Forecasting crop yield losses using satellite data and machine learning

E. Bueechi, M. Fischer, L. Crocetti, M. Trnka, A. Grlj, L. Zappa, W. Dorigo

Crop yields

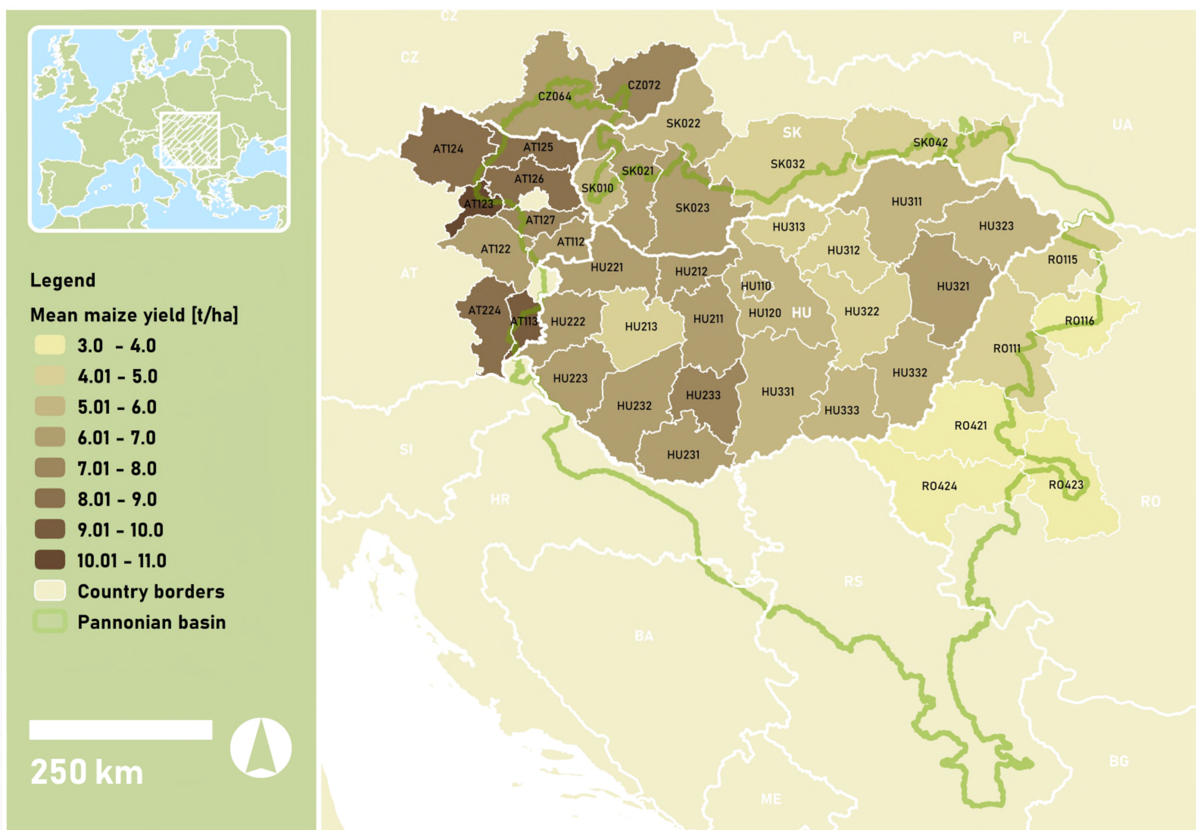

Fig. 1: Overview of the study area

Background

Droughts have heavily affected agricultural production in the Pannonian Basin. As most fields there are only rain-fed, local agriculture is particularly vulnerable to droughts. The already challenging conditions for crop production are expected to worsen due to climate change, by increasing frequencies and intensities of drought events. A potential tool to support the adaption to these challenging circumstances is crop yield forecasting. This has proven being a vital tool to minimize socio-economic impacts of crop losses. However, such forecasts tend to underestimate the impact of severe droughts on yield losses and, therefore, require improvements [1].

Predictors

Dataset	Source	Spatial resolution	Temporal resolution
Vegetation conditions			
VOD Ku-band	VODCA	0.25°	daily
NDVI	CGLS	0.01°	10-daily
LAI	CGLS	0.01°	10-daily
Precipitation			
Precipitation	E-OBS	0.25°	daily
Fraction of wet days	E-OBS	0.25°	monthly
Seasonal forecast (next 1 and 2 months)	ECMWF	1°	monthly
Temperature and radiation			
Temperature	E-OBS	0.25°	daily
Diurnal temperature	ERA5-Land	0.1°	daily
Surface net solar radiation	ERA5-Land	0.1°	daily
Seasonal forecast (next 1 and 2 months)	ECMWF	1°	monthly
Soil moisture			
Soil moisture	ESA CCI	0.25°	daily
Soil water index	ESA CCI	0.25°	daily
Drought indicators			
ESI (1 and 3 months)	MODIS	0.05°	weekly
SPEI (1 and 3 months)	ERA5	0.25°	monthly

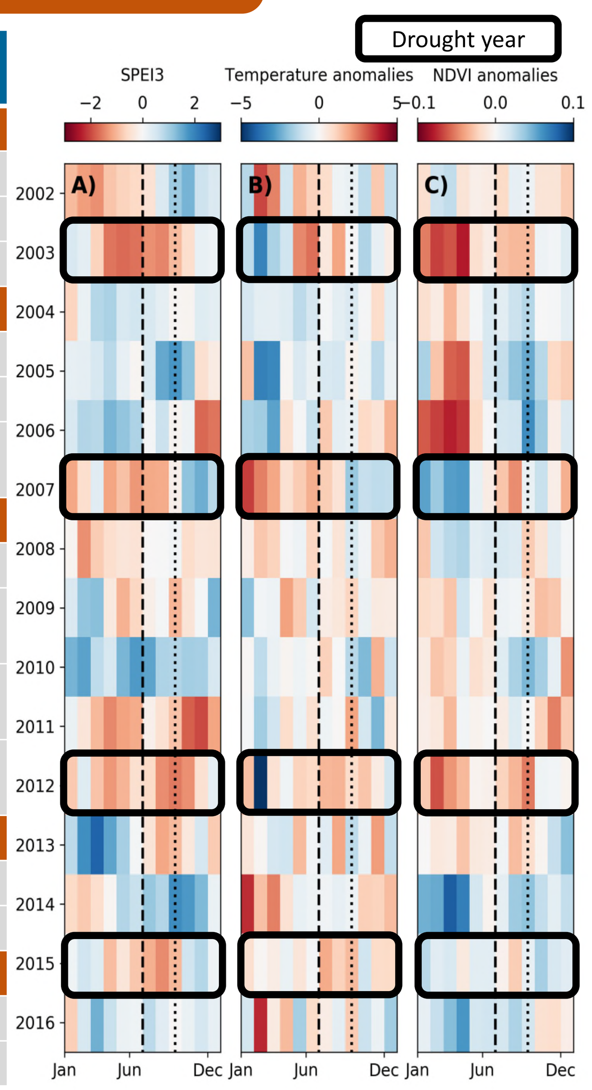

Fig. 2: List of used predictors and three examples on the right

Yearly yield anomalies of maize and winter wheat are forecasted for various districts in the Pannonian Basin from 2002-2016. Monthly forecasts are made for each growing season, starting around three months before harvest.
The forecasts are cross-validated by using leave-one-year-out as testing set. Different model optimization techniques were used like feature elimination and hyperparameter tuning.

Machine learning

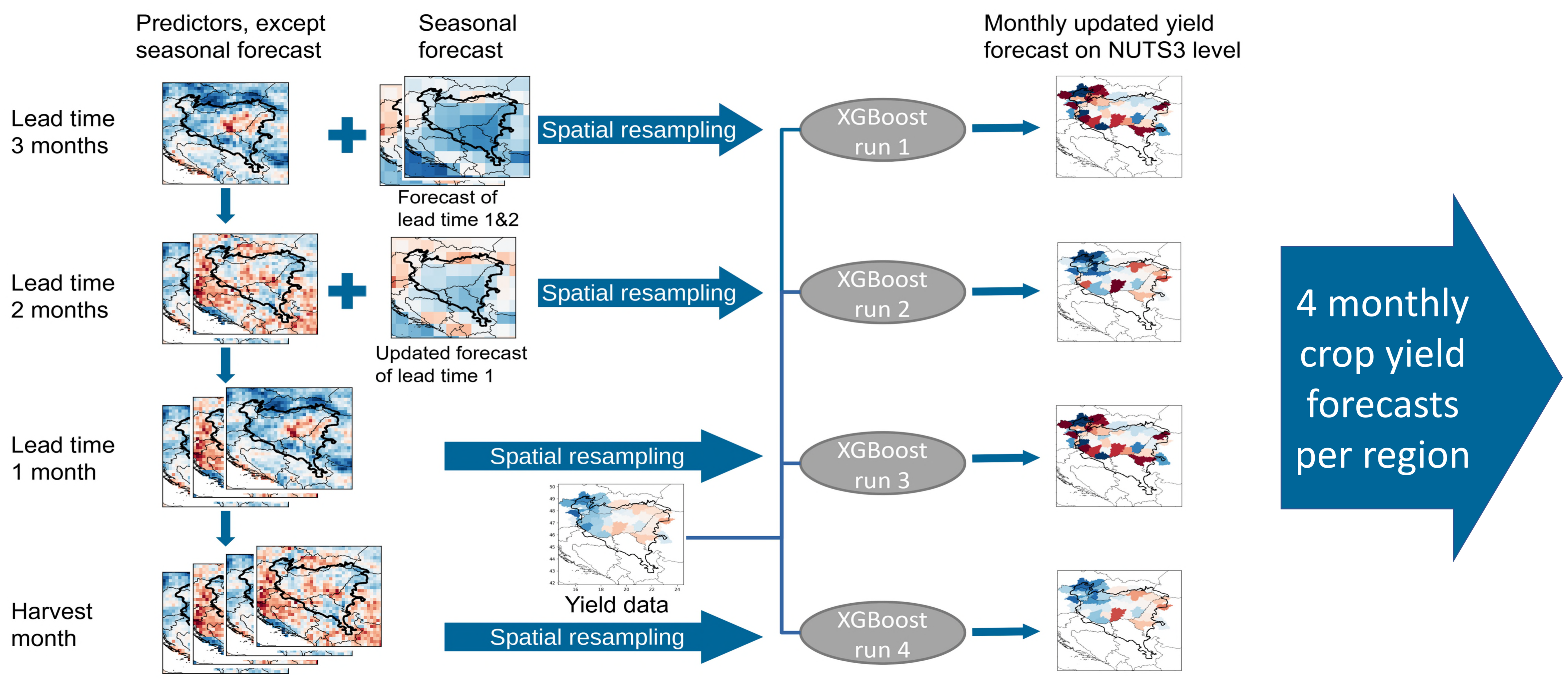

Fig. 3: Logic of the crop yield forecasting system

Key findings

- Wheat and maize crops can be forecasted around **two months before harvest** with a good performance (Fig. 4 & 5)
- Crop yield losses in years of **severe drought are underestimated by the forecasts** but the model correctly detects crop yield losses (Fig. 4)
- **Good performance** to predict **inter-annual variabilities** of the yields for the districts (Fig. 5)
- **Bad performance** to distinguish crop yields **between regions** within individual years
- **Wheat yields** largely dependent on **temperature; maize yields** on **water availability** (Fig. 6)

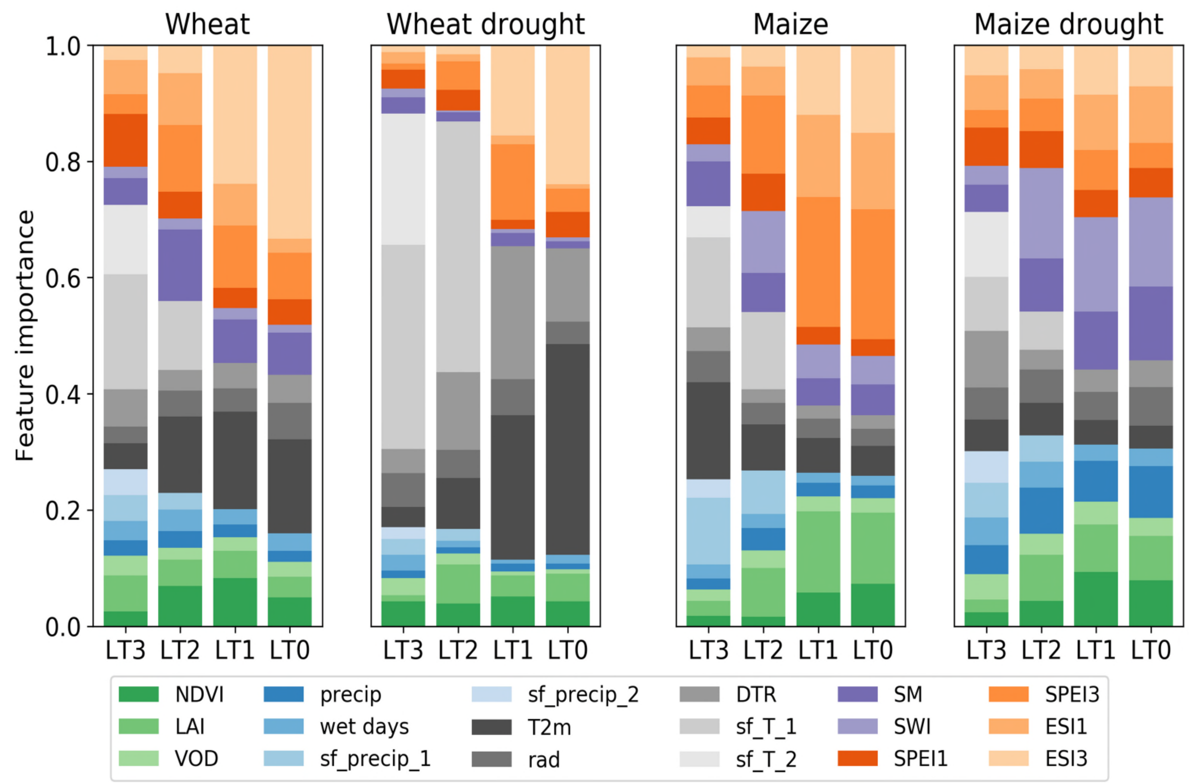

Fig. 6: Cumulative feature importance of the predictors

Key drivers of the wheat forecast model are temperature and ESI; water availability (SPEI/ESI) for maize. In drought years, seasonal temperature forecast has a large impact on the early wheat yield forecasts, while soil moisture is a key predictor for maize. Impacts of predictors are largely dependent on the forecast month.

Understanding the model

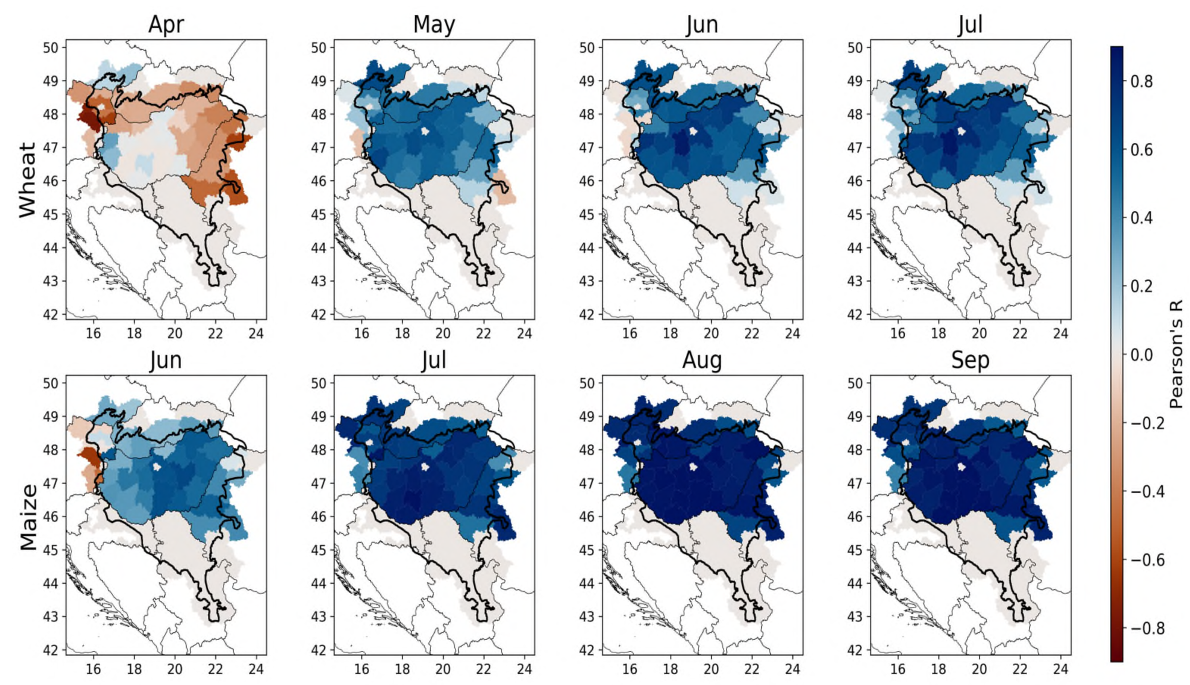

Fig. 5: Correlations of forecasted and observed crop yields

Crop yields of maize and wheat are highly dependent on the conditions in the last two months before harvest. This leads to highest performances of crop yield forecasts in these months. Seasonal weather forecasts enable longer lead times only to some degree.

Validation per region

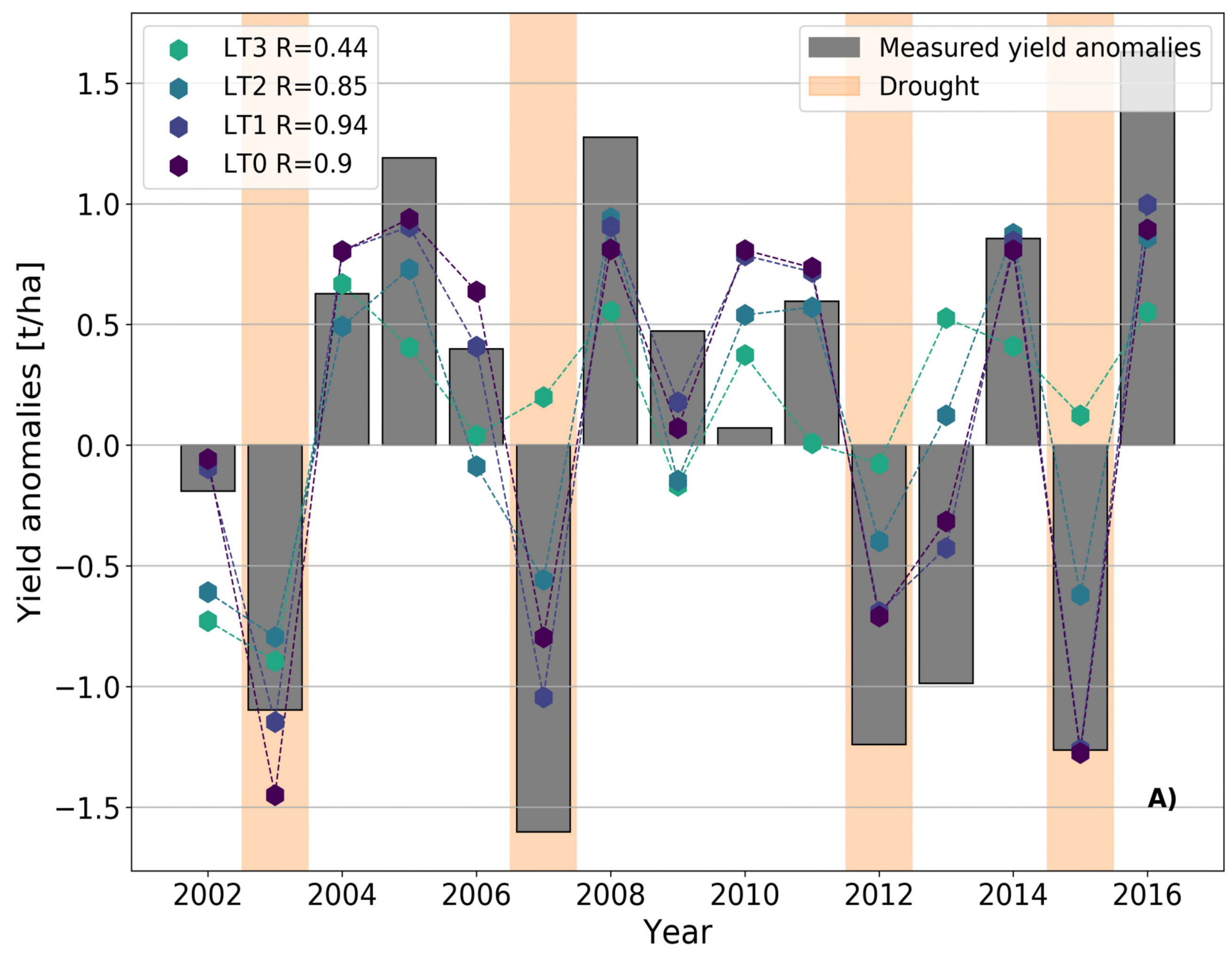

Fig. 4: Forecasted and observed maize yield anomalies per region

Contact
MSc Emanuel Büechi
CLIMERS - GEO - TU WIEN
Wiedner Hauptstrasse 8-10
1040 Vienna, Austria

Acknowledgement
The results presented here have been developed in the framework of the projects "YIPEEO: Yield Prediction and Estimation using Earth Observation" and "DryPan: Novel EO data for improved agricultural drought impact forecasting in the Pannonian basin" funded by the European Space Agency (ESA).

References

[1] E. Bueechi et al. (2023) Crop yield anomaly forecasting in the Pannonian basin using gradient boosting and its performance in years of severe drought, Agricultural and Forest Meteorology, 340, https://doi.org/10.1016/j.agrformet.2023.109596