3D Radar Echo Motion Estimation Using Deep Learning: A Case Study in Slovakia

5th ECMWF-ESA Machine Learning Workshop

Peter Pavlík
PhD Student / Research Assistant

Kempelen Institute of Intelligent Technologies

> As the climate warms, the consequent moistening of the atmosphere increases extreme precipitation.
>
> – W. Zhang et al. (2024)

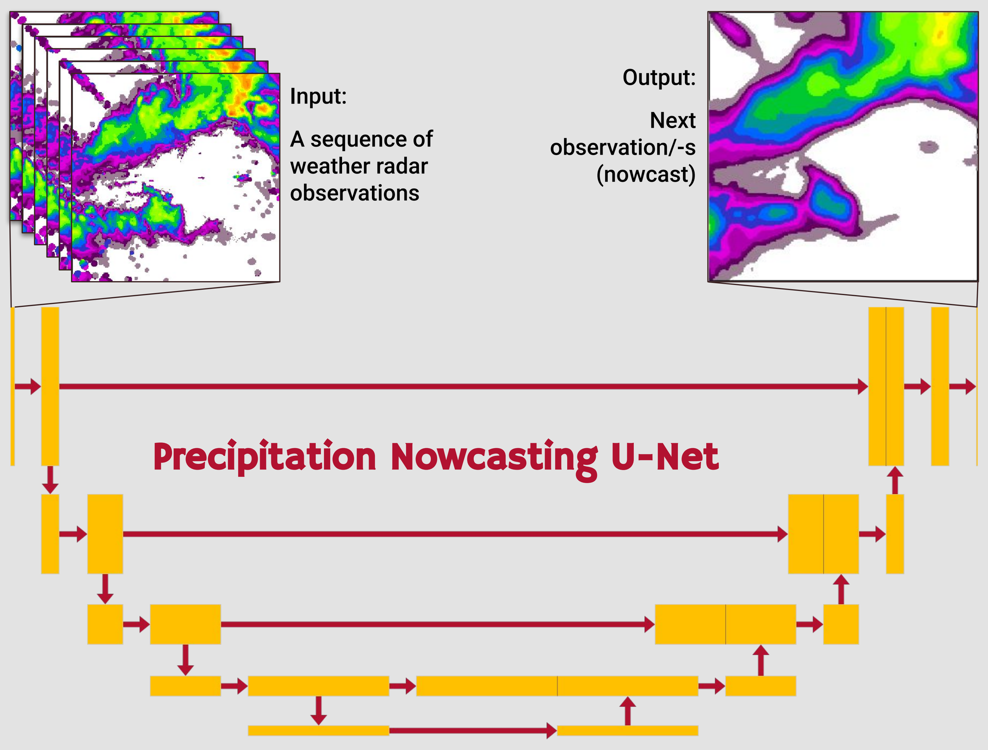
Neural Networks in Precipitation Nowcasting

Precipitation Nowcasting

- Nowcasting is a kind of forecasting:
  - with high local detail
  - over a period from the present to six hours ahead
  - including a detailed description of the present weather

Usage of Convolutional Neural Networks

- Weather radar stations observe precipitation particle reflectivity
- The precipitation maps can be processed by common computer vision models – nowcasting is analogous to video prediction

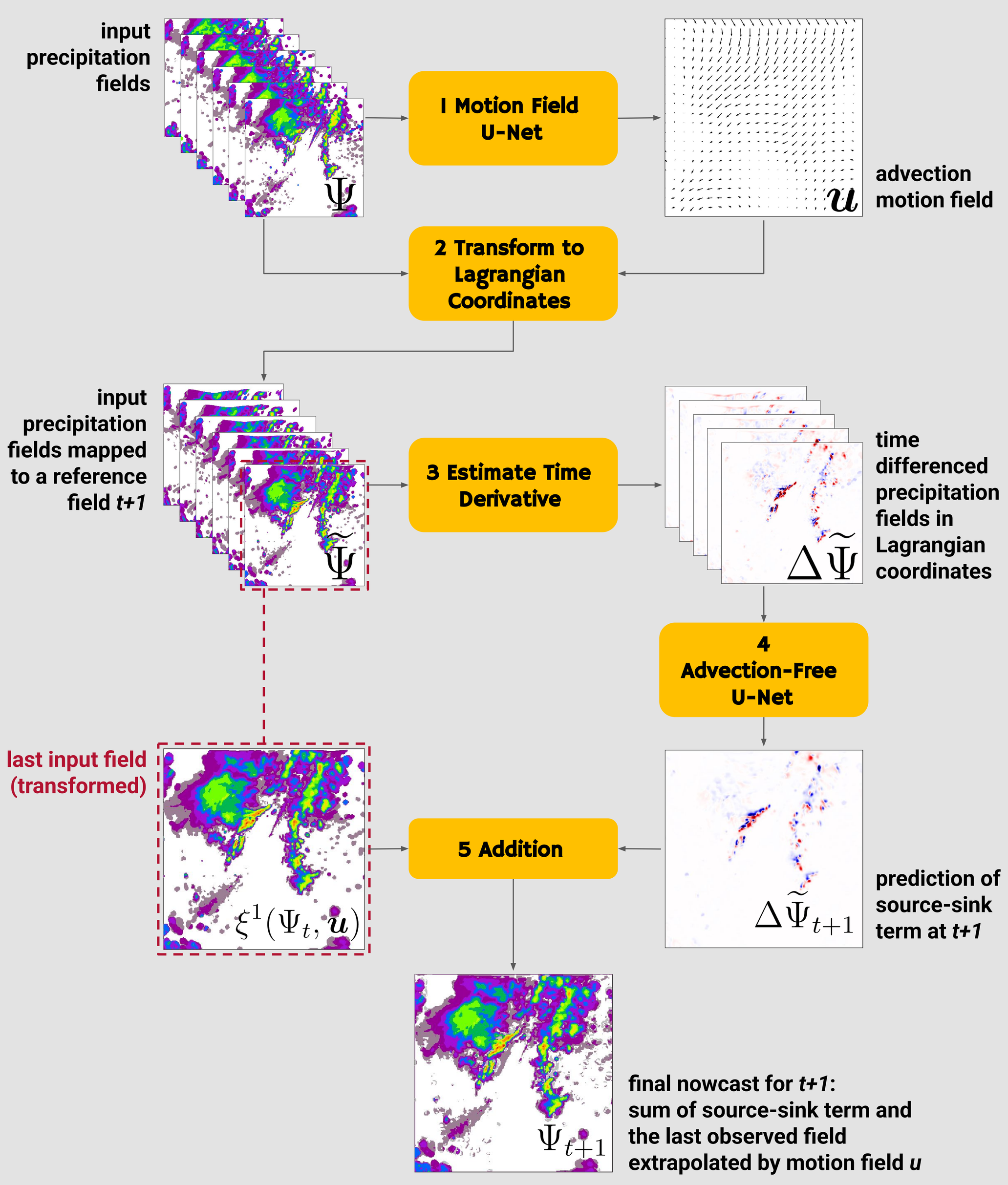


LUPIN: Lagrangian Convolutional Neural Network

Our Previously Published Work

Machine Learning + Physics-Informed Domain Knowledge

- Draws from traditional extrapolation-based nowcasting methods
- Motion Field U-Net learns to produce mesoscale advection fields thanks to differentiable semi-Lagrangian extrapolation operator
- Advection-free U-Net captures the growth and decay of precipitation



Why Volumetric Motion?

- Existing approaches mostly operate on two-dimensional radar composites
  - **Assumption:** Precipitation advection is sufficiently represented in a vertically collapsed space – *is it?*

Volumetric Deep Learning Approaches Are Under-Explored

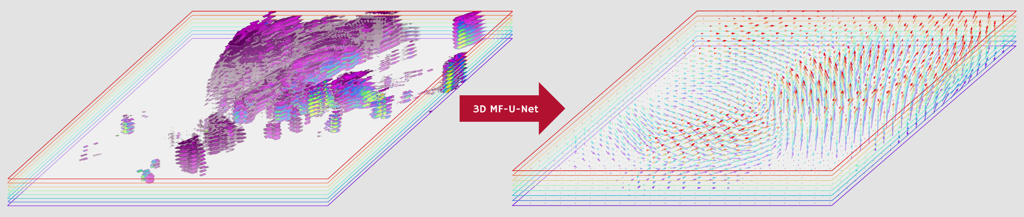
- CNNs have been applied to volumetric radar data and shown improvements in nowcast skill
- How much explicitly modeled volumetric motion fields can help general nowcasting performance remains unclear

3D Motion Field U-Net: Altitude-wise Volumetric Motion Field Estimator

Preprint Available

Model Architecture

- 3D U-Net that processes the 4D data (3D volume + time) as independent altitude-wise 3D volumes (horizontal 2D slices + time)
- Outputs altitude-wise horizontal motion fields that are stacked into a pseudo 3D motion field (no vertical component)
- Processes all altitudes in parallel as a single batch – rapid inference



Training

- Able to train thanks to fully differentiable semi-Lagrangian extrapolation operator (implemented in PyTorch)
- A few modifications for learning stable and smooth motion fields
  - Sequence-consistent motion estimation – a single motion field for the whole input sequence

$$\mathcal{L}_{\mathrm{MP}} = \sum_{t=1}^{n+m-1} \mathcal{C}(\xi^1(\Phi_t, \boldsymbol{u}), \Phi_{t+1}), \text{ where } \Phi = \{\Psi_1, \dots, \Psi_n, \hat{\Psi}_{n+1}, \dots, \hat{\Psi}_{n+m}\}$$

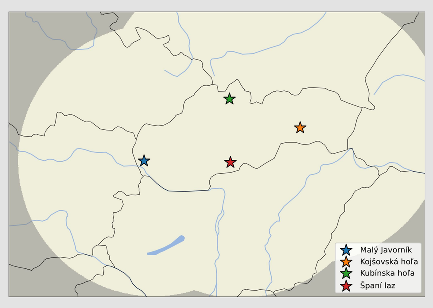
  - Multi-scale pooling – extrapolation loss is evaluated across multiple spatial scales and averaged (1, 2, … , 64, 128 km)
  - Physics-informed regularization – penalizes violations of the conservation of mass (discourage sources and sinks in the estimated motion fields)

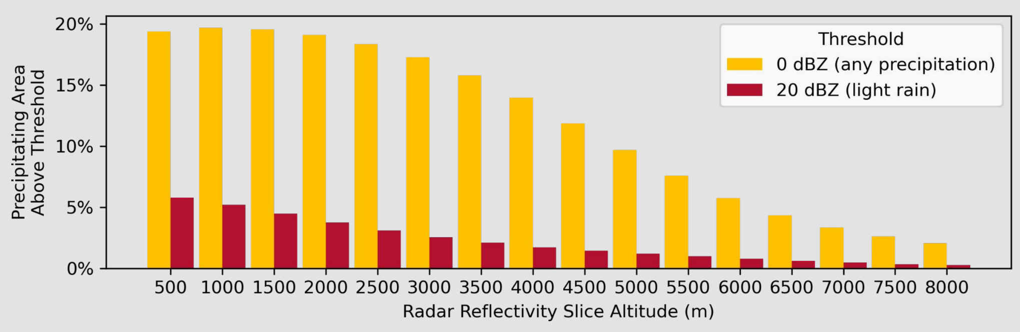
$$\mathcal{L}_{\mathrm{PI}} = |\nabla \cdot \boldsymbol{u}|, \text{ where } \nabla \cdot \boldsymbol{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y}$$

Slovak Volumetric Radar Dataset

Dataset Info

- Provided by Slovak Hydrometeorological Institute (SHMÚ)
- Data from four C-band Doppler radar stations located at Malý Javorník, Kojšovská hoľa, Kubínska hoľa, and Španí laz
- Spans the period of four years from 12 June 2018 to 22 August 2022
- Horizontal resolution of 1 km × 1 km, images of 517×755 pixels
- Vertical resolution of 16 altitude levels from 500 to 8000 m above ground (pooled to 8 for training the model)

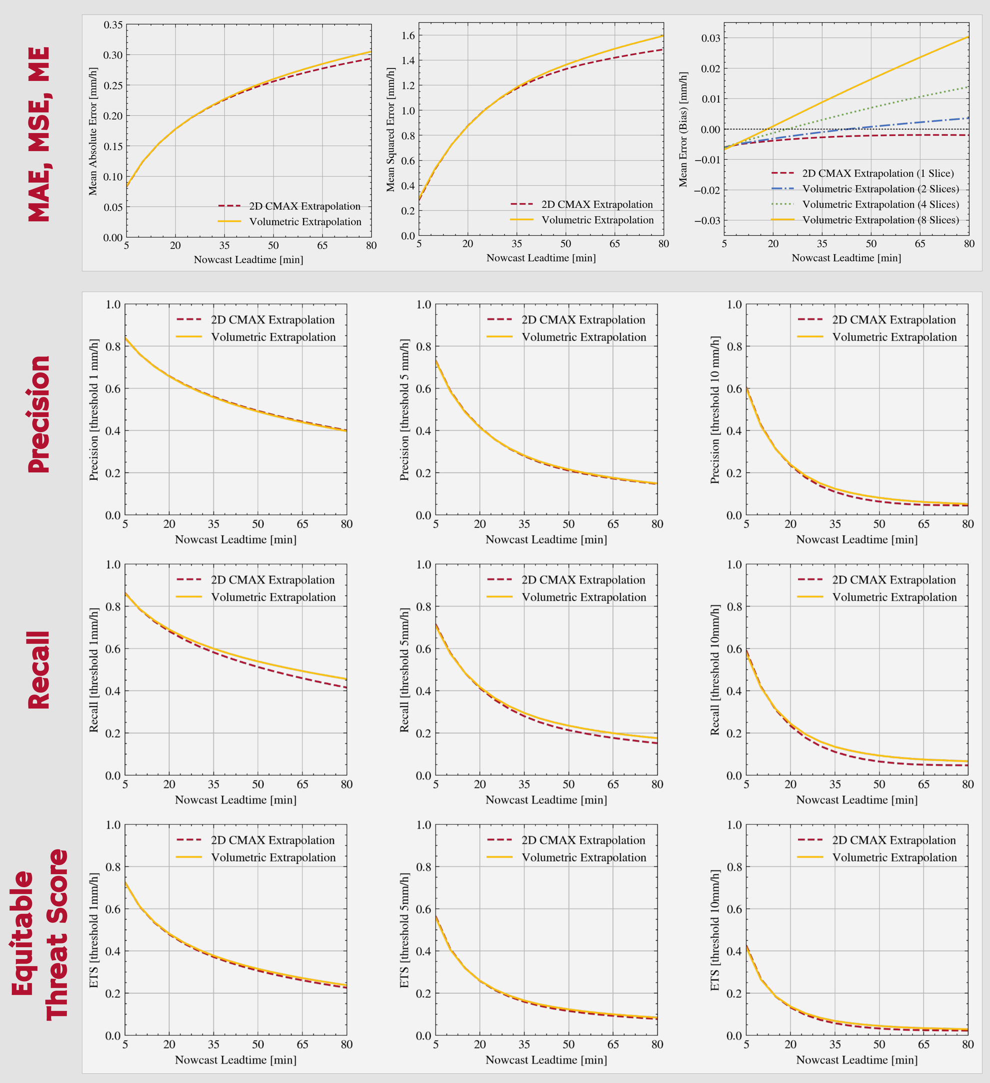




Nowcasting Performance: 2D vs. 3D Extrapolation

Comparison with a 2D Model

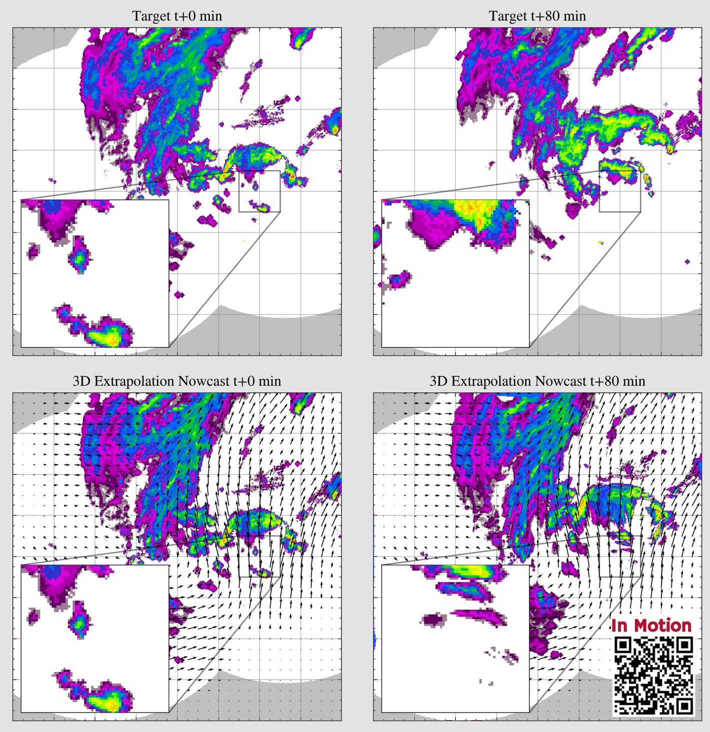
- Architecturally identical 2D CMAX baseline (3D MF U-Net) to compare against the proposed volumetric model
- The forecasts produced by the volumetric model are subsequently collapsed into two-dimensional CMAX for fair evaluation



- Volumetric approach does not reduce MAE and MSE, growing positive bias
- Improvement in Recall and ETS without a substantial loss in Precision (but mostly gained from non-physical cell-splitting artifacts)

Non-physical Artifacts

- Artificial cell-splitting increases the spatial extent of the predicted precipitation field and contributes to a growing positive bias and better recall
- Consequence of evaluation methodology – estimating altitude-wise motion fields independently, combined with the evaluation on vertically pooled products



Inter-altitude Motion Field Correlation Analysis

Full Dataset Correlations

- Investigate the degree of similarity between motion fields at different altitude levels
- The analysis focuses not on predictive performance but on the variability of the estimated motion fields, so the whole dataset was analyzed, including training data
- Pearson correlations were computed independently for all pairs of altitude levels and averaged over the entire dataset

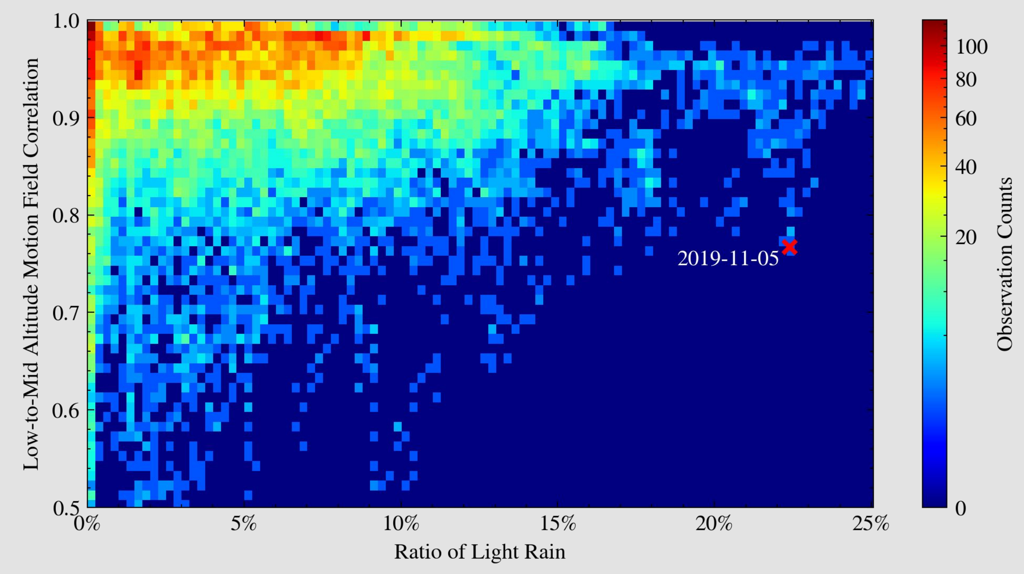
Radar Reflectivity Slice Altitude (m)

	1000	2000	3000	4000	5000	6000	7000	8000
1000	1.00	0.96	0.91	0.83	0.72	0.59	0.44	0.31
2000	0.96	1.00	0.96	0.88	0.76	0.61	0.46	0.32
3000	0.91	0.96	1.00	0.94	0.82	0.65	0.48	0.33
4000	0.83	0.88	0.94	1.00	0.90	0.73	0.53	0.35
5000	0.72	0.76	0.82	0.90	1.00	0.83	0.60	0.39
6000	0.59	0.61	0.65	0.73	0.83	1.00	0.72	0.46
7000	0.44	0.46	0.48	0.53	0.60	0.72	1.00	0.63
8000	0.31	0.32	0.33	0.35	0.39	0.46	0.63	1.00

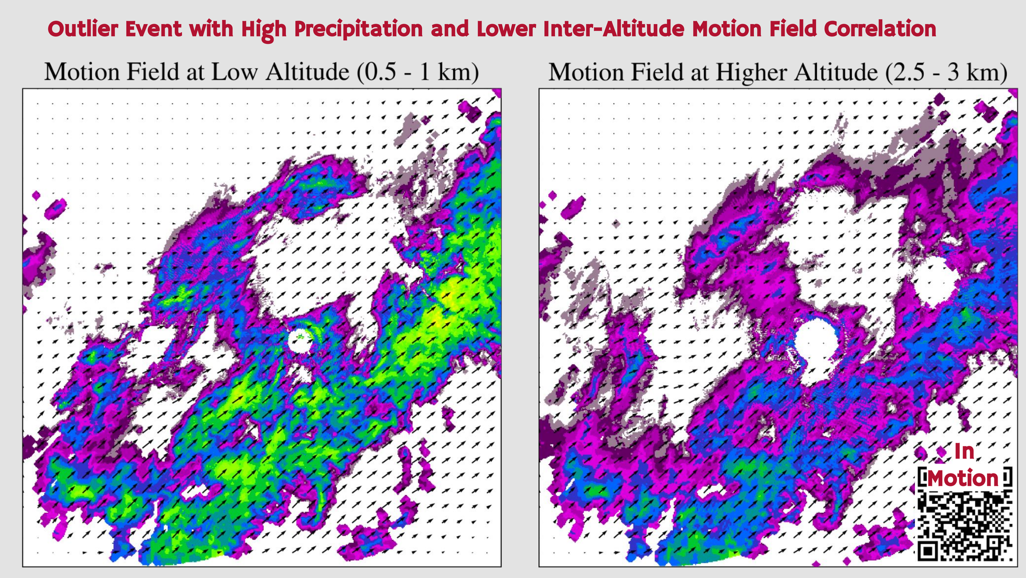
- The estimated motion fields are highly vertically coherent, horizontal advection is shared across vertical levels for most precipitation events

Outlier Inspection

- Identify samples with high overall precipitation coverage and low inter-altitude motion correlation – most promising scenarios in terms of volumetric motion field utility
- Outlier event identified using a joint ranking procedure – 2019-11-05 20:30



Outlier Event with High Precipitation and Lower Inter-Altitude Motion Field Correlation



- Even in the selected outlier sample motion fields remain strongly aligned across altitude levels, despite deliberately selecting an event that maximizes the likelihood of observing vertical differences in horizontal advection (more examples in the manuscript)

Conclusions

- The volumetric model underperforms the 2D baseline in continuous error metrics, but it improves categorical scores. However, these gains largely stem from non-physical cell-splitting artifacts.
- The model's computational efficiency enabled large-scale analysis of the estimated volumetric motion fields across the entire dataset
  - Strong inter-altitude correlations throughout the radar volumes across the dataset
  - The strongest inter-altitude correlations during the warm season
- Independent altitude-wise motion field estimation yields no clear benefit in any of the examined precipitation events from the Slovak radar dataset

Acknowledgements

This work was partially funded by the European Union, under the project LorAI - Low Resource Artificial Intelligence, GA No. 101136646 and by V4Grid, an Interreg Central Europe Programme project co-funded by the European Union, project No. CE0200803. The research was performed in cooperation with Softec, Ltd. and Slovak Hydrometeorological Institute.

Research results were obtained using the computational resources procured in the national project National competence centre for high performance computing (project code: 311070AKF2) funded by European Regional Development Fund, EU Structural Funds Informatization of society, Operational Program Integrated Infrastructure.