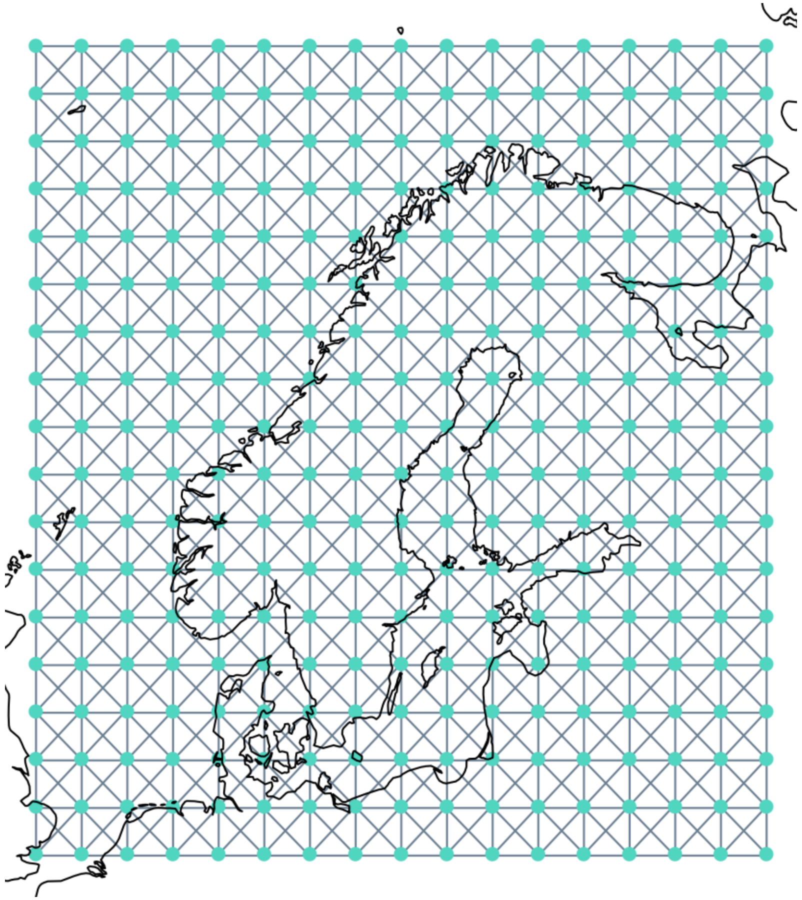
Neural LAM: Next steps

- Connecting global and LAM models
  - What forcing to use?
  - How to integrate?
  - How to train?



Summary

- Machine learning for NWP
- Graph-based LAM models
  - Adapting the graph
  - Boundary forcing
- Hi-LAM: Hierarchical graph



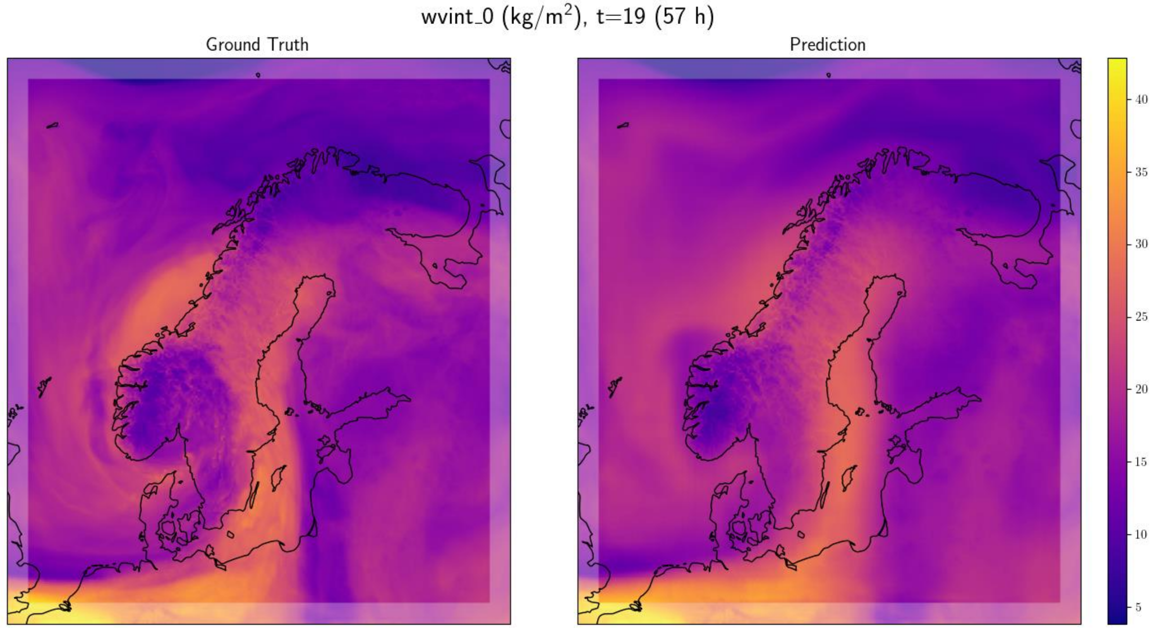
Graph-based Neural Weather Prediction for Limited Area Modeling

Joel Oskarsson
joel.oskarsson@liu.se
joeloskarsson.github.io

Tomas Landelius
tomas.landelius@smhi.se

Fredrik Lindsten
fredrik.lindsten@liu.se
lindsten.netlify.app

Paper
https://arxiv.org/abs/2309.17370



Code
https://github.com/joeloskarsson/neural-lam

LINKÖPING
UNIVERSITY