Karlsruhe Institute of Technology

Chair of Statistical Methods and Econometrics
Institute of Economics (ECON)
Blücherstr. 17 | 76185 Karlsruhe
https://statistik.econ.kit.edu/
christopher.buelte@student.kit.edu

Probabilistic data-driven weather forecasting

Christopher Bülte, Nina Horat, Sebastian Lerch, Julian Quinting

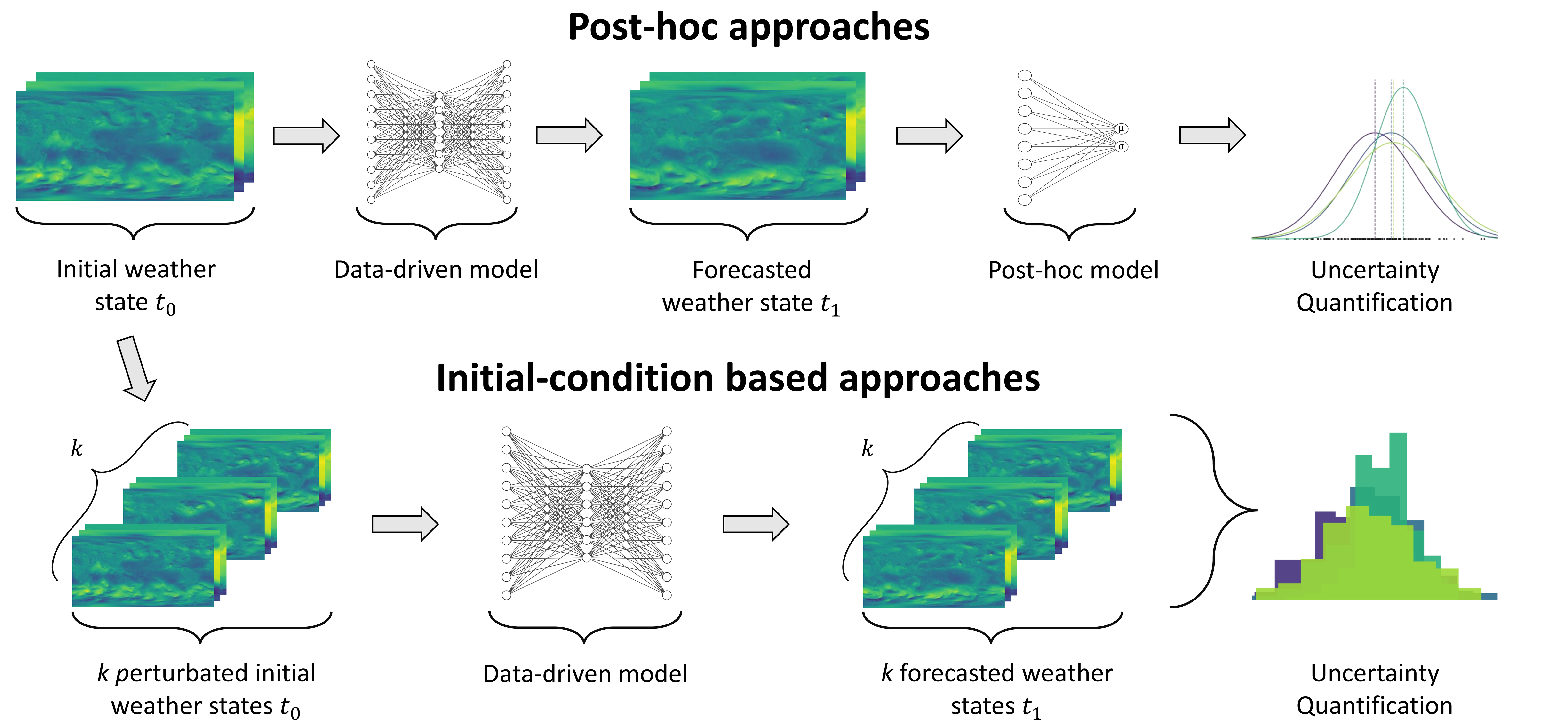
Motivation

- Physical weather models (NWP) vs new data-driven weather models
- Focus of developments so far mostly on deterministic predictions → **Need to incorporate uncertainty**

Data

- Use FourCastNet model [3] trained on ERA5 data to explore different approaches for incorporating probabilistic forecasts
- Analyze different methods on ECMWF European grid for the year 2022

Uncertainty quantification in data-driven weather models



Initial-condition based approaches

Gaussian perturbations (baseline):
Create initial conditions using Gaussian random noise

Physics-based initial conditions (PICS):
Use advanced initial conditions from numerical model

Random Field perturbations (RFP):
Use differences of past observations:

$$d_{noise} = \alpha \frac{a_{d1} - a_{d2}}{\|a_{d1} - a_{d2}\|_{\mathrm{Etot}}},$$

with state vectors a corresponding to the date d_i and tuning parameter α [2].

Post-hoc approaches

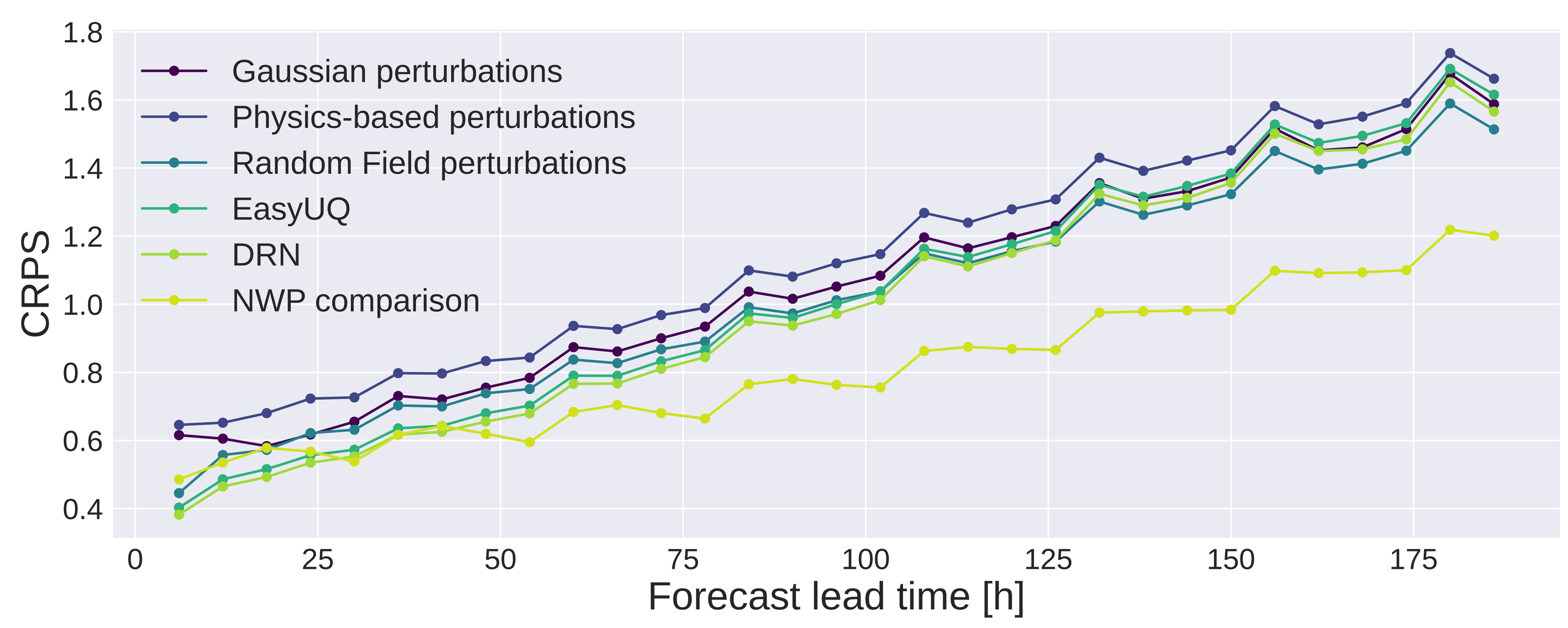
EasyUQ:
Generate predictive distribution based on isotonic distributional regression [1]. For a given pair of model output and outcome (x_i, y_i) prediction is given as:

$$\hat{F}_{x_j}(y) = \min_{k=1,\dots,j} \max_{l=j,\dots,n} \frac{1}{l-k+1} \sum_{i=k}^{l} \mathbf{1}\{y_i \le y\}, \quad j = 1, \dots, n,$$

DRN:
Use a Distributional regression network (DRN) to map the model output to the parameters of a distribution [4].

Results

CRPS (per lead time) for surface temperature



	Gaussian	PICS	RFP	EasyUQ	DRN	*NWP*
u10 (CPRS)	1.338	1.507	**1.290**	1.354	1.345	*1.037*
t2m (CRPS)	1.037	1.323	0.992	0.996	**0.970**	*0.808*

Summary

- Improved approaches useful and outperform baseline
- Performance varies depending on variable and timescale
- Both approaches can be incorporated into the workflow of a data-driven weather forecasting model

References

[1] Alexander Henzi, Johanna F. Ziegel, and Tilmann Gneiting. "Isotonic Distributional Regression". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 83.5 (Aug. 2021), pp. 963–993.

[2] Linus Magnusson, Jonas Nycander, and Erland Källén. "Flow-dependent versus flow-independent initial perturbations for ensemble prediction". In: *Tellus A* 61.2 (2009), pp. 194–209.

[3] Jaideep Pathak et al. *FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators.* 2022.

[4] Stephan Rasp and Sebastian Lerch. "Neural Networks for Postprocessing Ensemble Weather Forecasts". In: *Monthly Weather Review* 146.11 (2018), pp. 3885–3900.