Can Machine Learning Models be a Suitable Tool for Predicting Central European Cold Winter Weather on Subseasonal Timescales?

Karlsruhe Institute of Technology

WAVES TO WEATHER

Link to Kiefer et al. (2023)

Sharing is encouraged

Selina M. Kiefer[1], Sebastian Lerch[2], Patrick Ludwig[1], Joaquim G. Pinto[1]

[1] Institute of Meteorology and Climate Research Troposphere Research, [2] Institute of Statistics

1 AIMS

1. Suitability of purely data-driven machine learning (ML) models for forecasting daily mean 2-meter temperature and occurrence of days with cold waves (> 2 days with unusually cold temperatures)
   - lead times of 14, 21 and 28 days
   - predictors based on meteorological knowledge, input region see Fig. [1]
   - target region is Central Europe (without first coast gridpoints and terrain above 800m altitude), outline shown on Fig. [1]

2. Explainability of ML-models' forecasts and possible physical relevance of the learned pattern in the data

2 METEOROLOGICAL INPUT AND BENCHMARK

Predictors:
ERA-5 reanalysis [a],
wind (u10, u300), geopotential (z100, z250, z500, z850), temperature (t850), specific humidity (H850), pressure (msl), month

[1] Input region — Target region

Ground Truth:
E-OBS V23.1e [b],
2-meter temperature (tg) and cold wave days

Climatological Ensemble:
ground truth data from 1970 – 1999, each winter serves as one ensemble member;
used as the benchmark model

Numerical Forecasts:
ECMWF's S2S Reforecasts [c]

Input structure of ML-models:

Model Abbr.	Input Features
_all	u10, z100, z250, z500, z850, t850, H850, u300, msl, month
_sel	u10, z100, z250, month
QRF_ RFC_	all grid points of the fields + month
QRF_stat RFC_stat	minimum, mean, maximum and variance of the fields + month
QRF_pca RFC_pca	first 10 principal components of the fields + month

Skill Measures [e]:

the Continuous Ranked Probability (Skill) Score (CRPS, CRPSS) is used for continuous forecast evaluation of 2-meter temperature and the Brier (Skill) Score (BS, BSS) for binary forecasts of cold wave days

REFERENCES

[a] e.g. DOI: 10.24381/cds.bd0915c6
[b] DOI: 10.1029/2017JD028200
[c] e.g. DOI: 10.24381/cds.bd0915c6
[d] https://skranger.readthedocs.io/en/stable/
[e] DOI: 10.1111/j.1467-9868.2007.00587.x
[f] https://shap.readthedocs.io/en/latest/

3 MACHINE LEARNING MODELS

Quantile Random Forests (QRFs) [d]:
100 predicted equidistant quantiles as ensemble members, 100 decision trees, continuous forecasts of 2-meter temperature

Random Forest Classifiers (RFCs) [d]:
predictions of 100 decision trees as ensemble members, binary forecasts of cold wave days

Decision Tree Working Principle:

splitting criterion is learned — random subset of predictors — RF = multiple trees

as similar as possible — as heterogeneous as possible — as similar as possible

... ...

in the final parts: ground truth data belonging to the predictors; predictions: quantiles (→ QRF) or fractions (→ RFC) of the data in the respective final part

Training:
1950 – 2020, Oct – Apr, the evaluated winter (one from 2000-2020) left out

Shapley Additive Explanations (SHAP) [f]:

used to analyze which predictors contribute most to the models' predictions during a certain time period (visualized for 14d lead on Fig. [3] and [5])

5 CONCLUSIONS

1. in the 20-winter mean, skill can be found for some models at lead times of 14, 21 and 28 days compared to the climatological ensemble (positive values on Fig. [2], [4] and [6])

2. the skill of the ML-models compared to the climatological ensemble decreases with increasing lead time

3. the skill of all models and which model performs best strongly depends on the winter to be forecasted (whiskers on Fig. [2], [4] and [6] show variability of skill between winters)

4. the SHAP analysis shows that predictors representing the large-scale atmospheric flow contribute highly to the ML-models predictions (Fig. [3], [5])

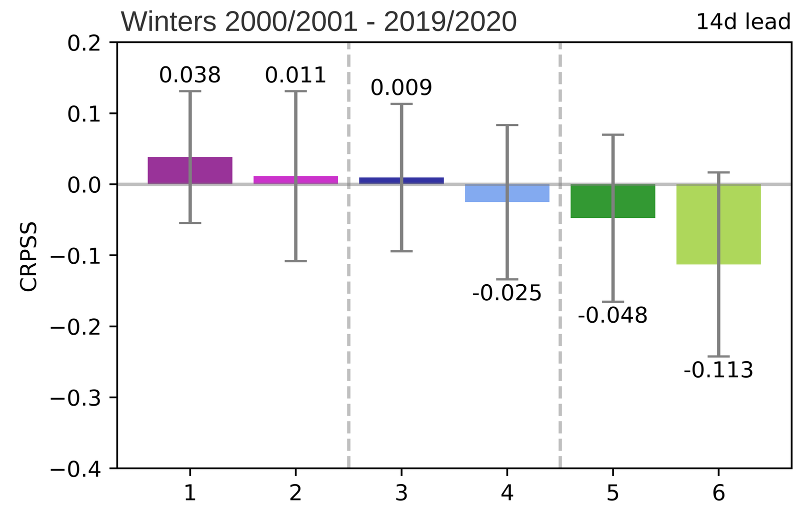
OUTLOOK *currently under review at AIES*

some of the analyzed ML-models show a higher skill than numerical forecasts at lead times of 21 and 28 days (for 28d lead see Fig. [6])

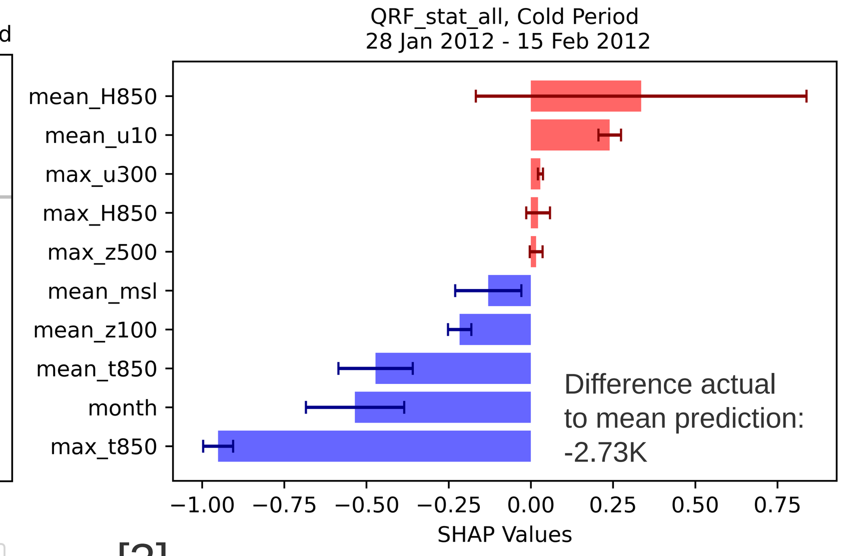
→ does the skill of these ML-models improve when including the S2S reforecasts directly as predictors?

4 RESULTS

CONTINUOUS TEMPERATURE FORECAST

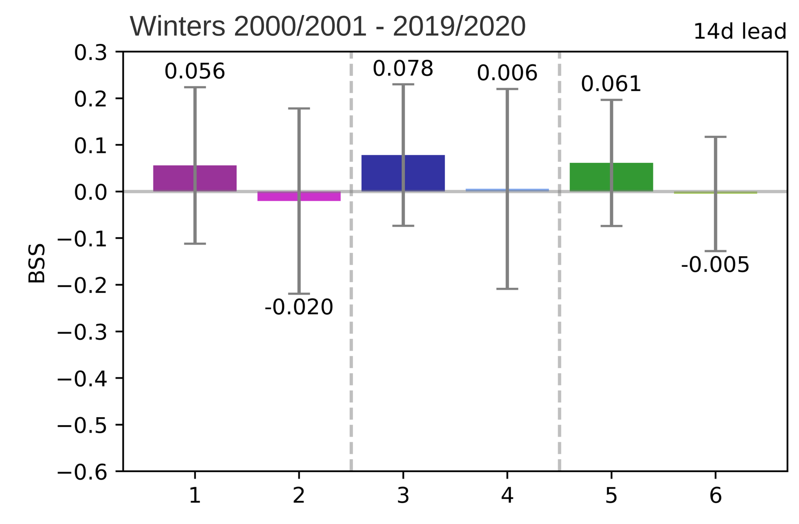


[2]

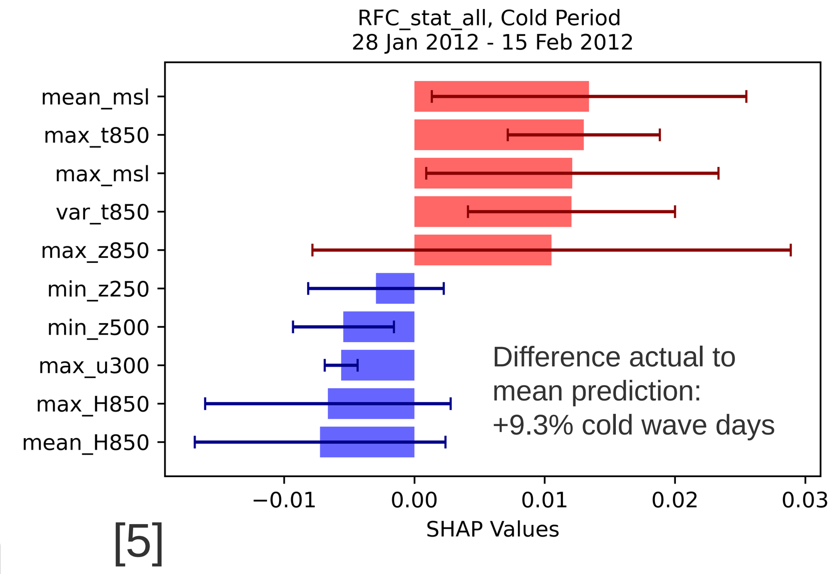


[3]

BINARY COLD WAVE DAY FORECAST

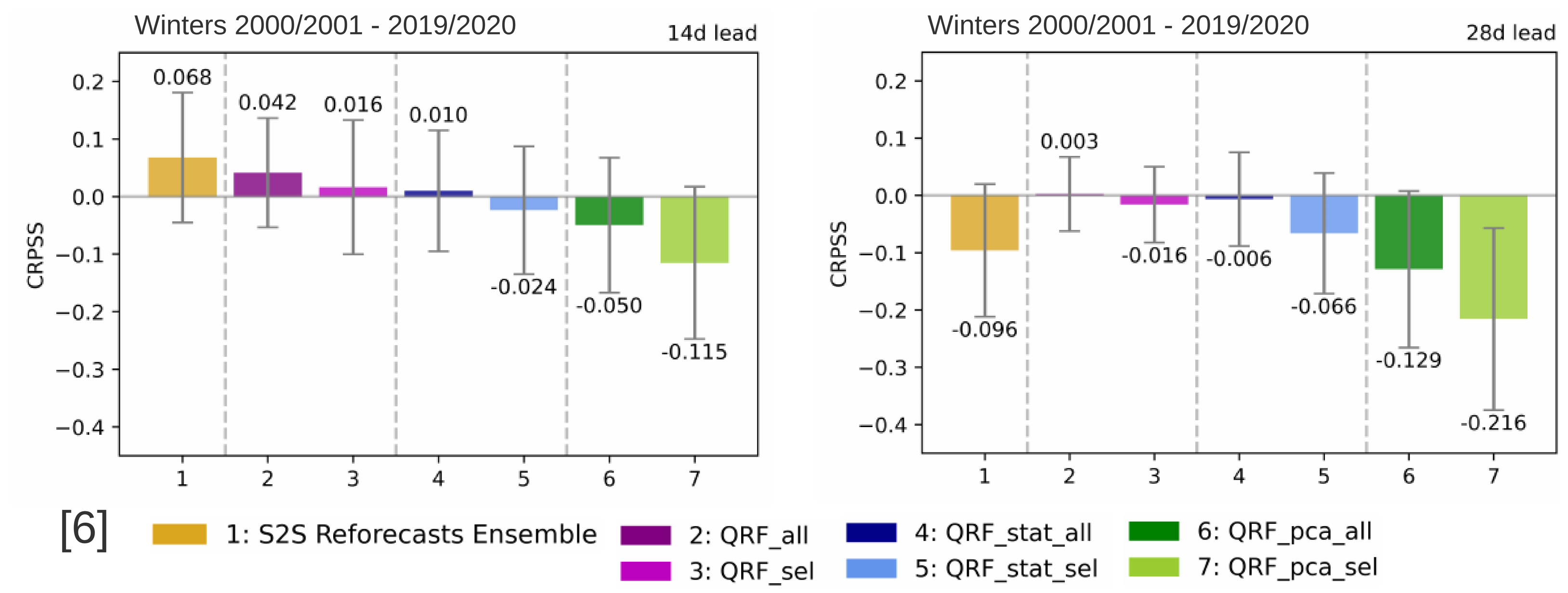


[4]



[5]

COMPARISON WITH NUMERICAL FORECAST



[6]

NOTE: only dates present in both, the S2S Reforecasts and the ML predictions are considered (52 dates/ winter)

Kiefer, S. M., S. Lerch, P. Ludwig, and J. G. Pinto, 2023: Can Machine Learning Models Be a Suitable Tool for Predicting Central European Cold Winter Weather on Subseasonal to Seasonal Time Scales? Artificial Intelligence for the Earth Systems, 2 (4), e230 020, https://doi.org/10.1175/AIES-D-23-0020.1.