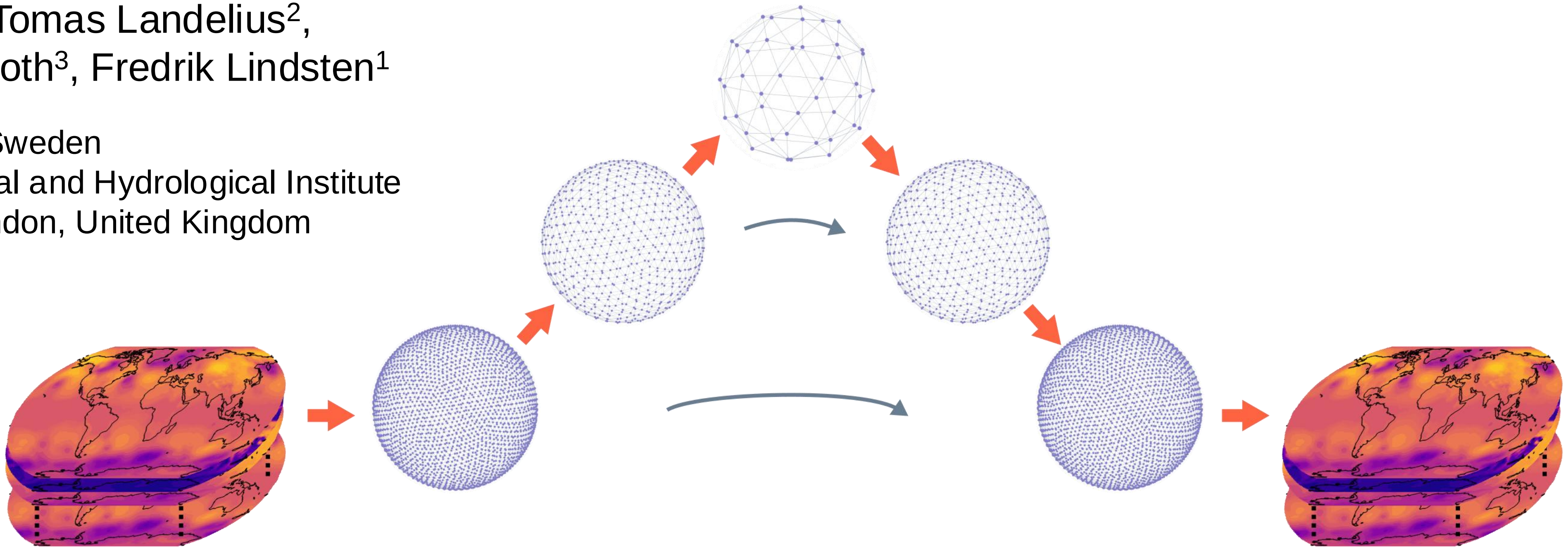
Probabilistic Weather Forecasting with Hierarchical Graph Neural Networks

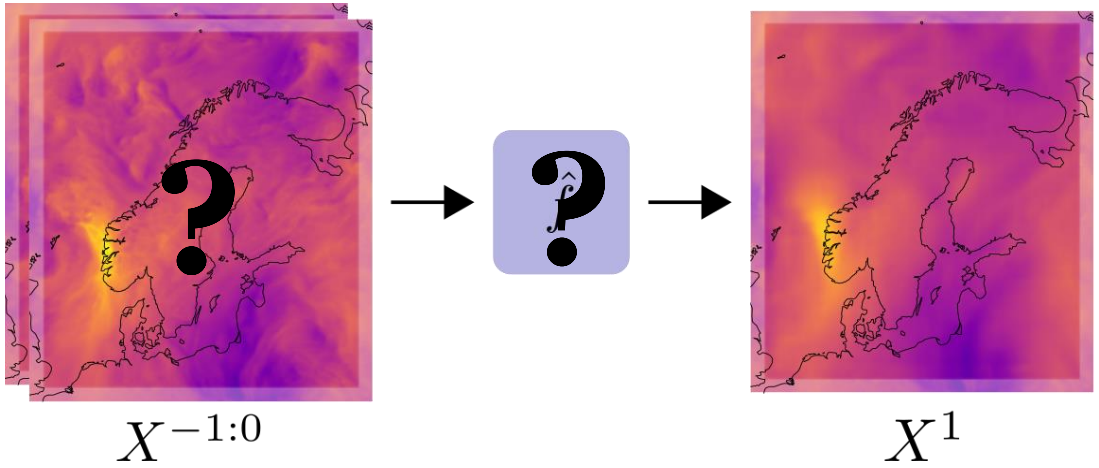
Joel Oskarsson[1], Tomas Landelius[2],
Marc Peter Deisenroth[3], Fredrik Lindsten[1]

[1]Linköping University, Sweden
[2]Swedish Meteorological and Hydrological Institute
[3]University College London, United Kingdom



LINKÖPING UNIVERSITY

Probabilistic weather forecasting



Currently:

- Deterministic models
- MSE loss

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

Want:

- Capture full distribution
  $p\left(X^{1:T}|X^{-1:0}, F^{1:T}\right)$
- Ensemble forecasting

Latent variable formulation

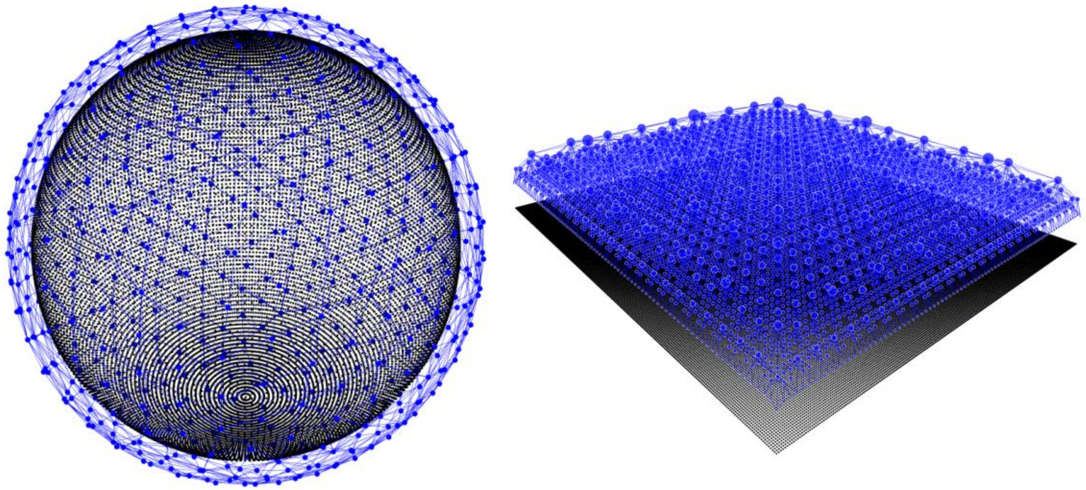
- Probabilistic + auto-regressive

$$p(X^t|X^{t-2:t-1}, F^t) = \int \underbrace{p(X^t|Z^t, X^{t-2:t-1}, F^t)}_{\text{Predictor}} \underbrace{p(Z^t|X^{t-2:t-1}, F^t)}_{\text{Latent map}} dZ^t$$

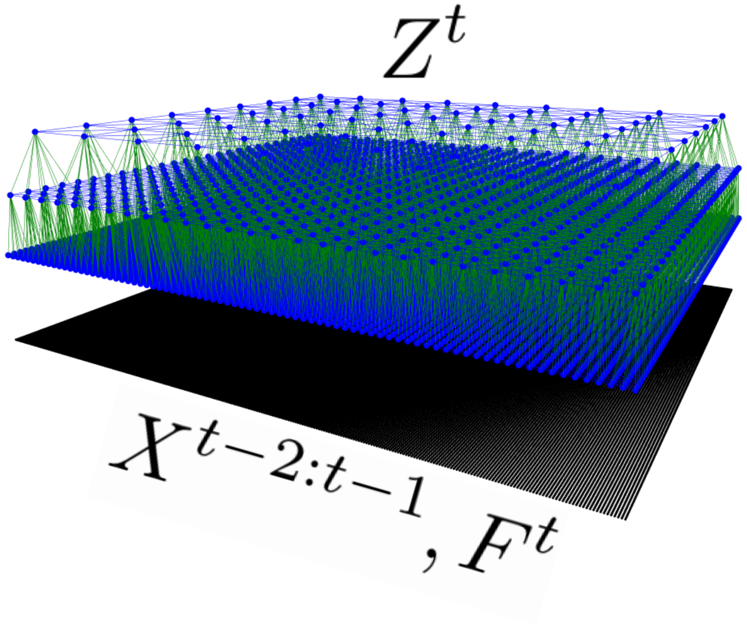
- Latent random variable Z^t
  - Captures uncertainty in single-step prediction

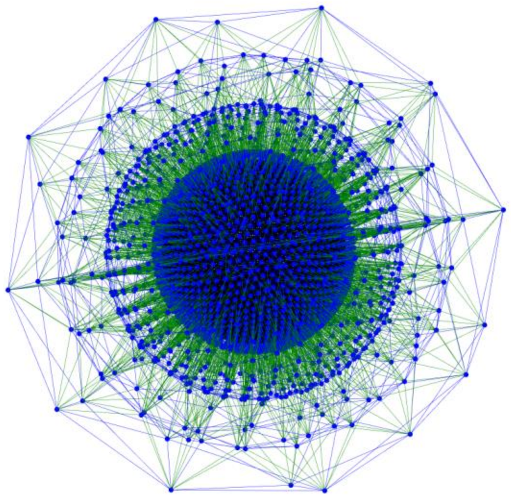
Graph-based weather forecasting

- Flexible framework for both global[1] and regional forecasting[2]



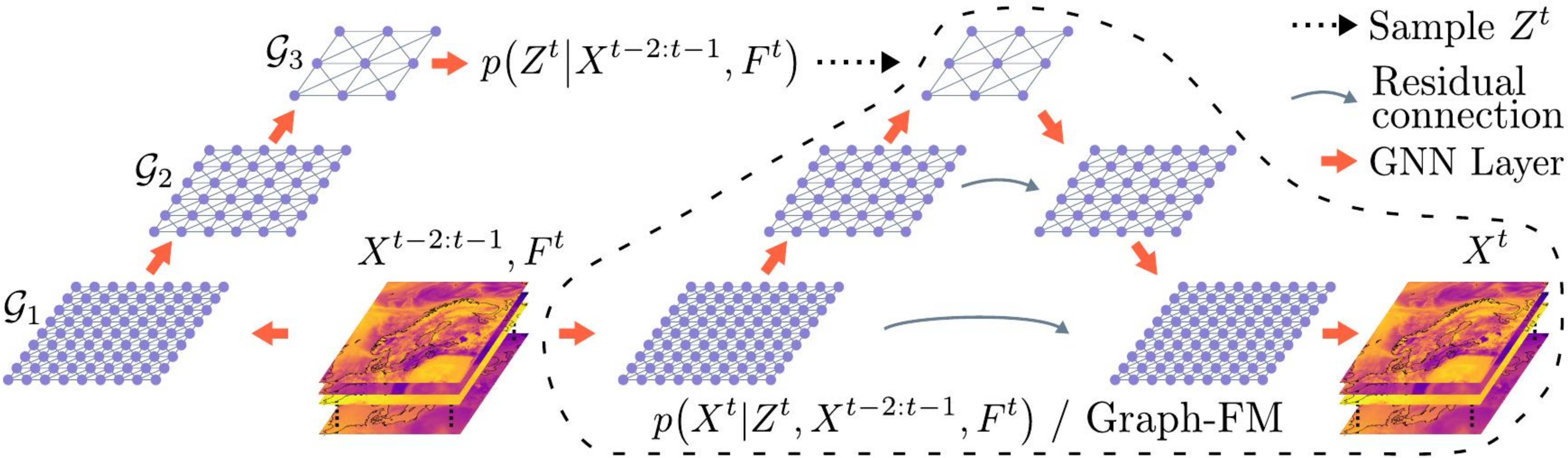
- Hierarchical graph construction





[1] Keisler, R. (2022). Forecasting global weather with graph neural networks. arXiv preprint. , Lam, R., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*.
[2] Oskarsson et al. (2023). Graph-based Neural Weather Prediction for Limited Area Modeling. *NeurIPS 2023 CCAI Workshop*.

Graph-EFM: Graph-based Ensemble Forecasting Model



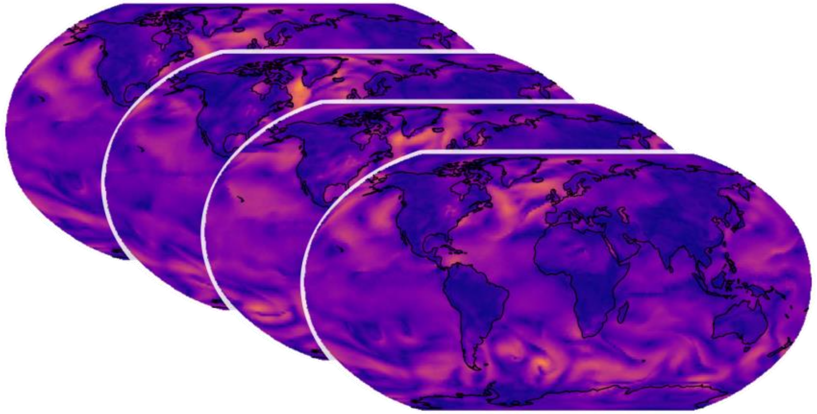
- Graph-FM: Deterministic model using hierarchical graph

Training and sampling

Training
- Maximize variational bound (ELBO)
- First on single-step prediction
- Finetuning on rollouts + using CRPS-based loss

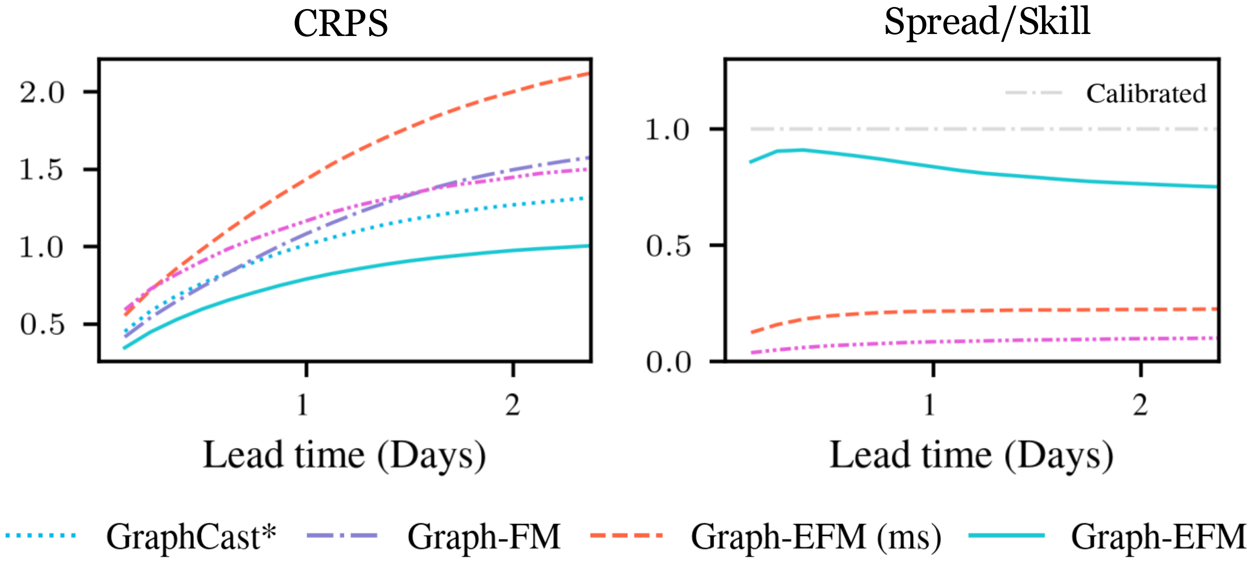
$$\mathcal{L} = \mathcal{L}_{\text{Var}} + \lambda_{\text{CRPS}} \mathcal{L}_{\text{CRPS}}$$

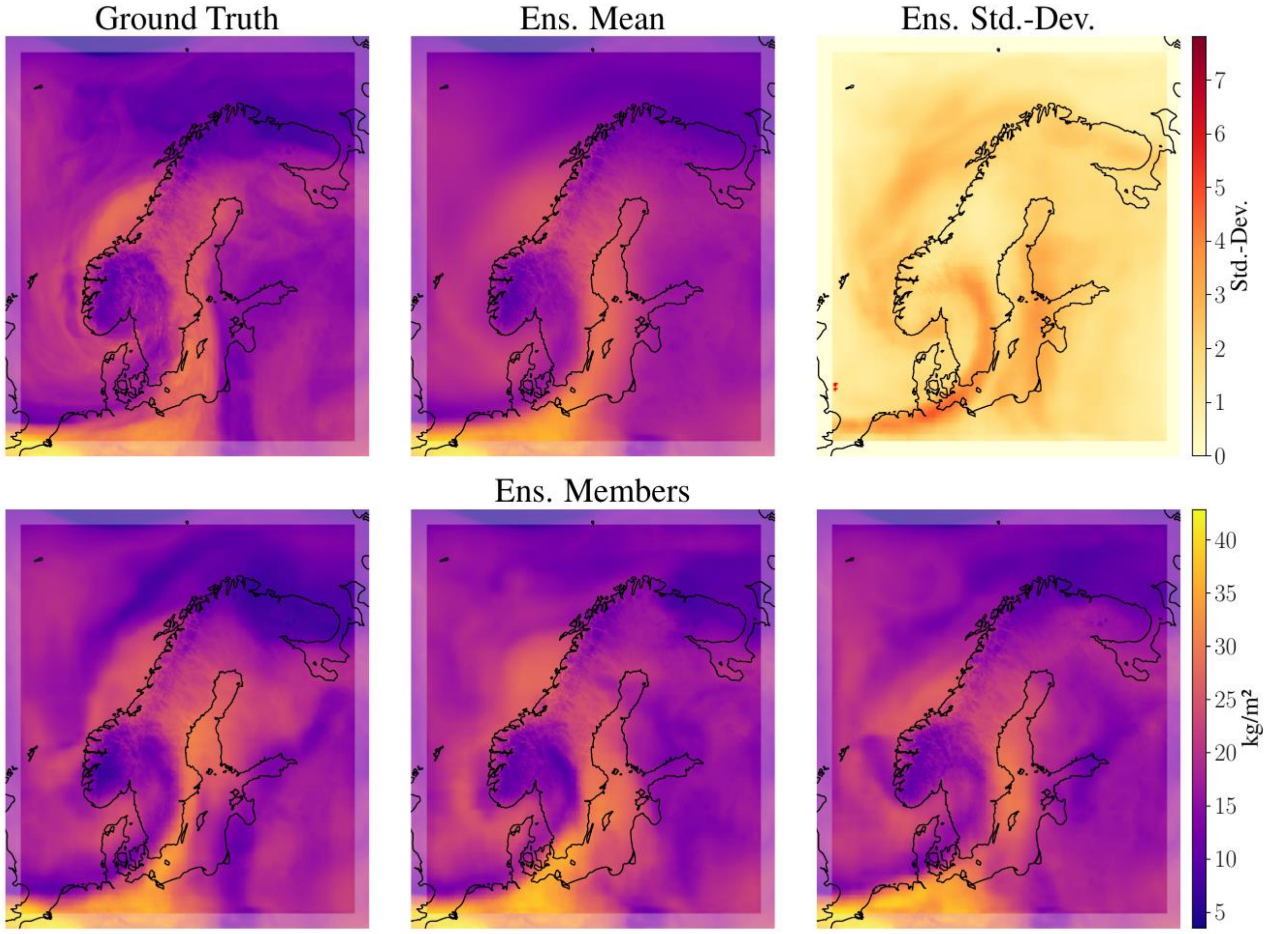
Sampling X^t
- Requires single forward-pass
- Contrast: Diffusion models



Results: Limited area modeling

- Surrogate model for forecasting Nordic region
    - Trained on dataset of 6000 forecasts
    - 57 h forecasts with 3 h time steps
    - 17 variables
- Boundary forcing

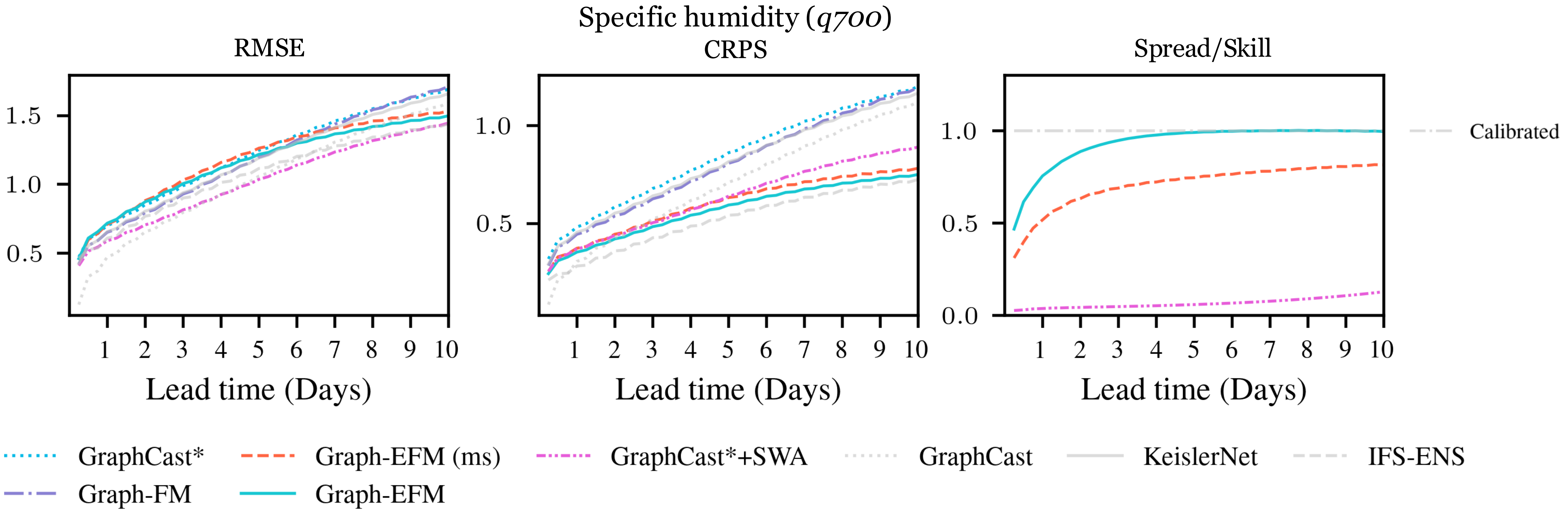




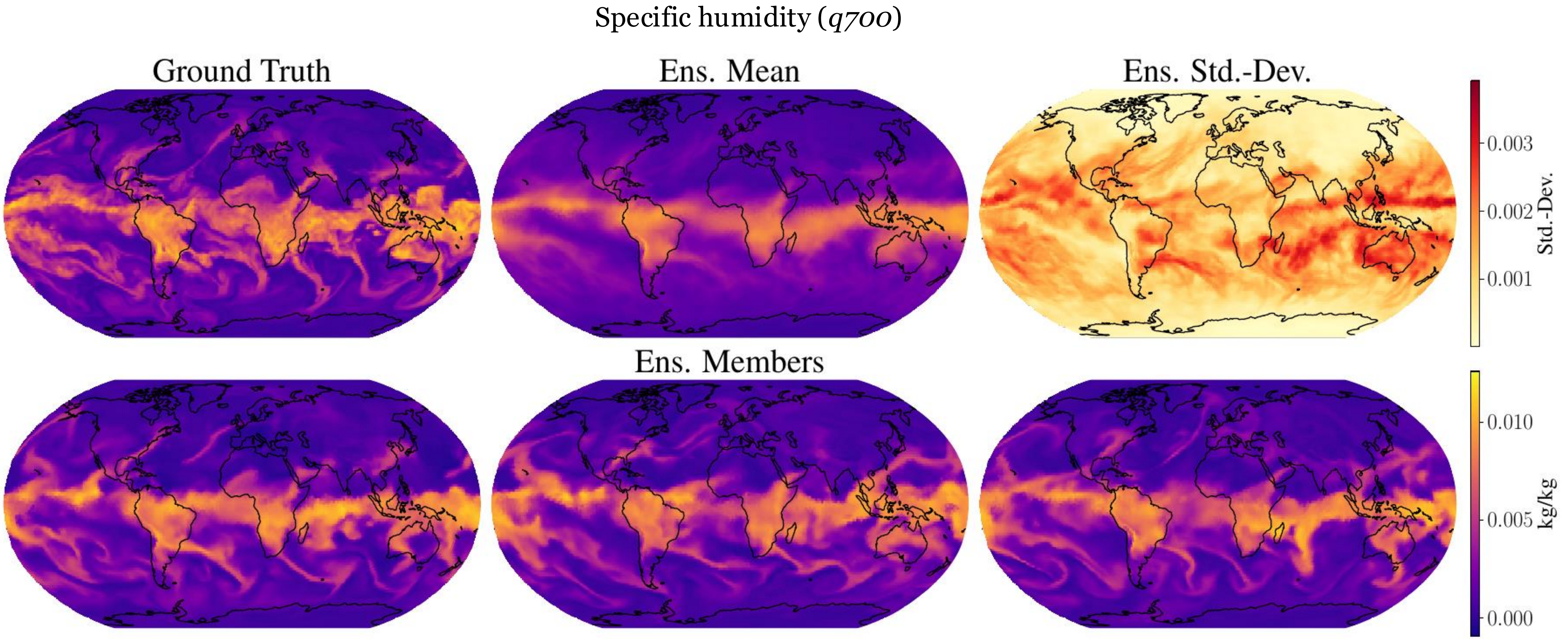
Water vapor (*wvint*)

Results: Global forecasting

- ERA5 on 1.5° grid
- 83 variables (5 surface + 6 atmospheric × 13 pressure levels)
- 10 day forecasts with 6 h steps



Results: Global forecasting

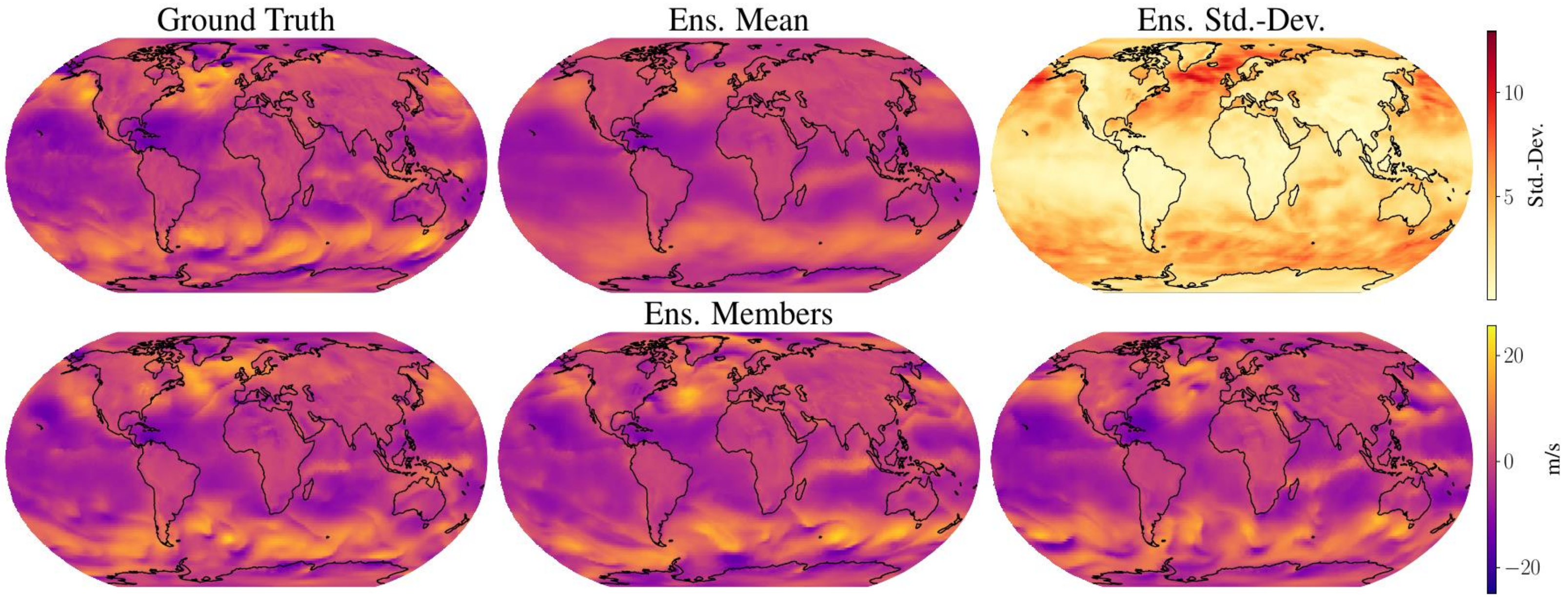


Probabilistic Weather Forecasting with Hierarchical Graph Neural Networks

Joel Oskarsson Tomas Landelius Marc Peter Deisenroth Fredrik Lindsten

Paper

https://arxiv.org/abs/2406.04759



Wind (10u)

Code

https://github.com/mllam/neural-lam/tree/prob_model_global