Flow matching for spatially-coherent in-situ weather forecast postprocessing

David Landry*, Claire Monteleoni and Anastase Charantonis
INRIA Paris

EGU

arXiv

EMAIL

Summary

We propose a novel **weather forecast postprocessing** methodology that improves the representation of **spatial and multivariate correlations**.

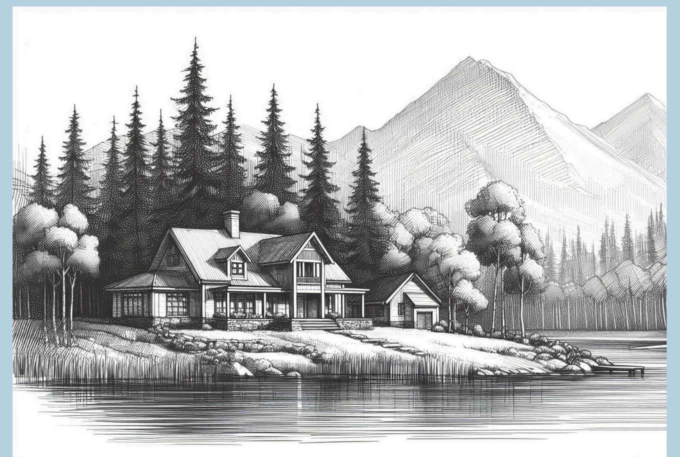
Our methodology is based on the **flow matching** generative model (which is equivalent to denoising diffusion models). Its neural network backbone is an implementation of the **transformer** architecture.

Experiments on the EUPPBench dataset show our method improves in situ weather forecast accuracy, while **better modeling the spatial and multivariate dependencies** between spatial locations.

Weather Forecast Postprocessing

In situ **weather forecast postprocessing** consist in **predicting future station observations** given a gridded weather forecast.

It is motivated by **local effects**: because the gridded forecast has a finite resolution, there are **systematic biases** between observations and the nearest gridpoint. These are typically removed through postprocessing before the forecast can be used in downstram applications.

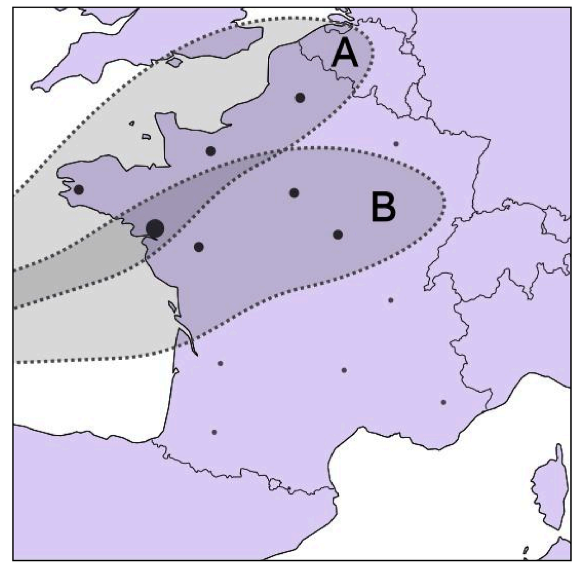


Spatial and multivariate coherence

Accurately modeling spatially-correlated effects is critical to many forecasting applications including **power production/consumption** and **hydrology**.

Most existing postprocessing methods can calibrate forecasts separately by stations, but in doing so they often **degrade spatial dependency structures**. Other methods, based on generative modeling, make progress but lack ensemble spread.

We hope to improve generative modeling for in situ weather forecast postprocessing using **flow matching**, with a particular focus on spatial and multivariate coherence.

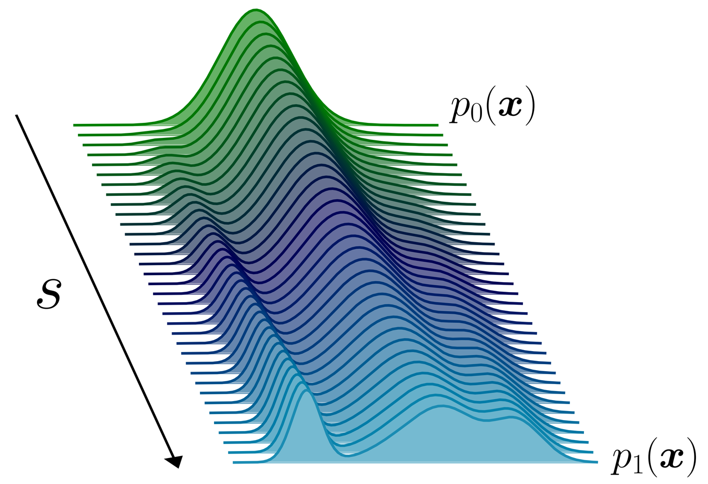


Flow Matching

Flow matching performs **distribution transport**. It displaces a well-known distribution (often a standard normal distribution) towards an unknown distribution for which we only have samples.

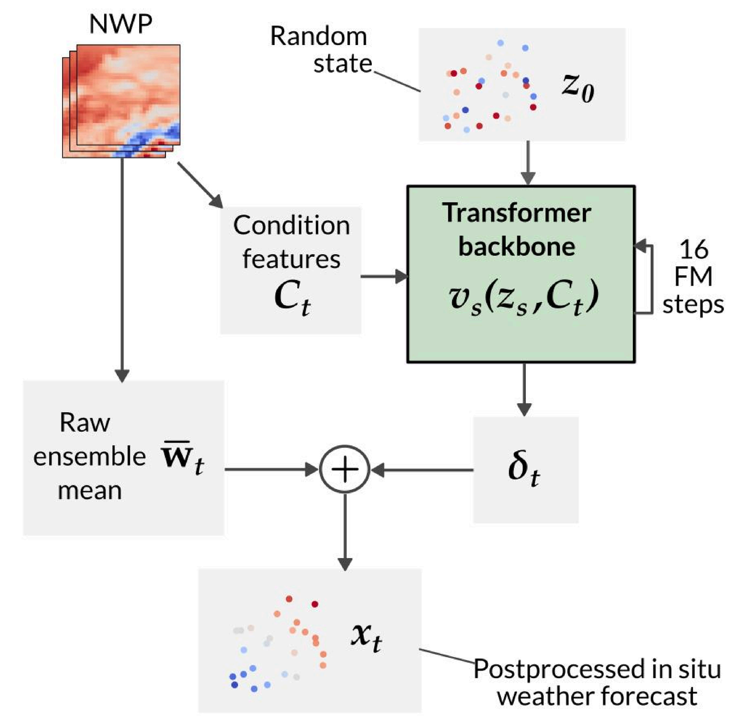
The transport is performed by **vector field** estimated by a neural network. The vector field is **integrated numerically** at inference time to generate new predictions. The NN is trained using a flow matching loss.

Flow matching is an alternative representation of the familiar **denoising diffusion** methods.

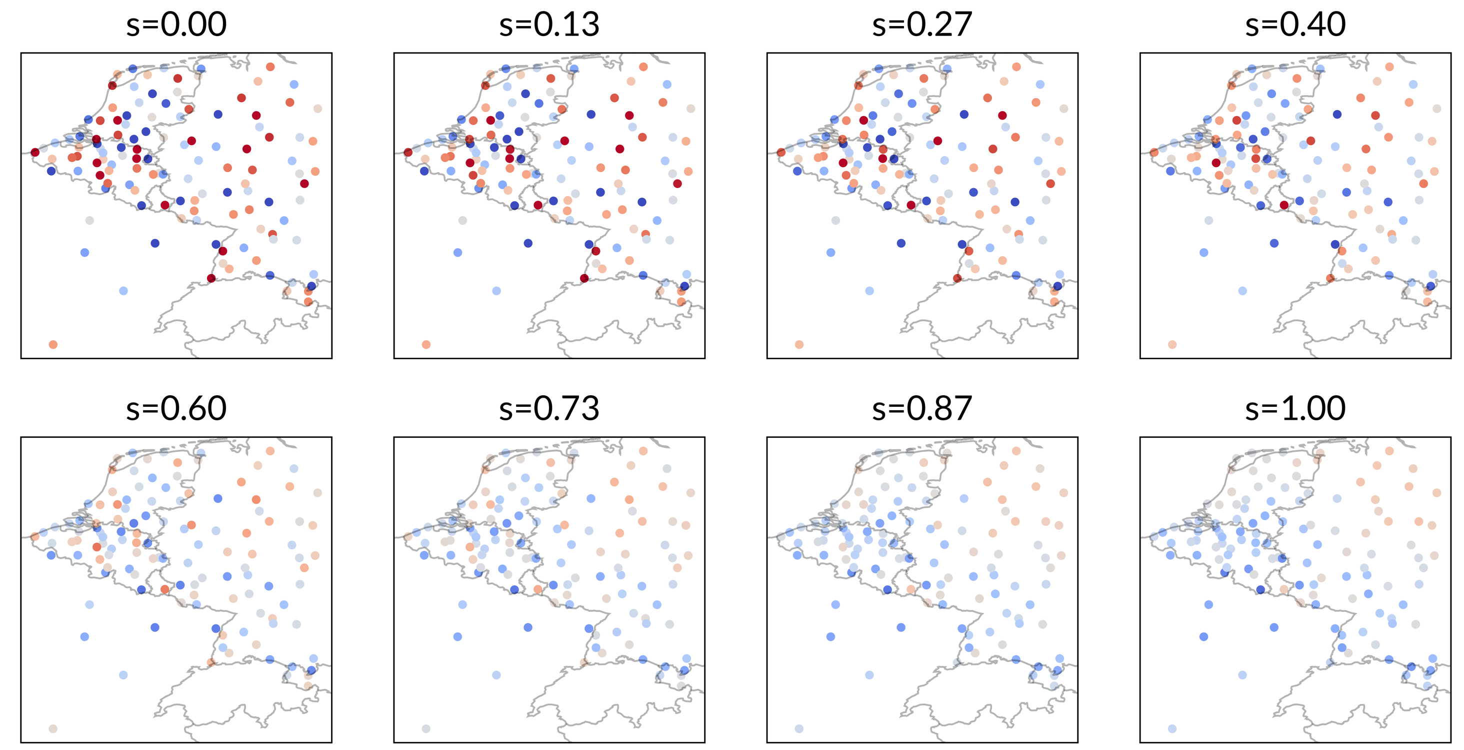


Method

Model overview



Flow matching visualized



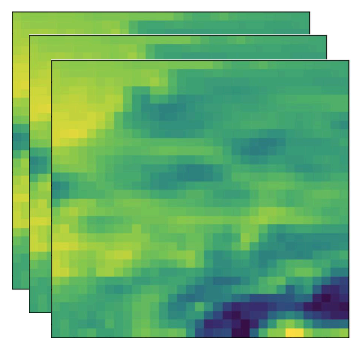
Flow matching is a great tool for *spatially coherent* in-situ weather forecasting

Experiments and Results

EUPPBench dataset

The EUPPBench dataset has **4180 reforecasts, 730 forecasts** over 20 steps. It has matching **surface temperature and wind speed** observations.

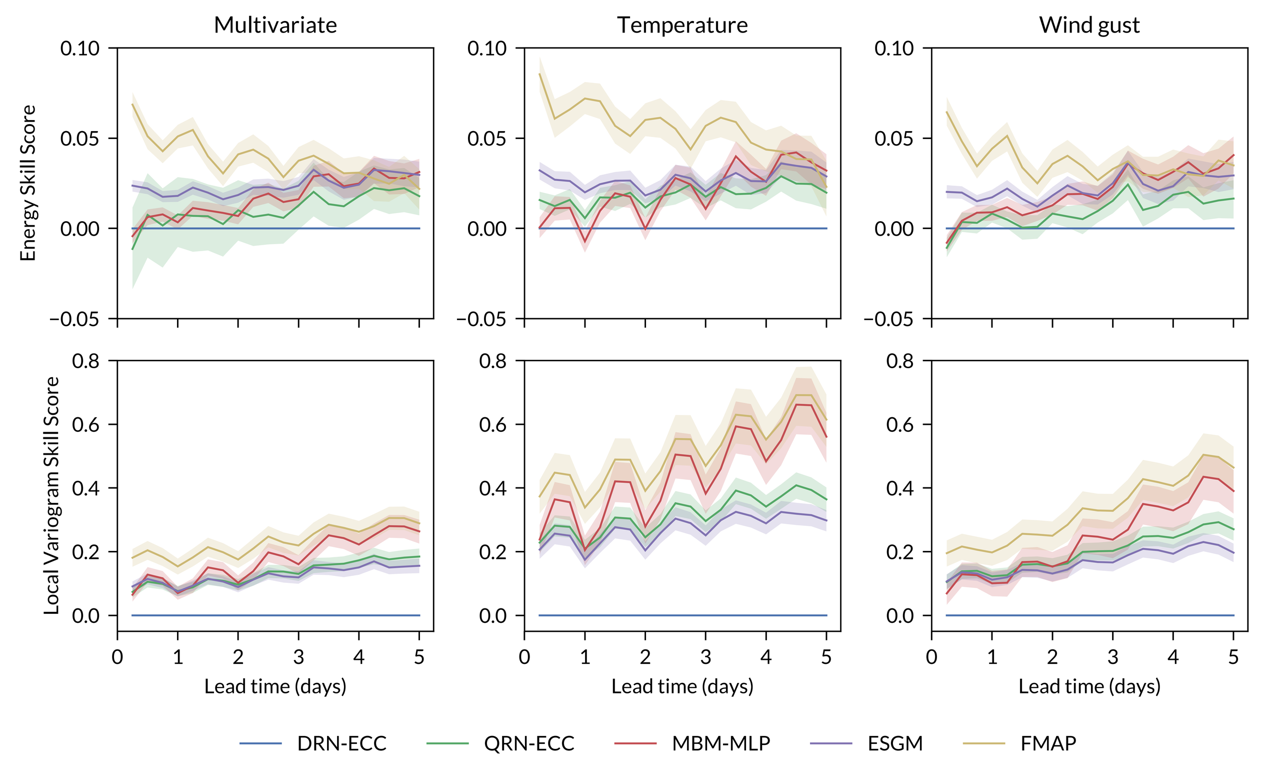
The reforecasts have 11 members, the forecasts 51. The model has a resolution of 0.25°.



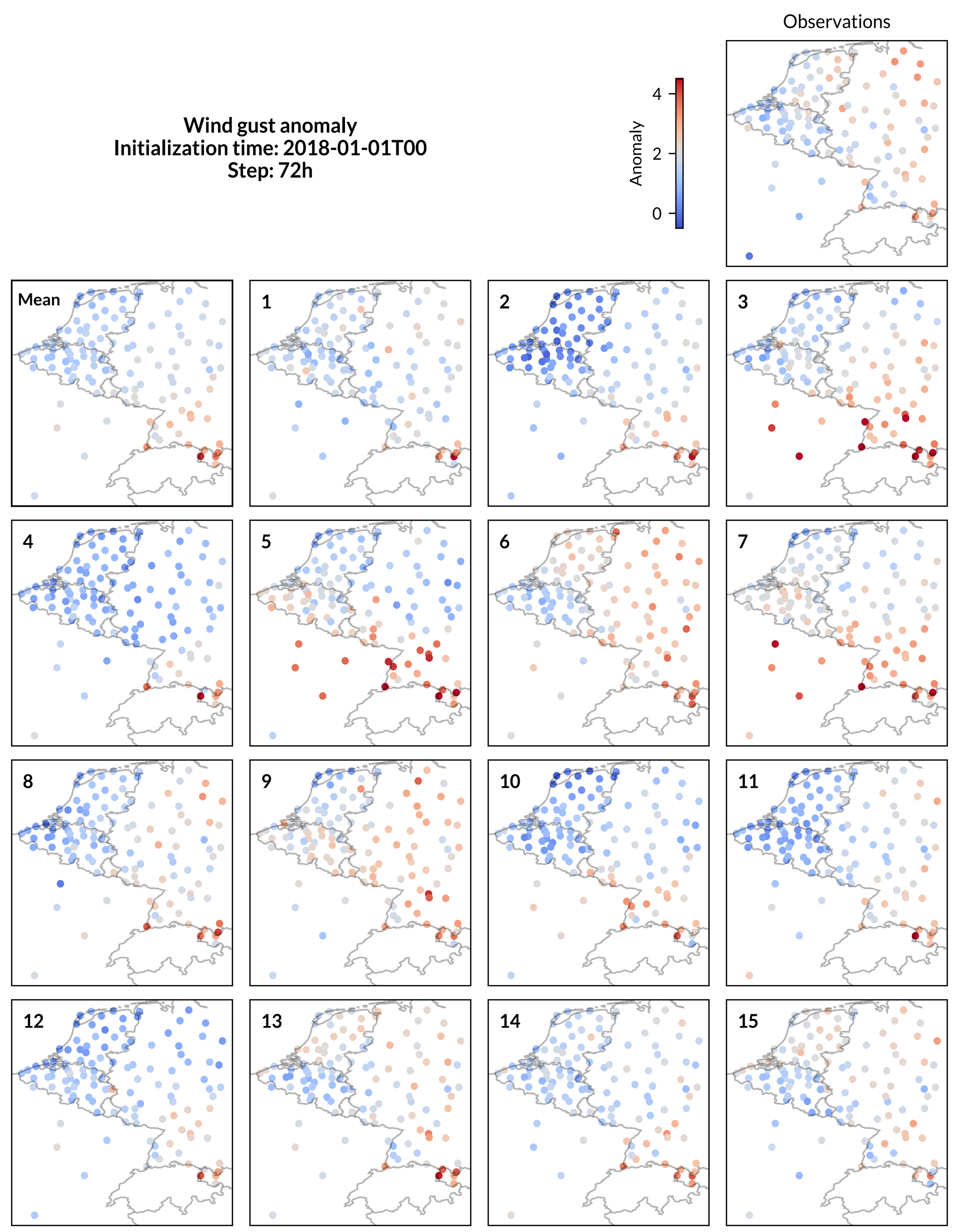
Reforecasts

Observations

Performance metrics



Sample forecast

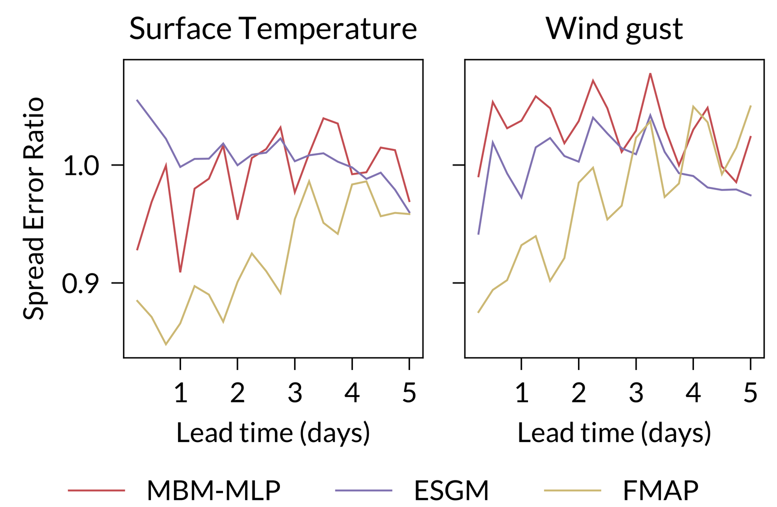


Conclusion and Perspectives

Our methodology, based on flow matching and a spatial attention transformer, **improved state of the art in weather forecast postprocessing**. The shift from supervised learning to **generative modeling** allows a better representation of spatial and multivariate dependencies.

The generated distributions are **slightly underdispersive** despite achieving the best energy scores, which constitutes an interesting limitation.

Future work: extend this approach to spatio-temporal modeling. Since the problem dimensionality is smaller than for full weather forecasting, we hope this will be manageable without auto-regression.



References

Vannitsem, S. et al. (2021) '**Statistical Postprocessing for Weather Forecasts: Review, Challenges, and Avenues in a Big Data World**', Bulletin of the American Meteorological Society, 102(3), pp. E681–E699. Available at: https://doi.org/10.1175/BAMS-D-19-0308.1.

Lipman, Y. et al. (2024) '**Flow Matching Guide and Code**'. arXiv. Available at: https://doi.org/10.48550/arXiv.2412.06264.

Chen, J. et al. (2024) '**Generative machine learning methods for multivariate ensemble post-processing**', The Annals of Applied Statistics, 18(1). Available at: https://doi.org/10.1214/23-AOAS1784.

Lakatos, M. et al. (2023) '**Comparison of multivariate post-processing methods using global ECMWF ensemble forecasts**', Quarterly Journal of the Royal Meteorological Society, p. qj.4436. Available at: https://doi.org/10.1002/qj.4436.