Evaluating Conformal Prediction in Weather Forecasting

Thomas Mortier, Cas Decancq, Diego G. Miralles

Hydro-Climate Extremes Lab, Faculty of Bioscience Engineering, Ghent University, Belgium

Abstract

Machine learning offers a promising alternative to traditional numerical weather prediction models, providing faster and more cost-effective forecasts. However, a major challenge remains the lack of reliable uncertainty quantification, which is essential for intepretable and trusthworthy predictions in complex climate systems. Existing Bayesian and frequentist approaches often face scalability issues or rely on restrictive assumptions. Conformal prediction (CP) offers a flexible, model-agnostic framework that yields valid prediction intervals without requiring assumptions about the underlying data distribution. Yet, its application to weather forecasting is limited by spatial and temporal dependencies that violate CP's exchangeability assumption. In this work, we assess the potential of CP for weather forecasting, discuss its main limitations, and highlight recent advances and future directions.

Introduction

In recent years, machine learning has emerged as a promising alternative to numerical weather prediction models, offering the potential for cost-effective and accurate forecasts. However, a significant limitation of current machine learning methods for weather forecasting is the lack of principled and efficient uncertainty quantification—a key element given the complexity of the Earth's climate system and the challenges in modelling its processes and feedback mechanisms. Inadequate uncertainty quantification and reporting undermines trust in and the practical use of current weather forecasting methods [Eyring et al., 2024].

Uncertainty quantification methods for weather forecasting typically use prediction intervals and can be categorized into Bayesian and frequentist approaches. Bayesian methods, while theoretically appealing, often involve restrictive assumptions and do not scale well to the complexity of spatio-temporal data. Frequentist approaches, such as ensemble-based methods, are widely used in weather forecasting and include techniques like perturbing initial states with noise [Bi et al., 2023, Scher and Messori, 2021], varying neural network parameters [Graubner et al., 2022], or training generative models [Price et al., 2025].

However, most frequentist methods provide only asymptotically valid prediction intervals, which may not suffice in all weather forecasting applications.

Conformal prediction is a promising uncertainty quantification framework that delivers valid and efficient prediction intervals for any learning algorithm, without requiring assumptions about the underlying data distribution [Vovk et al., 2005]. Despite its growing popularity in the machine learning and statistics communities, traditional CP methods are not tailored to spatio-temporal data in weather forecasting. This is due to challenges arising from spatial and temporal dependencies, such as spatial autocorrelation and temporal dynamics, that violate the exchangeability assumption underlying standard CP methods. Several recent studies attempted to address these challenges by introducing new CP algorithms specifically designed for various types of non-exchangeability [Oliveira et al., 2024]. However, these adaptations face several limitations, including high computational complexity, asymptotic guarantees, and/or the need for recalibration of prediction intervals.

In this work, we will evaluate CP methods in the context of weather forecasting and discuss several limitations. In addition, we will highlight recent advances and discuss potential future directions that could address challenges underlying the use of CP in weather forecasting.

References

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian. Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 619(7970):533–538, 2023.

Veronika Eyring, William D Collins, Pierre Gentine, Elizabeth A Barnes, Marcelo Barreiro, Tom Beucler, Marc Bocquet, Christopher S Bretherton, Hannah M Christensen, Katherine Dagon, et al. Pushing the frontiers in climate modelling and analysis with machine learning. *Nature Climate Change*, 14(9):916–928, 2024.

Andre Graubner, Kamyar Kamyar Azizzadenesheli, Jaideep Pathak, Morteza Mardani, Mike Pritchard, Karthik Kashinath, and Anima Anandkumar. Calibration of large neural weather models. In *NeurIPS 2022 Workshop on Tackling Climate Change with Machine Learning*, 2022.

Roberto I Oliveira, Paulo Orenstein, Thiago Ramos, and João Vitor Romano. Split conformal prediction and non-exchangeable data. *Journal of Machine Learning Research*, 25(225):1–38, 2024.

Ilan Price, Alvaro Sanchez-Gonzalez, Ferran Alet, Tom R Andersson, Andrew El-Kadi, Dominic Masters, Timo Ewalds, Jacklynn Stott, Shakir Mohamed, Peter Battaglia, et al. Probabilistic weather forecasting with machine learning. *Nature*, 637(8044):84–90, 2025.

Sebastian Scher and Gabriele Messori. Ensemble methods for neural network-based weather forecasts. *Journal of Advances in Modeling Earth Systems*, 13 (2), 2021.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*, volume 29. Springer, 2005.