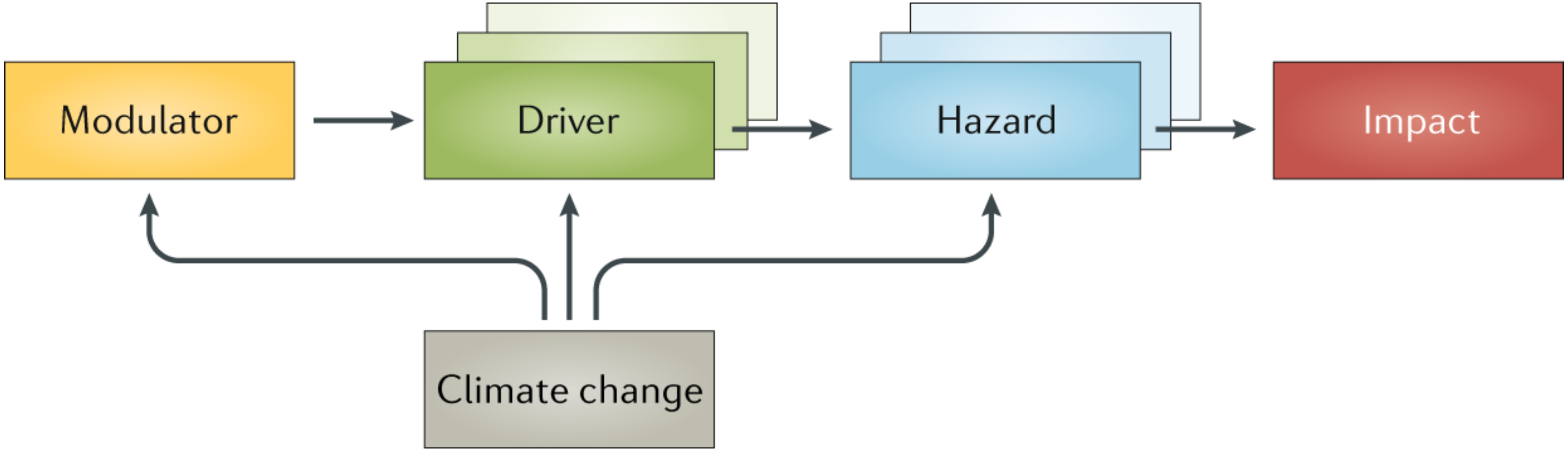
Using interpretable machine learning to identify compound meteorological drivers of crop failure

Lily-belle Sweet, Jakob Zscheischler

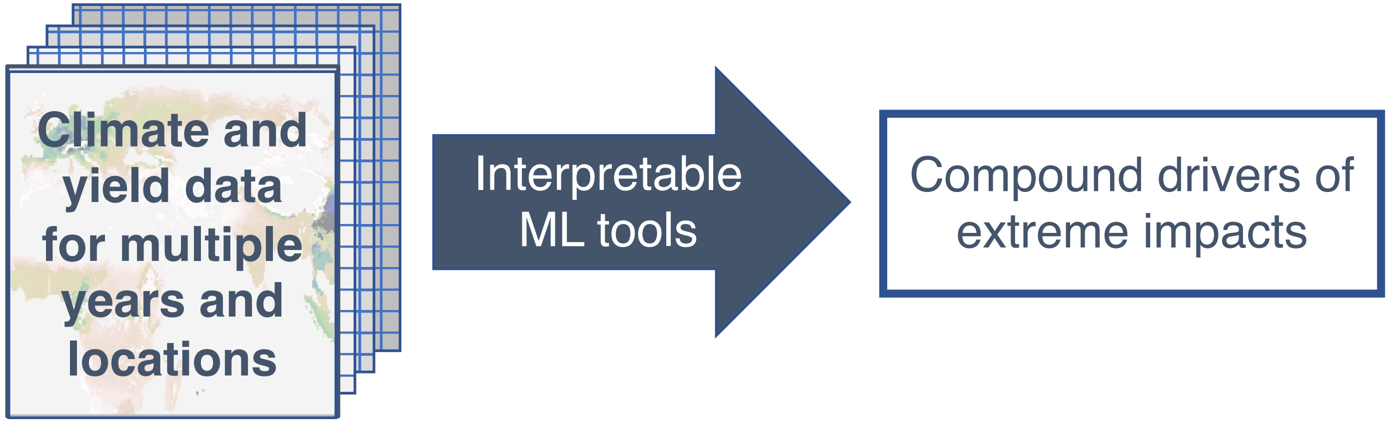


HELMHOLTZ
Zentrum für Umweltforschung

Department of Computational Hydrosystems,
Helmholtz Institute for Environmental Research - UFZ

Motivation

- Crop yield failure can be caused by a combination of non-extreme weather events
- Observed yield datasets are short and not well-distributed globally
- Including data from multiple locations would permit use of more complex models



- Good model performance is a prerequisite for useful interpretations
- **Aim:** Investigate the impact of cross-validation method on model performance metrics and interpretations for spatiotemporal data

UFZ HELMHOLTZ Zentrum für Umweltforschung

Data and model

Climate data

- Global daily reanalysis data
- Covers 1948-2008
- 0.5 degree resolution
- Variables used: pr, tas, averaged monthly
- Use 3 months before planting date plus the duration of growing season

Crop yield data

- Maize yield data from LPJmL
- Covers 1948-2008
- 0.5 degree resolution
- No irrigation, adaptation, fertilization etc considered
- Current cropping areas only

Yield failure

- Maize yield data detrended at each gridpoint
- Lowest 10% of years at each gridpoint are considered to be yield failure years

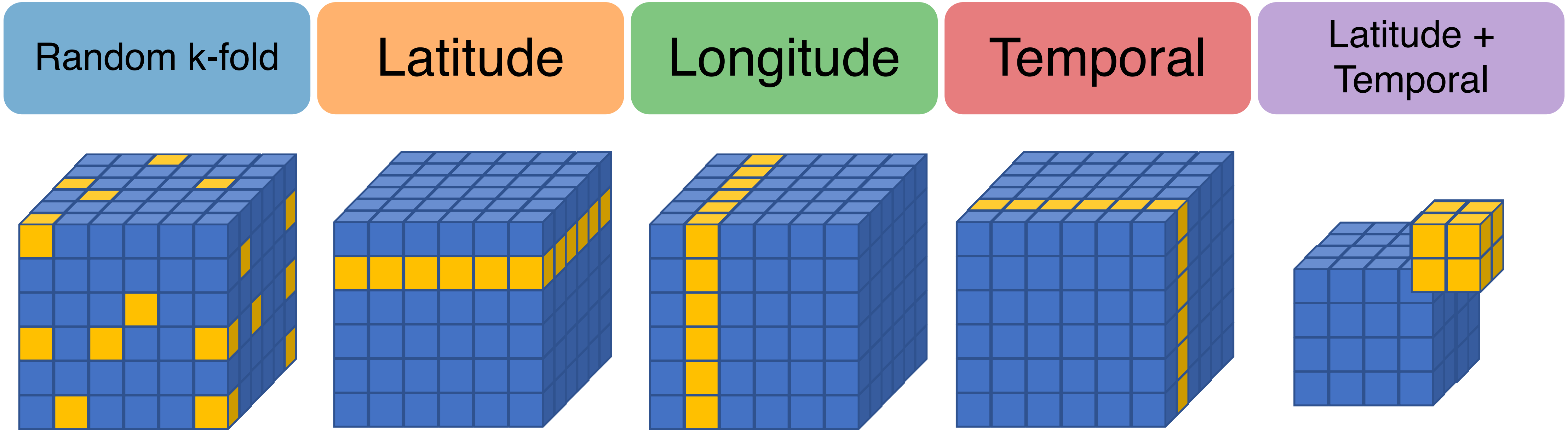
Random forest

Performance evaluation

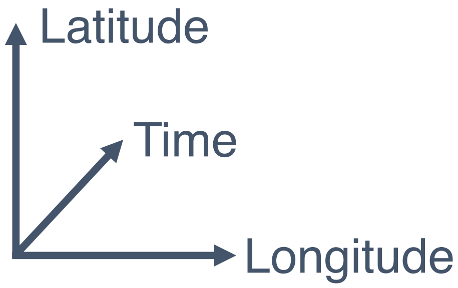
Model interpretation

HELMHOLTZ
Zentrum für Umweltforschung

Cross-validation strategies

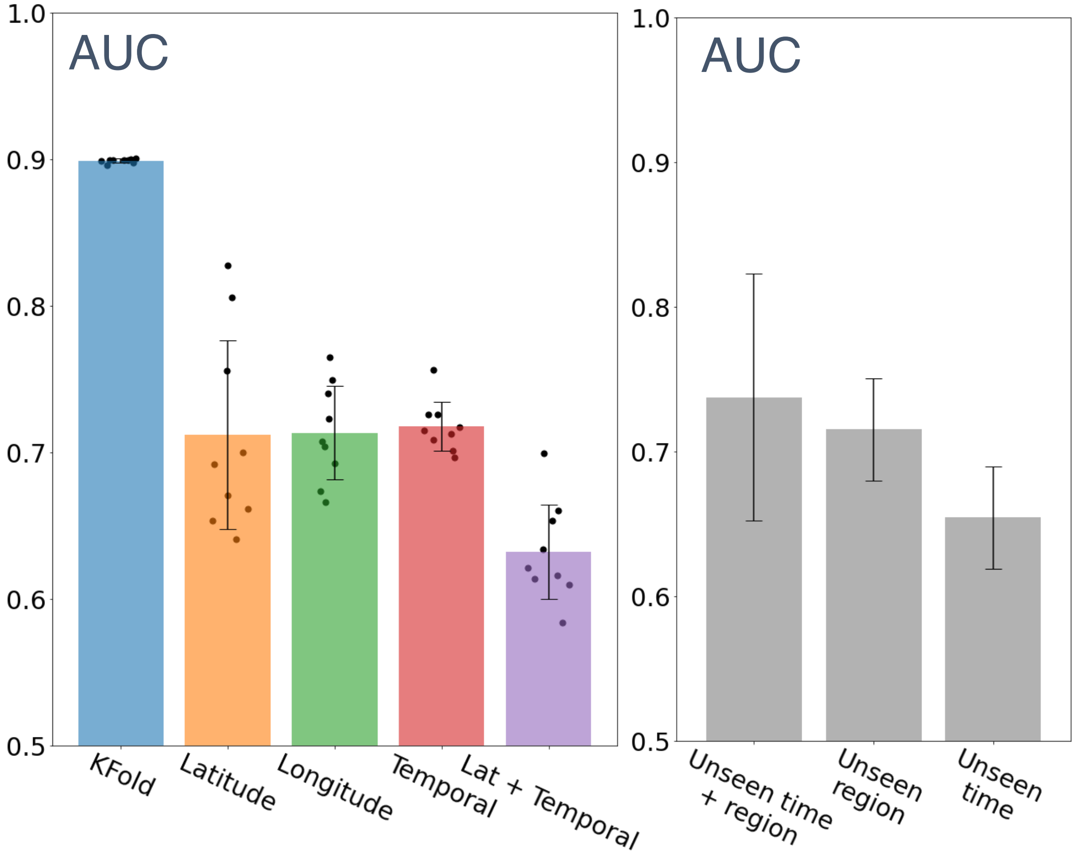


- Germany and Bolivia were kept out of the initial dataset, along with the last 6 years of data, as an additional test set.



UFZ HELMHOLTZ
Zentrum für Umweltforschung

Results

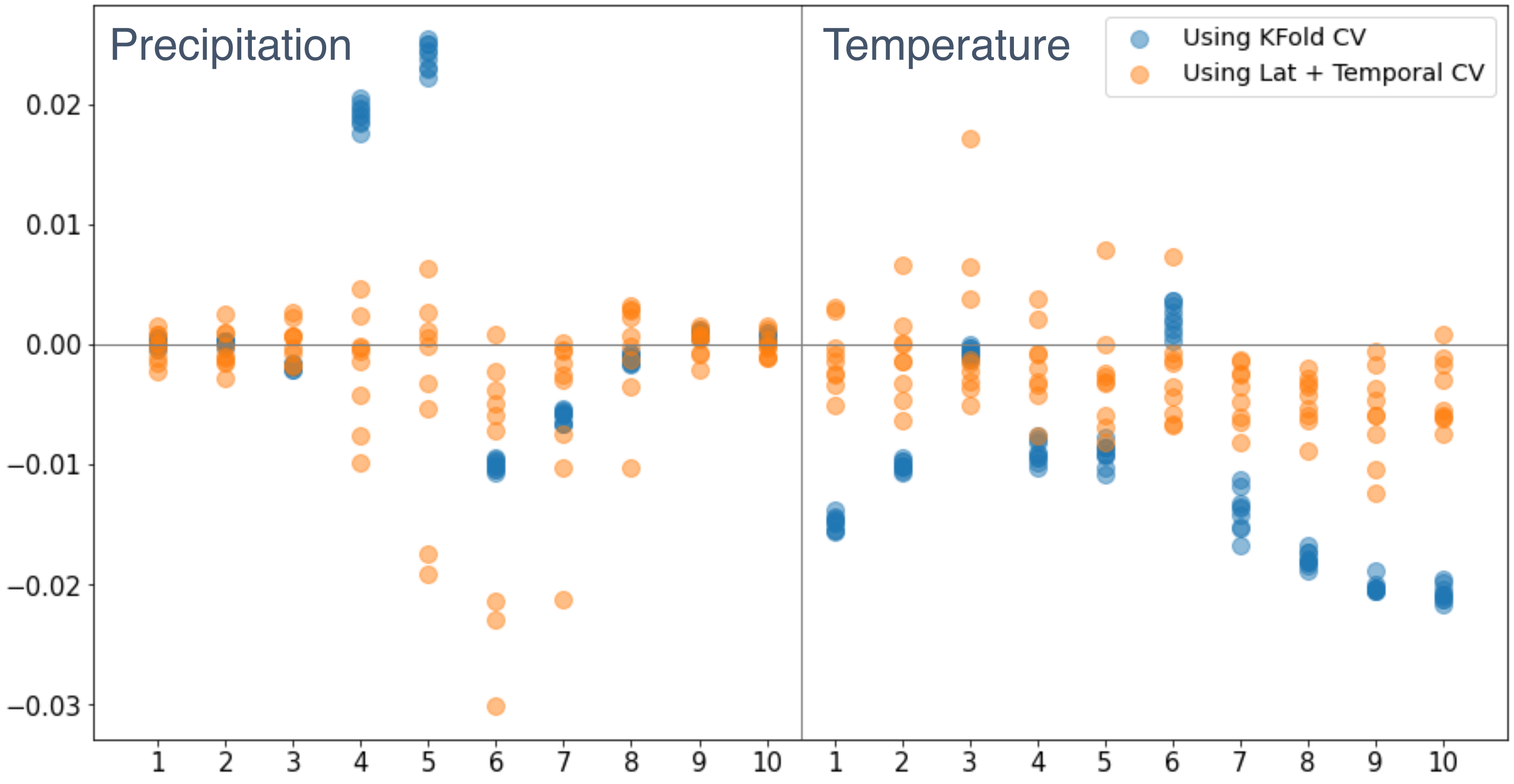


- Performance metrics calculated using random k-fold CV are more optimistic than from spatial or temporal CV
- Spatial and temporal CV better reflects model skill on unseen regions and times

HELMHOLTZ
Zentrum für Umweltforschung

Model interpretation

- Choice of cross-validation strategy has a large impact on permutation importances



HELMHOLTZ
Zentrum für Umweltforschung

Conclusions

- When using spatiotemporal climate + yield datasets, random k-fold cross-validation can overestimate model skill.
- Model skill measured using spatial + temporal cross-validation is more representative of performance on new data.
- Choice of cross-validation method should be carefully considered when using interpretable ML methods on spatiotemporal climate data.

References
1. Zscheischler, J. *et al.* Nature Reviews Earth & Environment (2020)
2. Müller, C. *et al*. Sci Data (2019)
3. Sheffield, J. *et al.* J. Climate (2006)
4. Portmann, F. *et al.* Global Biogeochemical Cycles (2010)

UFZ HELMHOLTZ Zentrum für Umweltforschung