EMS Annual Meeting Abstracts
Vol. 21, EMS2024-267, 2024, updated on 16 Feb 2025
https://doi.org/10.5194/ems2024-267
EMS Annual Meeting 2024

Machine learning-driven assessment and prediction of sub-seasonal forecast skill

Ana-Cristina Mârza[1,2], Daniela I.V. Domeisen[2,3], and Angela Meyer[1,4]
[1]Bern University of Applied Sciences, School of Engineering and Computer Science, Bern, Switzerland
[2]University of Lausanne, Lausanne, Switzerland
[3]ETH Zurich, Zurich, Switzerland
[4]TU Delft, Delft, Netherlands

Certain atmospheric conditions provide windows of opportunity for sub-seasonal prediction, such that forecasts initialized at these times maintain good predictive skill further out into the future than usual. Such windows of opportunity are for example given by certain phases of the Madden-Julian oscillation and the state of the stratospheric polar vortex at forecast initialization.

These findings come from retrospective analyses of forecast skill that often investigate only one or two drivers of sub-seasonal predictability at a time. What is lacking is investigating multiple drivers and comparing their relative importance, which we propose to achieve by means of explainable machine learning (ML) techniques. Furthermore, whereas previous ML-based studies of atmospheric predictability quantified uncertainty in terms of deterministic error or ensemble spread as a proxy for forecast skill, we here aim to predict the probabilistic forecast skill itself. The result is an operational decision support tool that can inform users, at forecast initialization time, of the skill expected at sub-seasonal lead times.

Therefore we are presenting, for the first time to our knowledge, an explainable ML model (based on random forests) that learns to predict directly the skill of the ECMWF extended-range forecasts at several weeks' lead time, given the atmospheric state at initialization. We use reanalysis data as ground truth and focus on the skill of weekly aggregated statistics of geopotential height at 500 hPa, a variable of particular interest to the renewable energy industry since it tracks the synoptic-scale weather patterns controlling European energy production and demand.

We expect that our forecast skill prediction model will enable enhanced risk management for sub-seasonal forecast users, such as stakeholders in the renewable energy industry, and thereby accelerate the transition to a decarbonized energy system. From a scientific perspective, interrogating our ML model with explainability techniques should yield new insights into the sources of sub-seasonal predictability.