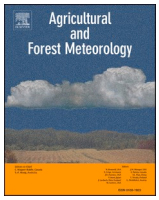
Agricultural and Forest Meteorology 340 (2023) 109596

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet



Crop yield anomaly forecasting in the Pannonian basin using gradient boosting and its performance in years of severe drought

E. Bueechi [a,*], M. Fischer [b], L. Crocetti [c], M. Trnka [b], A. Grlj [d], L. Zappa [a], W. Dorigo [a]

[a] Department of Geodesy and Geoinformation, TU Wien, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria
[b] Global Change Research Institute CAS, CzechGlobe, Bělidla 986/4a, 603 00 Brno, Czechia
[c] Institute of Geodesy and Photogrammetry, ETH Zurich, Robert-Gnehm-Weg 15, 8093 Zurich, Switzerland
[d] SPACE-SI, Slovenian Centre of Excellence for Space Sciences and Technologies, Aškerčeva 12, 1000 Ljubljana, Slovenia

ARTICLE INFO

Keywords:
Crop yield forecast
Remote sensing
Machine learning
XGBoost
Drought

ABSTRACT

The increasing frequency and intensity of severe droughts over recent decades have led to substantial crop yield losses in the Pannonian Basin in southeastern Europe. Their socioeconomic consequences can be minimized by accurate crop yield forecasts, but such forecasts often underestimate the impact of severe droughts on crop yields. We developed a gradient-boosting-based crop yield anomaly forecasting system for the Pannonian Basin and examined its performance, with a focus on drought years. Winter wheat and maize yield anomalies are forecasted for 42 regions in the Pannonian Basin using predictor datasets from Earth observation and reanalysis describing vegetation state, weather, and soil moisture conditions.

Our results show that crop yield anomaly estimates in the two months preceding harvest have better performance (maize errors 14–17%, wheat 13–14%) than earlier in the year (maize errors 21%, wheat 17%). The forecast models can satisfactorily capture the interannual yield anomalies, but spatial yield variability is only partially reproduced. In years of severe drought, the wheat model performs better than under average conditions with errors below 12%. The errors of the maize forecasts in drought years are larger than average forecast skill: 31% two months ahead and 20% one month ahead. However, for both crops the yield losses remain underestimated by the forecasts in severe drought years. The feature importance analysis shows that during the last two months before harvest, wheat yield anomalies are controlled by temperature and evaporation and maize by the combined effects of temperature and water availability as expressed by several drought indices. In severe drought years, during the two months before harvest the seasonal temperature forecast becomes the most important predictor for the wheat forecasts and soil moisture for the maize model. Overall, this study provides in-depth insights into the impact of droughts on crop yield forecasts in the Pannonian Basin.

1. Introduction

In recent decades, droughts have heavily affected agricultural production in the Pannonian Basin, a lowland area in southeastern Europe (e.g. Crocetti et al., 2020). Agriculture in this region is particularly vulnerable to droughts, as the majority of cultivated land is rain-fed (Crocetti et al., 2020). The already challenging conditions for crop production are expected to worsen due to climate change (Kis et al., 2020; Nistor et al., 2017; Zhu et al., 2021) and the area is even regarded as the region with the most negative impacts of climate change on crop production in Europe (Olesen et al., 2011). A potential tool to support the adaptation to these challenging circumstances is crop yield forecasting. This has proven being a vital method to minimize socio-economic impacts of crop losses (Ceglar et al., 2018) by allowing mitigation measures such as compensation planning, effective water management, or if known early enough, even sowing planning for more drought-resistant crops (Li et al., 2021; Udmale et al., 2014).

Crop yield forecasts have been implemented using machine learning (Filippi et al., 2019; Kern et al., 2018; Kogan et al., 2013; Peng et al., 2018; Potopová et al., 2020) or process-based crop yield models (Donohue et al., 2018; Pagani et al., 2019; Zhuo et al., 2019). While machine learning has advantages regarding flexibility, simplicity, and efficiency (Feng et al., 2020; Leng and Hall 2020) process-based models have their strengths regarding their lower use of training data and their

* Corresponding author.
E-mail address: Emanuel.bueechi@geo.tuwien.ac.at (E. Bueechi).

https://doi.org/10.1016/j.agrformet.2023.109596
Received 23 December 2022; Received in revised form 10 May 2023; Accepted 5 July 2023
Available online 11 July 2023

ability to account for interactions of the predictors (Droutsas et al., 2022; Zhang et al., 2022). Process-based models often showed to not optimally simulate the effects of extreme weather events (Eitzinger et al., 2013; Barlow et al., 2015; Feng et al., 2020; Song and Jin 2020) and proper parameter tuning requires extensive training data to reflect various conditions (Akhavizadegan et al., 2021). Machine learning models can provide reasonable results during the impact of extreme weather events, as long as they are properly trained (Droutsas et al., 2022). An optimal training does not only require enough data during extreme events (O'Gorman and Dwyer 2018), but meaningful predictors too. Key predictors for most machine learning-based crop yield forecast models are climatological drivers, using data from in situ measurements (Kogan et al., 2013; Potopová et al., 2020) or meteorological reanalysis (Kern et al., 2018; Pagani et al., 2019). An important additional source of input to crop yield models is Earth observation (EO). With its diversity of spectral ranges, EO data provides a unique source of information about the state of crops and the conditions they are facing (Li et al., 2022; Pagani et al., 2019; Zhuo et al., 2019). Useful crop yield forecasts can be obtained around two months before harvest (Li et al., 2022; Pagani et al., 2019; Zhuo et al., 2019) but most forecasting models tend to underestimate the impacts of severe droughts on crop yield (Bussay et al., 2015; Kang et al., 2020; Pagani et al., 2017; van der Velde et al., 2018). Accurate modeling of yield anomalies in drought years is of particular importance because crop losses in these years can be substantial and may have significant economic consequences (Guarin et al., 2020). Mathieu & Aires (2018) and Potopová et al. (2020) improved crop yield forecasts in severe drought years by including the Standardized Precipitation-Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010) as an explanatory variable of drought conditions in their models. A challenge for crop yield anomaly forecasting is an intensification of drought conditions after the moment of the forecast, which can still significantly reduce the forecast performance (Bognar et al., 2011). To mitigate this, seasonal weather forecasts can help by providing information on how the weather conditions are expected to change, and they have already proven their capability to improve crop yield forecasts in general (Filippi et al., 2019; Peng et al., 2018). Hence, a proficient crop yield forecast model for drought conditions requires the combination of a comprehensive set of predictors describing crop, soil, and climatic conditions, including drought indices, EO data, and seasonal weather forecast data. To our knowledge, such combination has not yet been reported in literature to forecast drought-related yield losses.

Therefore, this study aims to fill this gap by developing machine learning-based crop yield forecast models for the Pannonian Basin and testing their reliability, with a focus on drought years. Given the widespread cultivation of winter wheat and maize in the Pannonian Basin and their economic relevance for the region, we constrain our analysis to these two crops. More specifically, we want to answer to the following questions:

How does a machine learning-based crop yield forecasting system in the Pannonian Basin perform in normal years and in severe drought years?

How do various explanatory variables contribute to the model performance at different times during the growing season, and for different crops?

Answers to these questions could guide mitigation measures addressing the socioeconomic impact of large crop yield losses, and at the same time improve our understanding of drivers and processes affecting crop yield variability.

2. Study area

The Pannonian Basin covers the central section of the Danube Basin. With a total area of 445,000 km^2, it extends over multiple countries in central and southeastern Europe (Hungary, Romania, Serbia, Austria, Czech Republic, Slovakia, Ukraine, Bosnia & Herzegovina, Croatia, Slovenia, and Austria) and is surrounded by the Alps in the west, the Dinaric Alps in the south, and the Carpathians in the north and east (Lukić et al., 2019). The Pannonian Basin has a continental climate with warm to hot summers and cold and snowy winters (Mohammed et al., 2022), which corresponds to the classes Cfa and Cfb of the Köppen-Geiger classification (Kottek et al., 2006). The dominant land use type of the Pannonian Basin is agriculture, with croplands covering more than 70% of its area (Crocetti et al., 2020), and of which only a small fraction is irrigated (Bussay et al., 2015). Mean annual precipitation ranges from 400 to 600 mm (Nistor et al., 2017), equally distributed over all months, with a small peak in summer (Crocetti et al., 2020; Jakubinski et al., 2019). Especially in summer months, precipitation is characterized by large inter-annual variability, e.g., rainfall in July can range from close to 0 to 150 mm (Nagy et al., 2018). In combination with the relatively high annual potential evapotranspiration rates of 586 to 739 mm in the last three decades (Nistor et al., 2017) these varying precipitation rates in summer often lead to a shortage of water availability (Nagy et al., 2018). The Pannonian Basin has experienced several severe drought episodes over the last two decades, i.e., in 2003, 2007, 2012, 2015, 2017, and 2018 (Alsafadi et al., 2020; Crocetti et al., 2020; Trnka et al., 2020). These years are characterized as severe drought years by crop yield losses of up to 40% (Nagy et al., 2018) and outstanding numbers of reported drought impacts (Crocetti et al., 2020).

3. Data and methods

3.1. Crop yield data

Crop yield data on the district level, i.e., Nomenclature of Units for Territorial Statistics level 3 (NUTS3) (Fig. 1) have been provided by the national statistical offices of the individual countries (Potopová et al., 2023). The dataset consists of yearly yields, expressed in t/ha, for winter wheat, spring barley, potatoes, sugar beet, and maize (Jakubínský et al., 2019) for the period 2000–2017. However, due to gaps in the data, only the period 2002–2016 is considered in this study. We focus on winter wheat and maize since these are the dominant crop types present in the region (Bognar et al., 2017; Kern et al., 2018). Furthermore, we excluded districts with less than ten years of data available. Besides, districts where the given crop (either winter wheat or maize) was cultivated over an area smaller than 1000 ha in a given year were removed from further analysis. The quality control of the yield data included a consistency check with data at a higher spatial aggregation level. The higher-level data includes FAO yield data from the FAOStat database (https://www.fao.org/faostat/en/#data/QCL) on the country and Eurostat data (https://ec.europa.eu/eurostat/web/agriculture/data/database) on NUTS2 level for countries within the EU. It was evaluated if the interannual variability of the yield data at NUTS3 level is in line with the FAOStat and Eurostat data. In addition, the quality control included a consistency check of the neighboring areas and testing for deviations from the long-term mean. For the former it was checked if adjacent NUTS3 areas are not significantly deviating from each other. For the latter the crop yields of each NUTS3 area were compared with their individual long-term mean. All the yield reports at the NUTS3 level passed these quality checks.

Typical harvest periods in the Pannonian Basin are beginning of July for winter wheat (Kern et al., 2018), and September for maize (Bussay et al., 2015). As spatial autocorrelations in the crop yield data can affect the forecasting negatively (Ferraciolli et al., 2019) the spatial autocorrelations in the yield data are quantified by calculating Moranś I (Table 1). Moranś I compares the values of individual regions with the adjacent regions. It ranges from −1 (perfectly dispersed) to 1 (perfectly clustered) (Upton and Cook 2014).

3.2. Predictor data

A wide range of predictors is used to capture the wide range of meteorological conditions and the status of the crops (Table 2). The

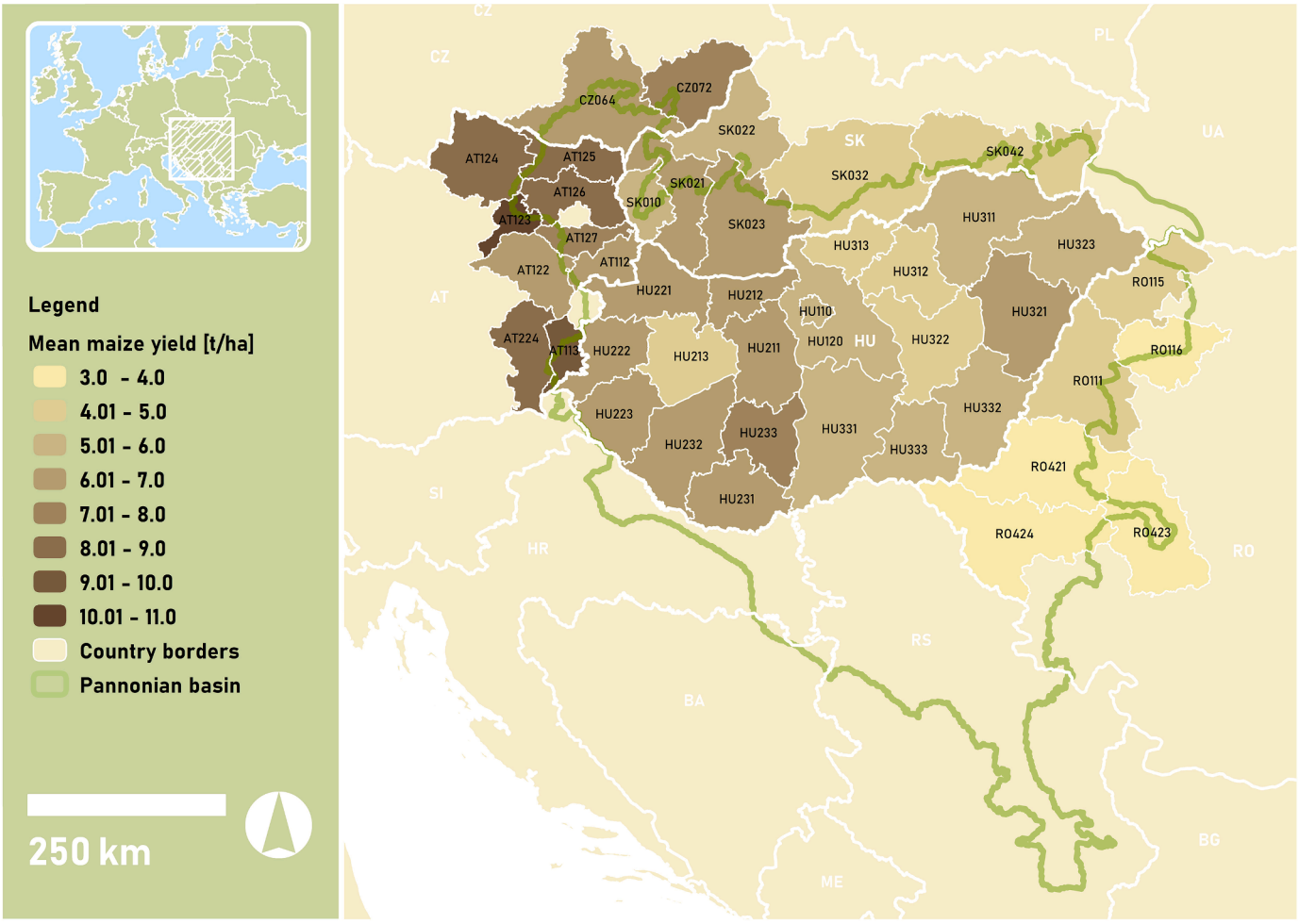


Fig. 1. Overview of the study area and the considered NUTS3 level regions, showing the mean maize yield from 2002 to 2016 per NUTS3 level region [2-column, color].

Table 1
Characteristics of observed yield data, showing the median crop yield, the spatial variance, which depicts the variance of the crop yields between the regions per year (median of all years is shown), temporal variance showing the variance of the crops per region between the years (median of all regions is shown), and Moranś I shows the spatial autocorrelation of the crop yields (median of all years is shown).

Crop	Wheat	Maize
Median crop yield [t/ha]	4.26	6.18
Spatial variance [t/ha]	0.53	2.05
Temporal variance [t/ha]	0.62	2.46
Median Moranś I	0.39	0.55

selection of the datasets to be used as predictor was largely motivated by the availability of long-term records (i.e., covering the period 2002–2016).

3.2.1. Canopy status

The canopy status is described by Leaf Area Index (LAI), Normalized Difference Vegetation Index (NDVI), and Vegetation Optical Depth (VOD). These three variables provide complementary information (Jones et al., 2011) and are used to get a comprehensive representation of the status of the crops: NDVI is sensitive primarily to green leaf biomass and its area (Tucker, 1979), LAI is a proxy for the leaf surface area that is available to photosynthesis (Norman and Jarvis 1975), and VOD shows the water content and biomass of the canopy (Moesinger et al., 2020). The NDVI and LAI datasets are provided by Copernicus Global Land Service (CGLS) and have a spatial resolution of 1 km and a temporal sampling of 10 days (CGLS, 2020). VOD Ku-band data from the VOD Climate Archive is used as it provides a unique long time series of VOD retrievals from multiple sensors (Moesinger et al., 2020). In addition, it is independent from optical sensors and provides more regular information than the NDVI and LAI datasets (Vreugdenhil et al., 2022), i.e., approximately once per day. Ku-Band data is chosen as it reflects the fine structure of crops well (Chen et al., 2020; Teubner et al., 2018).

3.2.2. Meteorological data and drought indicators

Key meteorological drivers of crop growth are temperature, radiation, and water availability (Nagy et al., 2018; Papagiannopoulou et al., 2017; WMO 2010). On large spatial scales such information can be obtained from EO, reanalysis and interpolated in situ measurements (Cornes et al., 2018; Hersbach et al., 2020; Wan 2008). The main advantage of in situ data is its higher accuracy in data-rich areas (Bandhauer et al., 2021), while reanalysis and EO observations offer valuable data over data-poor regions and are generally available with short temporal lags. In this study, data from all three sources are used to combine their advantages. Daily precipitation and air temperature are obtained from E-OBS v20.0 gridded in situ data, because of the relatively high station density in the Pannonian Basin (Cornes et al., 2018). E-OBS provides meteorological variables with a spatial resolution of 0.25° based on the interpolation of several thousands of in situ stations over Europe (Cornes et al., 2018). From the daily precipitation, the fraction of wet days (daily precipitation $>$ 1 mm) per month is calculated to gain

Table 2
Summary of used predictor datasets.

Dataset	Name	Source	Spatial Resolution	Temporal Resolution
Yield data			NUTS3	yearly
Canopy status				
VOD Ku-Band	VOD	VODCA (Moesinger et al., 2020)	0.25°	daily
NDVI	NDVI	CGLS (CGLS, 2020)	0.01°	10-daily
Leaf Area Index	LAI	CGLS (CGLS, 2020)	0.01°	10-daily
Meteorological variables				
Precipitation	precip	E-OBS (Cornes et al., 2018)	0.25°	daily
Fraction of wet days	wet_days	E-OBS (Cornes et al., 2018)	0.25°	monthly
Seasonal forecast precipitation	sf_precip	ECMWF (Johnson et al., 2019)	1°	monthly
Air temperature	T2m	E-OBS (Cornes et al., 2018)	0.25°	daily
Surface net solar radiation	rad	ERA5 (Hersbach et al., 2020)	0.25°	daily
Diurnal temperature range	DTR	ERA5-Land (Munoz-Stabater et al., 2021)	0.1°	daily
Seasonal forecast temperature	sf_T	ECMWF (Johnson et al., 2019)	1°	monthly
Drought indices				
SPEI (1 month)	SPEI1	Based on ERA5 (Hersbach et al., 2020)	0.25°	monthly
SPEI (3 months)	SPEI3	Based on ERA5 (Hersbach et al., 2020)	0.25°	monthly
ESI (1 month)	ESI1	Based on MODIS (Anderson et al., 2011)	0.05°	weekly
ESI (3 months)	ESI3	Based on MODIS (Anderson et al., 2011)	0.05°	weekly
Soil water availability				
Soil Moisture	SM	ESA CCI (Gruber et al., 2019)	0.25°	daily
Soil Water Index	SWI	Based on ESA CCI (Gruber et al., 2019)	0.25°	daily

information about the distribution of precipitation.

Hourly temperature from ERA5-Land reanalysis (Munoz-Sabater et al., 2021) is used to compute the diurnal temperature range (DTR), i.e., the difference between daily minimum and maximum temperatures. Hernandez-Barrera et al. (2016) indicated that the DTR can have a larger impact on crop yield than daily minimum or maximum temperatures alone.

Incoming shortwave radiation, obtained from ERA5 reanalysis (Hersbach et al., 2020), is used as an individual predictor as well as in combination with other variables, i.e., air temperature, dewpoint temperature, wind speed, and precipitation, to compute SPEI. For consistency, all variables used to compute SPEI are taken from ERA5. SPEI is used as an indicator of the drought intensity. It determines the climatic water balance as the difference between the precipitation and the computed reference evapotranspiration (Allen et al., 1998). SPEI is calculated for the periods of one (SPEI1) and three months (SPEI3), to consider the short-term variations as well as conditions representing the cumulative effects of water availability during longer parts of the growing cycle. It is computed using the R-package SPEI (Vicente-Serrano et al., 2010).

Also the Evaporative Stress Index (ESI) is used as a drought indicator. ESI is well-suited for early warning of flash drought development (Otkin et al., 2013; McEvoy et al., 2016). In addition, it has been shown that ESI can provide key information in the Pannonian Basin to assess drought impacts on wheat yields (Jurečka et al., 2021). ESI is based on the actual and reference evapotranspiration (Allen et al., 1998) and quantifies the standardized anomalies of their ratio (Anderson et al., 2011). For this study, it is calculated using the Atmosphere-Land Exchange Inverse (ALEXI) model driven by day/night land surface temperatures from MODIS (Hain and Anderson, 2017). The anomaly is computed for windows with a width of one month (ESI1) and three months (ESI3).

Seasonal forecasts of mean air temperature (2 m) and total monthly precipitation are used as estimates of meteorological conditions beyond the forecast date. For this, the seasonal forecasts from the European center for Medium-Range Weather Forecasts (ECMWF) are used (Johnson et al., 2019), which have been shown beneficial for crop yield and drought forecasting (Ceglar and Toreti 2021; Portele et al., 2021). These forecasts are based on the SEAS5 model and provide forecasts for lead times between six hours and seven months. Seasonal forecasts exist since 2017 with hindcasts from 1993 to 2016 (Johnson et al., 2019). For this study, the hindcasts for lead times of one and two months are used. A validation study of the forecasts over Europe showed good performance of the seasonal temperature forecast in the Pannonian Basin (Crespi et al., 2021).For a one-month lead-time the correlations to ERA5 temperature values are up to 0.8 and a mean bias of around 1–2 °C in spring and summer. The precipitation forecasts are less reliable, with mean correlation below 0.2, and considerable biases during summer months (up to 100 mm in the period from June to August) (Crespi et al., 2021).

3.2.3. Soil moisture

EO-based soil moisture datasets are used as an indicator of the water available to the crops. ESA CCI Soil Moisture COMBINED product v7.1 (Dorigo et al., 2021) is used, as it provides consistent quality-controlled data with a unique combination of various active and passive microwave sensors (Dorigo et al., 2017; Gruber et al., 2019). It is globally available for the time span from 1978 to 2021 with a spatial resolution of 0.25° and a temporal resolution of one day (Dorigo et al., 2021). In addition, the Soil Water Index (SWI) is derived from ESA CCI surface soil moisture data by applying time-based filtering following Albergel et al. (2008). This serves as a proxy for root-zone soil moisture, which is the water actually available for root uptake (McElrone et al., 2013).

3.3. Data preparation

The data preprocessing is summarized in Fig. 2. The first step is to calculate detrended anomalies, to minimize the impact of long-term environmental and climatological changes and changes in agricultural management (Eck et al., 2020; García-León et al., 2019; Lu et al., 2017; Mathieu and Aires, 2018). This is done for both the crop yield data and all predictors, except for the drought indices. For those, the detrending is done too, but the anomalies are not calculated, as they are already standardized values (Anderson et al., 2011; Vicente-Serrano et al., 2010). The detrended anomalies are calculated as in Eq. (2) (Papagiannopoulou et al., 2016) before the temporal resampling of the predictors using a linear detrending function from the python package *scipy* (Virtanen et al., 2020).

$$A_t = D_t - S_t \tag{2}$$

with A_t describing the anomaly at time t, D_t the linearly detrended value at that time, and S_t the long-term climatology computed from the detrended data. The latter is calculated as the mean of all years (2002–2016) at the considered month including a moving average window of 30 days. In the case of crop yield data, S_t describes the mean of all crop yield values per location and crop, as there is no seasonality. For predictors with a monthly resolution, i.e., seasonal forecasts, monthly precipitation, and fraction of wet days, the moving average window is not applied.

Finally, spatial and temporal harmonization of the data is applied, as

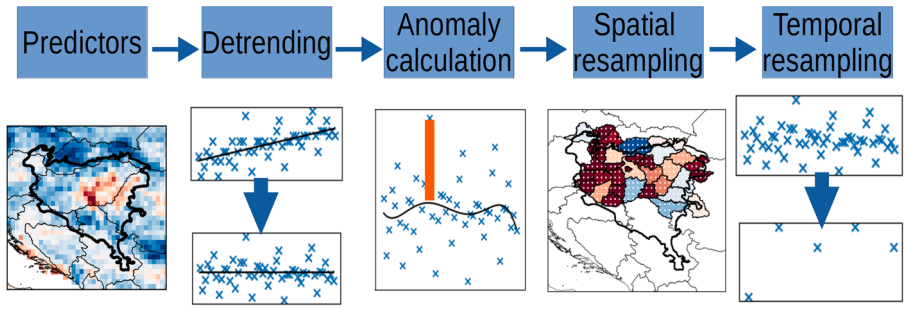


Fig. 2. Flowchart of the data preprocessing steps. The points in the spatial resampling plot represent the center points of the observations with the most coarse resolution (0.25°) [2-column, color].

the model is developed on a monthly basis and for each NUTS3 level region separately. For the spatial harmonization, the individual spatial resolutions of the explanatory variables are resampled to the areas of the NUTS3 level. Hence, all pixels of which the center coordinate lies within an individual region are averaged. For the temporal harmonization, monthly means of the detrended anomalies are calculated, except for precipitation and the variables that already have a monthly resolution (i.e., wet days, seasonal forecasts, and SPEI).

3.4. Correlation analysis between predictors and crop yield anomalies

First, the Spearman's rank correlation is calculated to assess how well the predictors correlate to the crop yield anomalies at various lead times. The correlations are calculated for each month at NUTS3 level. We further stratify our analysis for drought and non-drought years in order to compare how extreme conditions can potentially affect such correlations. The 95%-confidence intervals for the correlations are calculated using Fisher transformation (Upton and Cook 2014).

3.5. Extreme gradient boosting

Extreme gradient boosting (XGBoost) is a machine learning algorithm that uses scalable end-to-end tree boosting (Chen and Guestrin 2016). It is used in this study, as it outperforms other machine learning techniques for crop yield forecasting (Kang et al., 2020) and provides information about the importance of the predictors (Shahhosseini et al., 2021). Additionally, tree boosting methods can handle multicollinearity of predictors (Kotsiantis 2011; Piramuthu 2008) which is a major concern given the large number of predictors employed in this study. Nevertheless, we test two methods for reducing the number of predictors, i.e., remove the ones with high cross-correlations, thus reducing model complexity (Chapter 3.5.1.).

3.5.1. Feature elimination

Two methods are chosen to reduce the number of predictors due to possible cross-correlations between the input variables: Variance inflation factor (VIF) (Craney and Surles, 2002) and by assessing the cross-correlations between the predictors. Feature elimination based on VIF follows a simple logic: the model is run using all predictors and for each predictor the VIF is calculated. All predictors with a VIF exceeding a threshold are eliminated and the model is rerun with the remaining predictors. Again, the VIF is calculated for each predictor at all months and the ones exceeding the threshold are again eliminated. This procedure is repeated until all predictors get VIFs below a predefined threshold (Craney and Surles, 2002). The thresholds used in this study are 10 and 5, which are typical cut-off points (Craney and Surles, 2002).

Feature elimination based on the cross-correlation of the predictors is applied by removing features with correlations above certain thresholds. As for the VIF method, various thresholds are used: Pearsońs R of 0.8, 0.7, 0.6, 0.5. We do one model run for each of these thresholds excluding the concerned features. The feature elimination is done by excluding features that with absolute values of the cross-correlations exceeding these threshold. When two features have cross-correlations above these thresholds, the feature with the overall lower correlation to the other features is kept and the other eliminated. This rule is not applied in case that one feature has correlations exceeding the used threshold to more than one other predictor. In that case, the feature with the fewer correlations above the threshold is eliminated.

3.5.2. Calibration and validation

For both, wheat and maize, we setup the model at four different times starting with a lead time of three months (Fig. 3). The models are calibrated for each prediction month separately. Each month the models are recalibrated by adding new data (LT2, LT1, and LT0 for the harvest month) (Fig. 3). The models are implemented using the Python package *xgboost* (Chen and Guestrin 2016). Automated hyperparameter-tuning is applied to improve the performance of the model, by tuning the parameters max_*depth, learning_rate, n_estimator, colsample_bytree* (Table 3). The hyperparameters are tuned once per crop type using all predictors of all months. For the hyperparameter-tuning a random 20-fold cross validation is applied.

To validate the models, the datasets are divided into a training and a testing set. To get a robust assessment of the models' accuracy, training and test sets must be independent (Rebala et al., 2019). However, environmental data might be correlated (Mathieu and Aires, 2018): neighbouring regions (in this study NUTS3 regions) might be characterized by varying degrees of spatial autocorrelation; temporally, environmental conditions might persist in two or more consecutive years, e.g., drought impacts can last several years (Bales et al., 2018). Yet, it is generally accepted that the weather of two consecutive years is not heavily correlated (Mathieu and Aires, 2018). To reduce the impact of the spatial correlation on the validation of the model, many authors used entire years of all regions for the validation of their model (Johnson et al., 2016; Kern et al., 2018; Kogan et al., 2013; Mathieu and Aires, 2018), while others used cross-validation (CV) to improve the reliability of the validation (Filippi et al., 2019; Gómez et al., 2019; Johnson et al., 2016; Kogan et al., 2013). In this study, a leave-one-year-out CV is applied to ensure a sufficient amount of training data.

Yearly crop yield estimates obtained through CV are then examined in more detail. In particular, we assess: (1) the overall performance on NUTS3 level, which includes forecasts for all regions and years; (2) the performance to forecast spatial crop yield variability, where the crop yield forecasts of all NUTS3 regions are validated on a yearly basis, to evaluate the models' skill to reproduce spatial patterns of crop yield anomalies; (3) the regional performance, which compares time series of forecasted and observed crop yields per NUTS3 region, to evaluate the models' ability to capture temporal dynamics of yield anomalies; (4) the Pannonian mean forecast (PB means), showing if the forecasted yearly mean crop yields of all NUTS3 regions correspond to the observed mean. This allows for verifying if the forecasted small-scale deviations at NUTS3 level have better skill than those obtained over the whole Pannonian Basin. In (1), (2), and (4) drought years are also separately

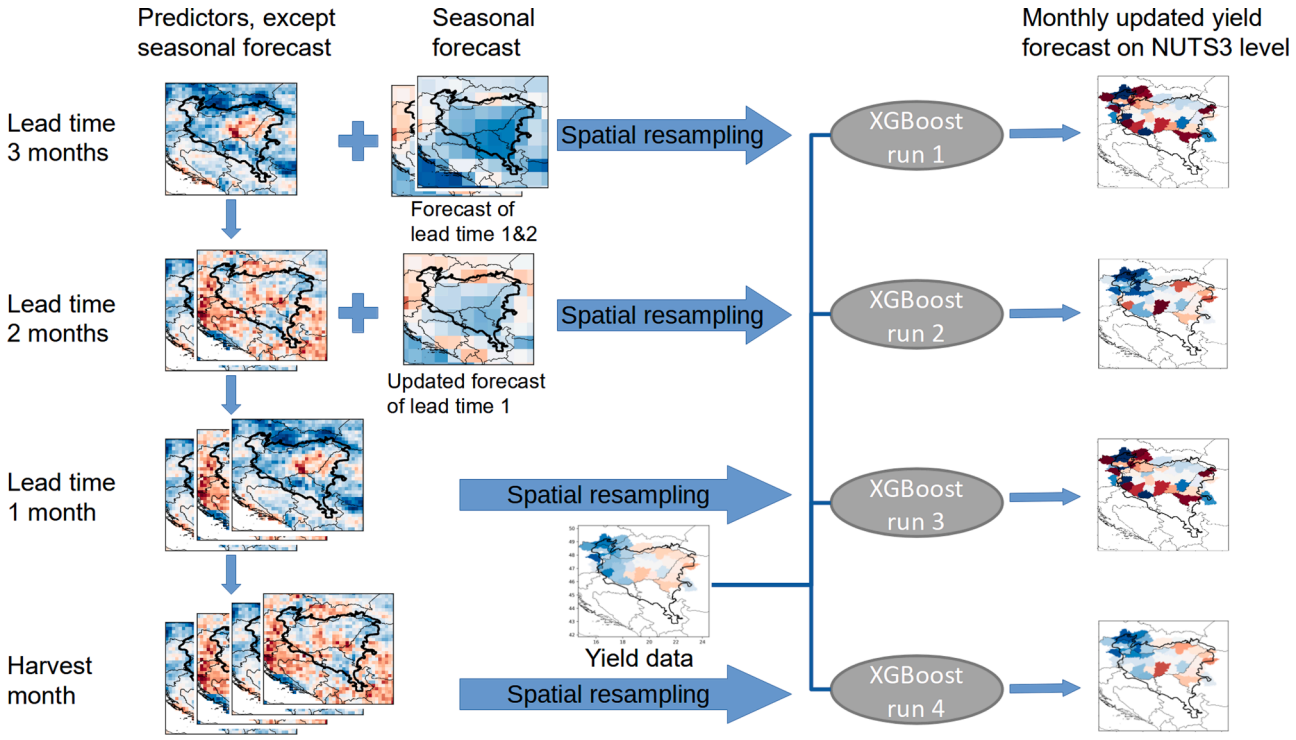


Fig. 3. Logic of the crop yield forecasting system. The initial model run is done with a lead time of three months before harvest. Only data that is available by the end of this month is used for the forecast. Hence, the observations of the meteorological, canopy, and soil moisture data, as well as the seasonal forecasts that are established by then until the harvest month [2-column, color].

Table 3
Used values of the hyperparameters for the tuning and the resulting values per crop type.

Hyperparameter	Tested values	Selected wheat	Selected maize	Definition
max_depth	3, 6, 10	6	10	Describes complexity of model
Learning_rate	0.01, 0.05, 0.1	0.05	0.01	Step size shrinkage
n_estimator	100, 500, 1000	1000	1000	Number of decision trees
colsample_bytree	0.3, 0.7	0.3	0.3	Subsample ratio of columns

analyzed. The years that are considered as drought years are 2003, 2007, 2012, and 2015. Those years are defined as severe drought years in this study based on a drought impact report database by Jakubínský et al. (2019). There are over 30 reported drought impacts in each of those years, which is more than twice the average of around 14 drought impacts per year in the Pannonian Basin (Crocetti et al., 2020). Reported drought impacts have been shown to be highly beneficial for drought monitoring (Bartošová et al., 2022) and allows to define drought years independently from datasets that are used in the model development, such as SPEI.

Model performance is assessed by calculating the Pearson's correlation coefficient and the normalized root mean squared error (NRMSE) between observed and forecasted crop yield anomalies. NRMSE is calculated as the ratio between the RMSE and the mean crop yields of all observations per crop type, and is selected instead of the RMSE because it allows to compare results characterized by different amplitudes. Pearsońs correlation depicts the linear dependencies of the forecast and observations well, but is of limited use when forecasted and observed values are not matched (Li, 2017).

3.5.3. Feature importance

As a final step, it is tested how important the different explanatory variables are for the models at different lead times. The feature importances are based on the average Gini impurity index (Rebala et al., 2019) taken from all the CV runs. An additional analysis is done by checking the feature importance when using only drought years for training the models. Finally, the feature importances obtained for each CV run (i.e., considering different years for training) are compared to assess the consistency of the individual models.

4. Results

4.1. Correlation analysis between predictors and crop yield anomalies

The correlations between predictor variables and winter wheat and maize yield anomalies are shown in Figs. 4 and 5, respectively. Important differences exist between the two crops. For wheat, absolute values of correlations are in most cases smaller than 0.5, and do not show major changes throughout the season. On the contrary, maize yield anomalies show high (absolute values of) correlations with a number of predictors three and two months before harvest. As expected, the correlations between yield anomalies and temperature-related variables (mean temperature, radiation, DTR, and seasonal temperature forecast) are mostly negative. Conversely, proxies of vegetation status (NDVI, LAI, and VOD), moisture availability (precipitation, wet days, seasonal forecast of precipitation, soil moisture, SWI), and drought indices, are generally characterized either by positive or non-significant correlations with crop yield anomalies.

The correlations in drought years (bottom panels in Figs. 4and 5) show less distinct patterns and vary a lot between months. The predictors with the highest correlations to wheat yield anomalies are ESI1 and ESI3 for LT3 and LT2, and soil moisture, LAI, and NDVI for LT2 and LT1 for maize. There are even some unexpected negative correlations between precipitation and wheat yield anomalies for LT3 and LT2. Overall, the correlations are higher for wheat than for maize during

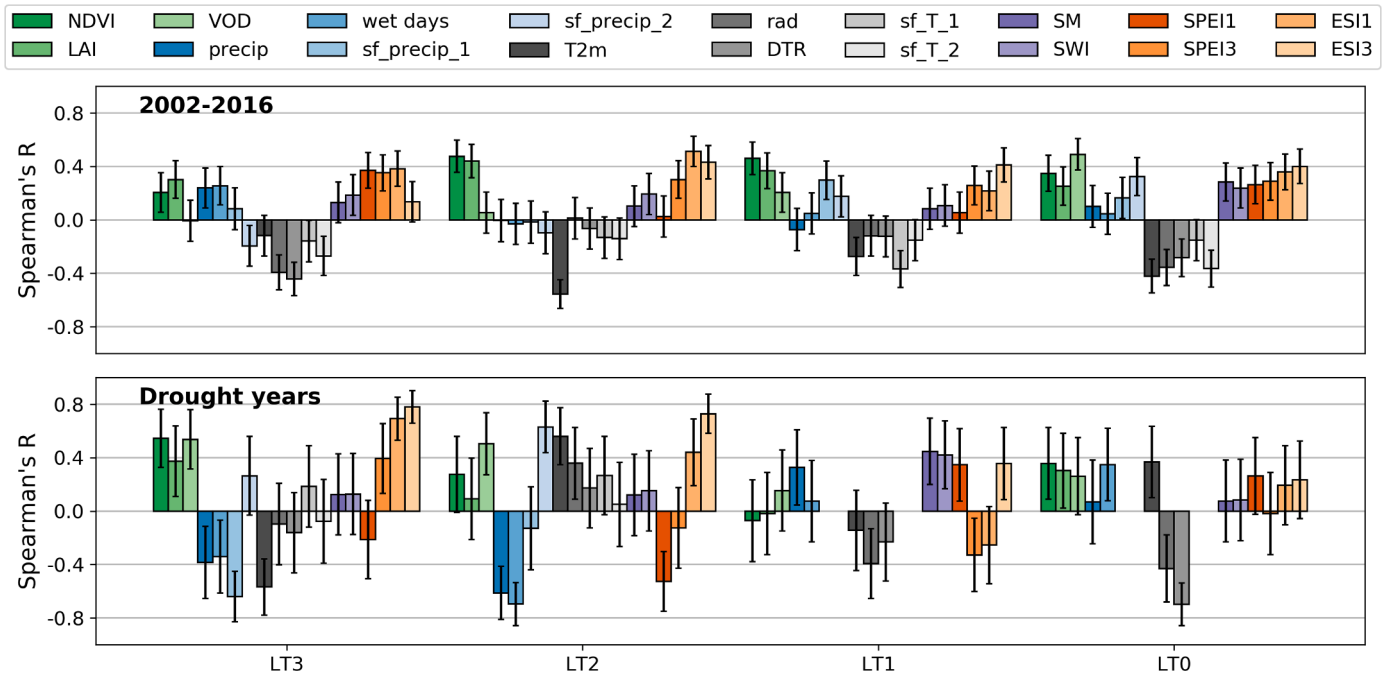


Fig. 4. Correlations of the monthly mean anomalies of the explanatory variables to wheat yield anomalies for all years on top and for the individual drought years below. Green colors show canopy status related variables, blue colors precipitation-related variables (total precipitation, ratio of wet days per month and the seasonal precipitation forecast for one and two months ahead), gray temperature-related predictors (land surface temperature, radiation, diurnal temperature range), violet soil moisture and soil water index, and orange the drought indices [2-column, color].

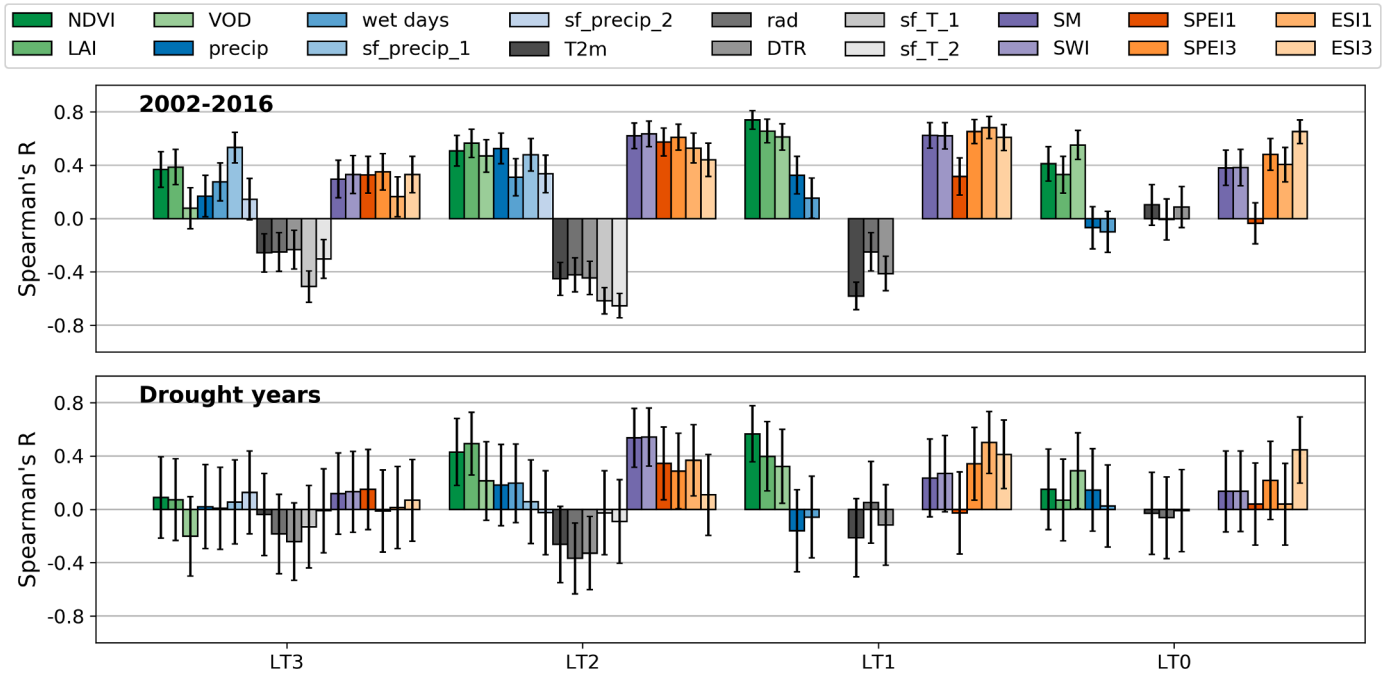


Fig. 5. Same as Fig. 4 for maize [2-column, color].

drought years.

4.2. Validation of crop yield forecasts

Fig. 6 shows the agreement, expressed in terms of Pearson R and NRMSE, between predicted and reference crop yield anomalies per NUTS3 region for different lead times, while Figs. 7A and 8A show results for the yearly average regional forecasts over the Pannonian Basin for wheat and maize, respectively. Both validations show that the best forecasts are obtained from around two months before harvest. NRMSE is below 17% for both crops, while the correlation between predicted and observed yields reach ~0.7 for maize, and 0.43 for wheat. Earlier forecasts, with a lead time of three months, yield considerably lower accuracies (Fig. 6 and Table 4). Indeed, maize shows medium correlations (Pearson's R around 0.4) and high errors (NRMSE around 25%). Early wheat forecasts are characterized by medium errors (NRMSE

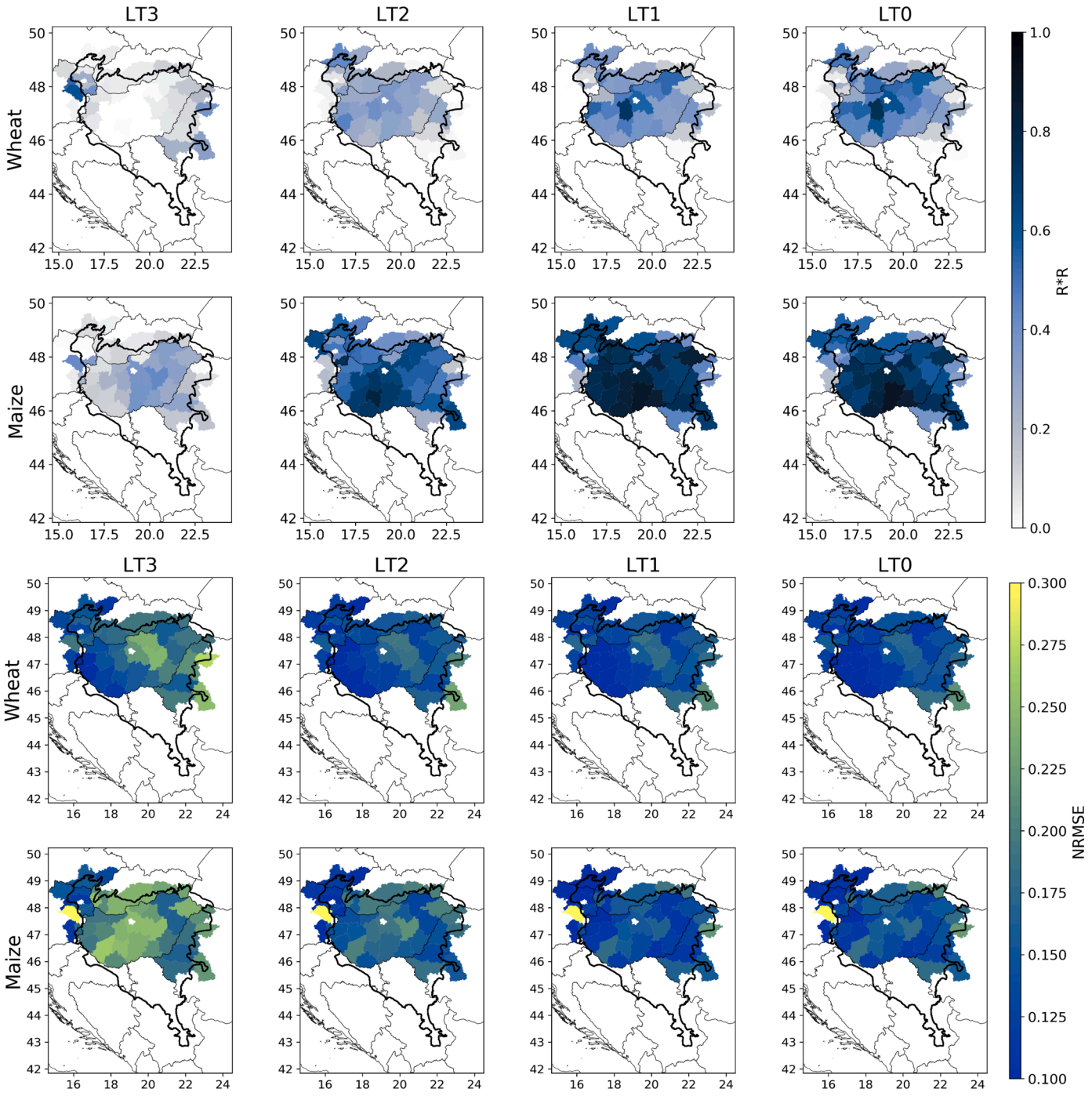


Fig. 6. Validation of the crop yield forecasts per NUTS3 region. The first two rows show the correlation of the predicted and measured crop yield anomalies and the last two show the errors of the forecast (NRMSE) [2-column, color].

around 18%) but low correlations (Pearson's R around 0) (Fig. 6). There is only one outlier with exceptionally high errors for a region in Austria (in the west of the Pannonian Basin) for maize. The remaining regions show a similar range of errors for both maize and wheat. The differences in performance between LT1 and LT0 are only minor. Overall, yield forecasts are more reliable for maize than for winter wheat regardless of the lead time (Table 4).

Negative yearly mean yield anomalies for the entire Pannonian Basin observed during drought years, are generally reproduced by our forecasting models (Figs. 6 and 7). For wheat, the forecasts during drought years have higher correlations than on average and the errors are smaller than 12% from two months before harvest (Table 4). Yet, the mean wheat yield losses of the Pannonian Basin are significantly underestimated (Fig. 7A). Conversely, the amplitude of maize yield anomalies is well captured during dry years, especially in 2003 and 2015 (Fig. 8A). It should also be highlighted that, the model underestimates positive and negative extremes alike. I.e., the range of forecasted wheat and maize yield anomalies is from around −1 to 1, −2 to 2 respectively, while the observed range of the yield anomalies are almost twice as large for both crops. Even though errors during drought years are on average higher than during non-drought years, the difference between errors in drought and non-drought years is decreasing closer to harvest (Table 5).

On a yearly basis, there is a large variability in the correlations between forecasted and observed yield anomalies (Fig. 9D). They range

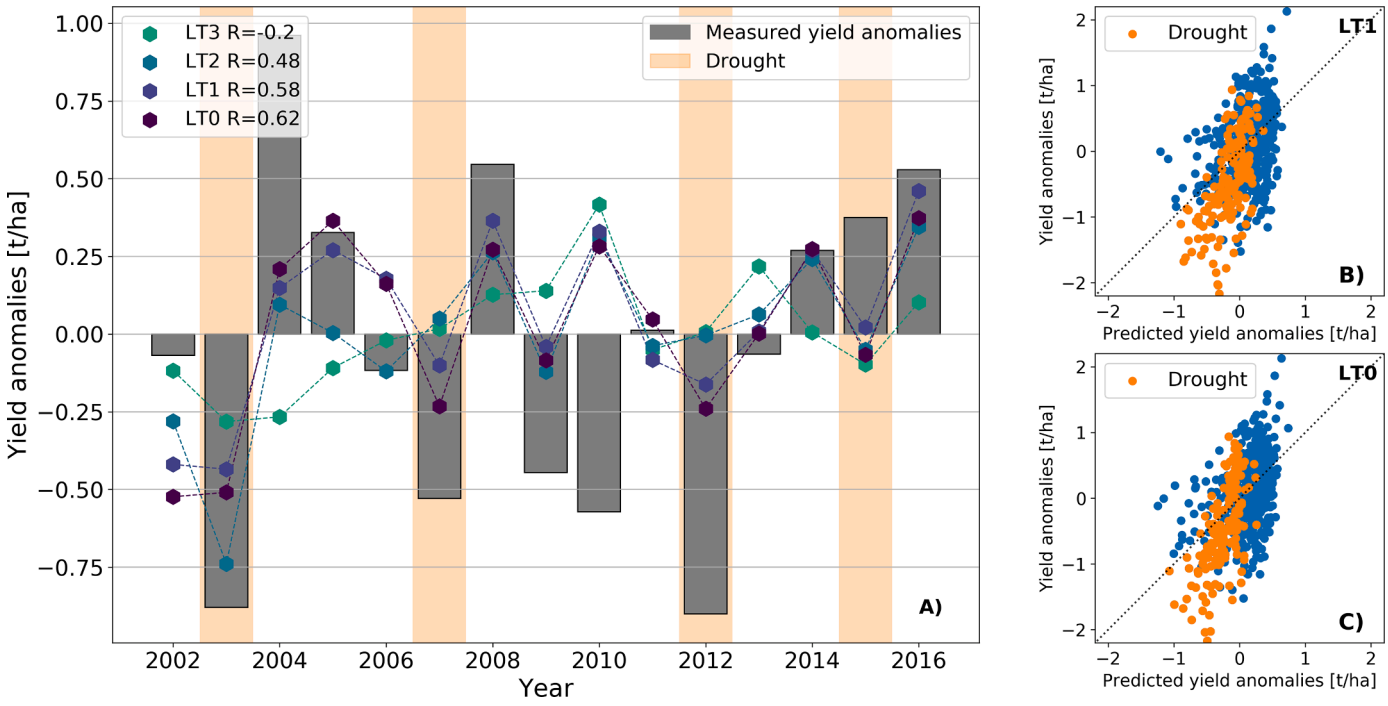


Fig. 7. (A) Measured (bars) and predicted (lines) wheat yield anomalies over entire Pannonian Basin. The color of the line indicates the month at which the forecast was calculated. The correlations of the predicted yearly mean anomalies to the actual ones are printed in the legend. (B) and (C) show all measured yield anomalies for all regions and all years and their forecasts in the last two months before the harvest. Drought years are shown with bright orange background in the bar plot and as orange points in the scatter plot [2-column, color].

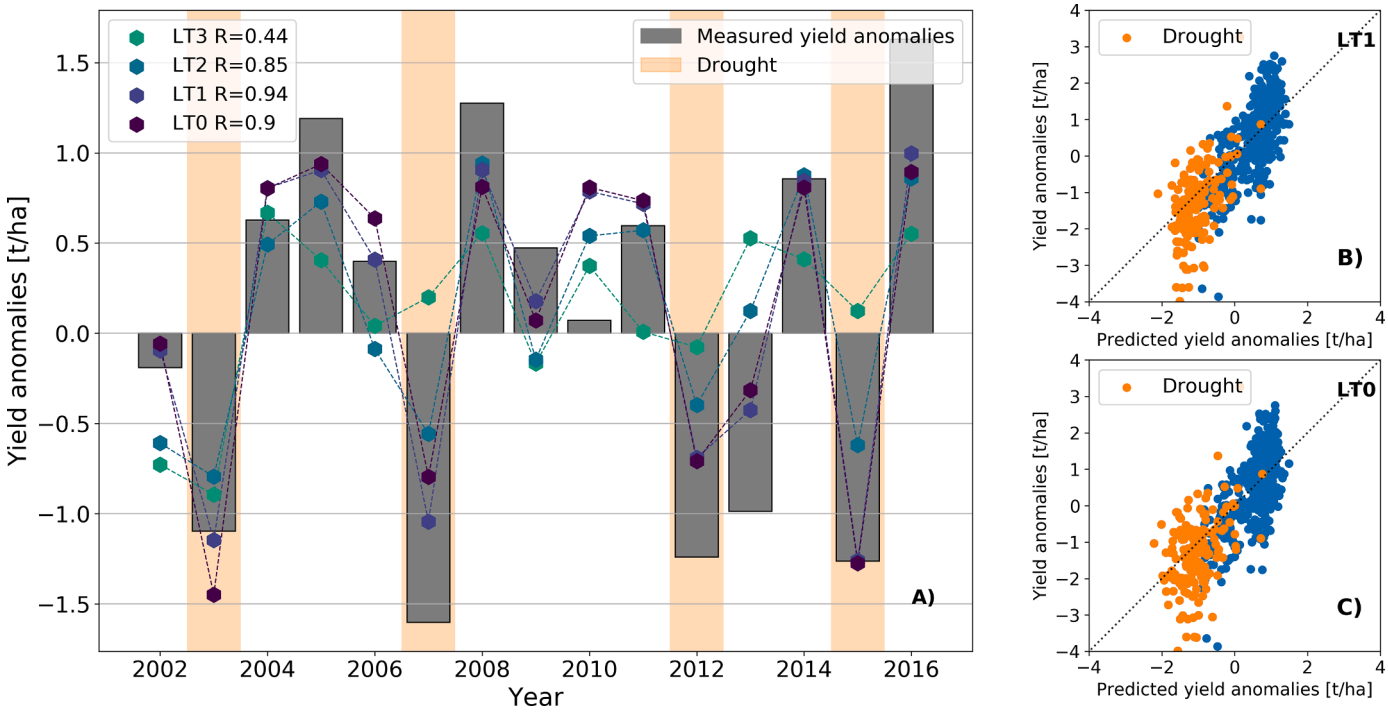


Fig. 8. Same as Fig. 7 for maize [2-column, color].

from around 0 to 0.75 and are in 2004 even strongly negative for maize at LT2. The impact of severe drought years is diverging too. While two drought years (2003 and 2012) are among the years with highest correlations, in the drought years 2007 and 2015 correlations are low.

Fig. 9A–C shows the development of some important explanatory variables (temperature, SPEI3, and NDVI) over the growing season to analyze how this impacts crop yield anomalies. All three variables show high fluctuations within and between the years. Droughts are clearly visible in SPEI3 and the temperature anomalies, whereas the negative NDVI anomalies are not so clearly distinguishable from non-drought years.

Table 4
Forecast validation of the model for different validation techniques as described in Section 4.5.1: overall for the comparison of all regions and years, PB means shows the mean forecasted observed yields over the Pannonian Basin, and Drought for the performance during all drought years. The scale bar for the color coding of the two metrics is shown in the last row.

		Overall	PB means	Drought		Overall	PB means	Drought
maize_LT3	Pearsońs R	0,33	0,44	−0,10	NRMSE	0,21	0,16	0,60
maize_LT2		0,69	0,85	0,39		0,17	0,10	0,31
maize_LT1		0,79	0,94	0,47		0,14	0,06	0,20
maize_LT0		0,76	0,90	0,36		0,15	0,08	0,22
wheat_LT3		−0,17	−0,20	0,07		0,17	0,14	0,16
wheat_LT2		0,43	0,48	0,43		0,14	0,11	0,12
wheat_LT1		0,50	0,58	0,63		0,13	0,10	0,10
wheat_LT0		0,54	0,62	0,64		0,13	0,10	0,09

Table 5
Median absolute error of the forecast of drought and non drought years, and p value of T-test comparing the distributions of the errors during drought and non drought years.

	Median error non-drought [t/ha]	Median error drought [t/ha]	T-test error non-drought and drought [p value]
Wheat_LT2	0.32	0.49	0.00
Wheat_LT1	0.34	0.43	0.01
Wheat_LT0	0.35	0.4	0.08
Maize_LT2	0.65	0.77	0.00
Maize_LT1	0.47	0.6	0.02
Maize_LT0	0.55	0.6	0.01

4.3. Feature elimination

Table 6 shows the results of the different model runs using the two feature elimination techniques. It shows that the model runs using all predictors outperform almost all model runs using feature elimination. Only the wheat model shows significant improvements at LT3 with increasing Pearsońs R and decreasing NRMSE for most model runs. The best feature elimination technique in this study seems to be the ones excluding features with cross-correlations above 0.8. In this case, wet days and SPEI1 are eliminated due to their high correlation to precipitation, and DTR because of its high correlation to radiation. As the improvements are not consistent, i.e. the forecasts for maize at LT3 and LT2 decrease substantially, it is decided not to use any feature elimination. This also has the advantage that no additional uncertainties are introduced by defining a cross-correlation threshold for exclusion.

4.4. Impact of the explanatory variables on the models

Fig. 9 shows the development of the relative importance of predictors throughout the growing period, for all years and drought years separately. For wheat, the variables with the highest importance are air temperature and its seasonal forecast, the drought indices ESI and SPEI, and NDVI. Notably, the importance of these predictors changes throughout the season: for instance, the impact of ESI3 increases strongly from LT3 to LT0, while the contribution of seasonal forecasts and precipitation-related variables decreases. For maize, the explanatory variables with the highest feature importance are SPEI3, temperature, and LAI. The impact of temperature is considerably decreasing after LT3, whereas the importance of SPEI3, ESI1, and ESI3 is increasing towards the harvest date. All these variables show relatively high correlations in the respective months in the correlation analysis (Fig. 5).

For drought years, feature importances are different from all years together. All temperature-related predictors together reach an importance higher than 50% for wheat in all LTs considered. The impact of the drought indices is lower than on average within the last two months before harvest for both crops (Fig. 10). On the other hand, in the wheat forecast model, the seasonal forecast of temperature in the first two months and temperature and DTR in the last two months have high predictive power. For maize, soil moisture, SWI, and precipitation are among the most important features during drought years.

Fig. 11 compares the feature importance of the different model runs during the CV, hence providing insight into the model consistency. Overall, there is good agreement between the feature importance obtained from the different test-train splits carried out through the CV. For wheat, most of the predictors with the high importance are from LT2, while in the maize model most of the features with highest importance (drought indices and LAI) are from LT1.

5. Discussion

5.1. Overall model performance

The maize models tend to perform better than the wheat models. They can capture the interannual yield variability. For most regions high correlations (R>0.7) and low NRMSE between forecasted and observed yields from around two months before harvest have been obtained (Fig. 6). The predictive skill from two months before harvest is in line with the correlation analysis between predictors and crop yields, which also showed a marked increase in correlation two months before harvest (Figs. 4 and 5). Some predictors have already similar large correlations with the crop yield anomalies in LT3 as in LT2. However, according to the feature importance analysis these “early” predictors do not have a large overall impact on the model: for wheat, the most important variable changes between from the seasonal temperature forecasts and soil moisture in LT3 to air temperature and ESI3 at LT1 (Fig. 10). For maize, the most important explanatory variable changes from observed and forecasted temperature in LT3 to soil moisture and SPEI3 in the last two months before harvest (Fig. 10). In both cases, the correlations between these predictors and crop yield anomalies is higher for the same months. Our results indicate that the crop yield anomaly forecasts in the last months before harvest largely depend on water availability for maize, and temperature and ESI3 for wheat. This can be explained by the phenology of the crops. The critical phenological stages are flowering (around 5–8 weeks before the start of harvest) and grain filling (1–5 weeks before the start of harvest) (Bussay et al., 2015). Winter wheat starts growing at a daily mean temperature of around 0 °C while conditions for crop growth become less favourable around daily mean temperature of around 25 °C (McMaster and Wilhelm, 1997). This condition is usually met in spring and early summer. Air temperatures above 30 °C, generally associated with low ESI values, negatively affect wheat yields (Hlavacova, 2017). Maize, on the contrary, mainly depends on soil water availability. Healthy growth of maize requires sufficient soil moisture at rooting depth throughout the growing cycle, but it is most vulnerable to water scarcity during flowering (WMO, 2010), which occurs in the summer months when potential evapotranspiration is normally at its peak.

5.2. Model behavior throughout the season

The Pannonian Basin mean forecasts (Figs. 7A and 8A) confirm that the specific conditions up to two months before harvest can significantly

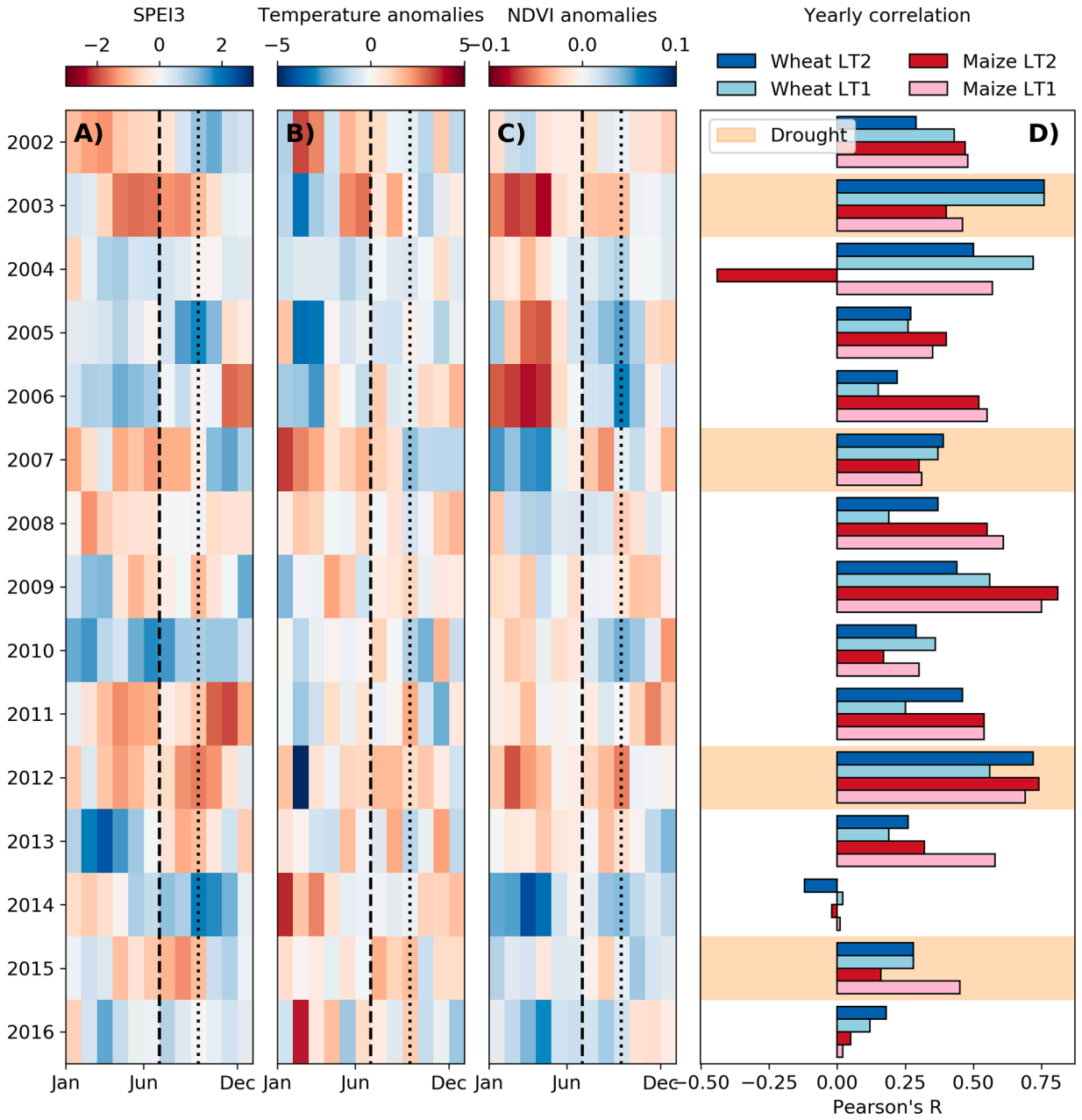


Fig. 9. Development of three explanatory variables during time span of the study (A-C) and the yearly correlations of the forecasted and observed crop yield anomalies of all available NUTS3 regions. The vertical lines in A-C indicate the approximate timing of the harvest – dashed line for wheat and dotted line for maize [2-column, color].

impact the final yield. For example, the beginning of the year and early spring in 2009 is characterized by high SPEI3 and average to low air temperatures, which cause the model to forecast positive wheat yield anomalies. In April and May, the air temperature anomalies become positive and SPEI3 drops (Fig. 9A and 9B) and so does the yield forecast. The subsequent high temperatures and low rainfall, which persisted until June (i.e. 1 month before harvest), caused large wheat yield losses over the Pannonian Basin (Fig. 7). Later in the growing season, i.e. between July and August, SPEI3 and air temperature anomalies return to normal. Consequently, the maize yield models of July (LT2) and August (LT1) forecast higher yields than in June (LT3), but still substantially underestimate the actual maize yield anomalies observed later in the year. Similar patterns can be observed in other years, e.g., in 2013 and 2015. These years start with high SPEI3 and average temperature values and only towards the summer months the conditions are getting adverse for crop growth, showing increasing positive temperature anomalies and lower SPEI3 (Fig. 9A and 9B). Wheat yields are not much impacted by these unfavorable conditions due to the earlier harvest, but maize yields show large negative anomalies. The development of maize in 2002 and 2011, which start very dry, shows that this works also the other way around. SPEI3 is low throughout the first half of the year and increases significantly in July and August (LT2 and LT1 for maize) causing average or even positive maize yield anomalies. Hence, the yield anomalies highly depend on the conditions during the last two months before harvest. This complicates accurate crop yield forecasts with lead times longer than two months. Hence, early season crop yield forecasts need to rely on seasonal weather forecasts which can be highly uncertain (Crespi et al., 2021). Still, including seasonal temperature forecasts to some degree adds value to the maize forecast at LT3, where the forecast shows medium correlations between forecasted and observed maize

Table 6
Validation of the model runs with applied feature elimination techniques. The values show the difference of the correlations and NRMSE values to the original model run using all predictors. The resulting predictors for each of the model runs are summarized in the lower half of the table. VIF_10 and VIF_5 show the used predictors when a Variance Inflation Factor threshold of 10, respectively 5, is applied and Cor_08 to Cor_05 represent the thresholds used for the cross-correlation analysis. E.g. for Cor_07 all predictors with cross-correlations above 0.7 are excluded. The seasonal forecasts are always included. The scale bar for the color coding of the metrics are shown in the row after wheat_LT0 for both metrics.

	Improvements to baseline					
	VIF 10 Δ R	VIF 5 Δ R	Cor_08 Δ R	Cor_07 Δ R	Cor_06 Δ R	Cor_05 Δ R
maize_LT3	0	−0,07	−0,09	−0,02	−0,06	−0,07
maize_LT2	−0,05	−0,08	−0,04	0,01	−0,04	−0,04
maize_LT1	−0,05	−0,03	−0,01	0	−0,02	−0,01
maize_LT0	−0,03	−0,01	0,01	0,01	0	−0,01
wheat_LT3	−0,19	0,15	0,14	0,18	0,2	0,4
wheat_LT2	−0,12	−0,05	0,04	0,04	−0,14	−0,12
wheat_LT1	−0,06	−0,14	0	−0,01	0,03	−0,05
wheat_LT0	−0,14	−0,16	0,01	−0,03	−0,01	−0,03
	Δ NRMSE	Δ NRMSE	Δ NRMSE	Δ NRMSE	Δ NRMSE	Δ NRMSE
maize_LT3	0,01	0,02	0,01	0	0	0
maize_LT2	0	0,01	0	−0,01	0	0
maize_LT1	0,01	0,01	0	0	0,01	0
maize_LT0	0,01	0	−0,01	0	0	0
wheat_LT3	0,03	0,01	0	0	−0,01	−0,02
wheat_LT2	0,01	0,01	−0,01	0	0,01	0,01
wheat_LT1	0,01	0,02	0	0	0	0,01
wheat_LT0	0,01	0,01	0	0	0	0,01

Used predictors.
Reference: all.
VIF 10: NDVI; LAI; VOD; precip; wet days; Diurnal; SM; SWI; ESI1; sf_tg; sf_precip.
VIF 5: NDVI; VOD; Diurnal; SM; ESI1; sf_tg; sf_precip.
Cor_08: NDVI; LAI; VOD; precip; LST; rad; SM; SWI; SPEI3; ESI1; ESI3; sf_tg; sf_precip.
Cor_07: LAI; VOD; LST; SM; SPEI1; SPEI3; ESI1; sf_tg; sf_precip.
Cor_06: LAI; VOD; LST; SWI; SPEI1; ESI1; sf_tg; sf_precip.
Cor_05: LAI; VOD; SPEI1; ESI1; sf_tg; sf_precip.

yield anomalies in many regions (Fig. 6).

5.3. Comparison to similar studies

Our results are in line with previous studies carried out in the same region. For instance, various authors found an increase in performance around two months before harvest (Bussay et al., 2015; Kern et al., 2018; Potopová et al., 2020), while other studies could predict maize yields consistently better than wheat yields (Frieler et al., 2017; Kern et al., 2018; Nagy et al., 2018; Pinke and Lövei, 2017). Generally, these studies obtained better overall performance than our results. Bussay et al. (2015) forecasted maize yields over Hungary on NUTS3 level regions obtaining a correlation between forecasted and observed wheat yields of 0.82 seven weeks before harvest. They used a regression-based approach to forecast crop yields and evaluated it using leave-one-year-out CV. The main difference to this study is that they used a biophysically-based crop growth simulation model and used different predictors. Similar results to Bussay et al. (2015) were obtained by Bognár et al. (2017), who forecasted wheat yields in Hungary on a NUTS3 level with a correlation of 0.83 around 45 days before harvest. They applied a statistical model on NDVI data with a spatial resolution of 250 m and a temporal resolution of 5 days. They validated each year separately by using all antecedent years as training data. However, direct comparisons are hampered by different validation techniques, study region, and observed periods. Another important aspect is that the above-mentioned studies employed absolute crop yields as reference data, which generally leads to higher validation scores compared to using crop yield anomalies, as in this study. Still, what we learn from the comparison to other studies is that our models could be further improved by considering spatial autocorrelations in the data, different temporal and spatial resolutions, considering cropland masks, or by using different detrending methods. In addition, the time series length and the diversity of the crop yield data are crucial for training the model. Hence, more crop yield data than the 10 to 15 years of data used here would most likely improve the results.

5.4. Forecasting spatial crop yield variability

The performance of the model to forecast spatial crop yield variability (Fig. 9D) clearly shows the weakness of the model to distinguish the crop yields between different NUTS3 regions within the years. This is likely related to high spatial autocorrelations of the predictors and the crop yield anomalies (Table 1) and the rather coarse resolution of the predictors. In addition, the spatial autocorrelation leads to a rather low range of yield anomalies (Table 1). All together, these factors make forecasts of the spatial anomaly patterns between regions challenging. On the other hand, at interannual time scales, the range of yield anomalies is larger and the autocorrelation lower. This makes a reliable forecast easier, which is reflected in the higher regional performance. It shows correlations around 0.6 for most regions two months before harvest (Fig. 6). This is in agreement with the findings of Li et al. (2019), who state that the performance of the model to distinguish yields between different regions is strongly related to the spatial yield variability. This is related to a general issue of gradient boosting and other machine learning techniques, which do not consider the timing of the predictors, i.e., many climatic conditions can both reduce and increase crop growth in adjacent regions depending on the antecedent conditions and the local boundary conditions (Lischeid et al., 2022). An additional factor that could contribute to the issue of a low performance to forecast spatial crop yield variabilities are the different spatial resolutions of the predictor datasets which are harmonized to NUTS3 level regions. Parameters with coarse spatial resolutions can have a significant part of their footprint outside the actual region and inside neighbouring regions. This makes an accurate comparison of neighbouring regions even more difficult.

This issue of the model to forecast the spatial variability impacts the overall performance significantly (Table 4). Especially as there are more regions (around 40) than years (15), high performances can only be achieved when the spatial crop yield variabilities are well reflected by the models. Under these circumstances, the overall correlations of 0.79 for maize of LT1 and errors below 0.15 for wheat of LT2 (Table 4) are reasonable results.

5.5. Performance in years of severe drought

Despite the availability of only four years characterized by severe droughts, the availability of crop yield data for more than 40 regions per year results in more than 160 observations per crop and allows for some cautious conclusions. During the four drought periods under investigation the forecast errors are higher than during non-drought years (Table 5), caused by a general underestimation of the magnitudes of the crop yield losses. For example, the mean Pannonian Basin forecast of maize with LT1 overestimates the actual yield on average by around 20% in drought years. For wheat, the forecasted yearly mean Pannonian Basin yield losses in drought years are only around 50% of the observed losses. This could be improved by using more training data, especially during extreme conditions, as machine learning models require a good coverage over all conditions to provide reasonable results for each condition. Nevertheless, the wheat forecast model performs in drought years even better than on average for all years. The errors are below 12% two months before harvest and the correlations even reach 0.63 in the last month before harvest as opposed to only 0.5 for all years together.

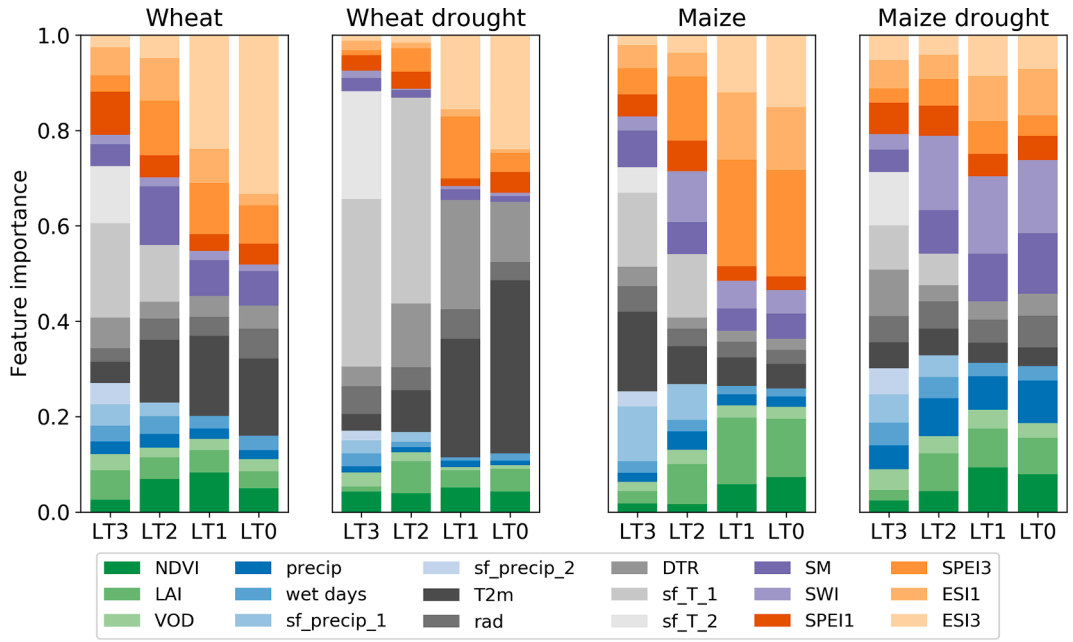


Fig. 10. Temporal evolution of the feature importance of the models. The four plots show the mean feature importance of the predictors of the wheat model of all years, the mean feature importance of the predictors of the wheat model trained in drought years, and then the same for maize. For color coding of the predictors, see caption of Fig. 3 [2-column, color].

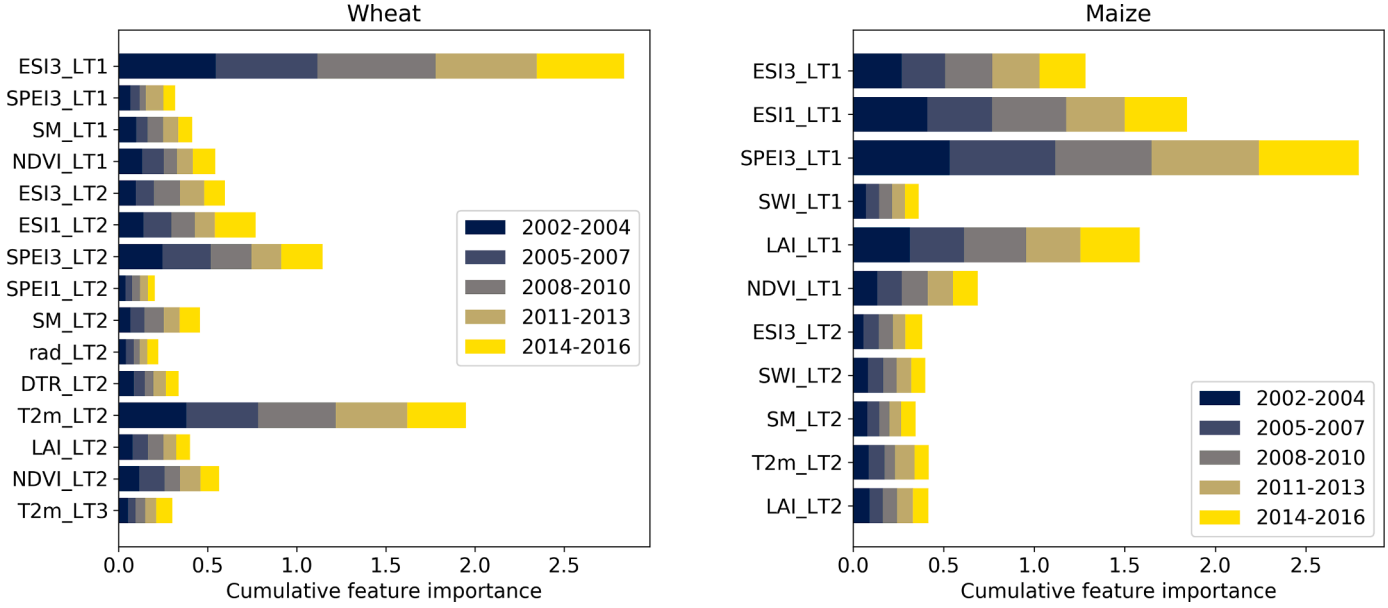


Fig. 11. Development of the feature importance of the models over the different calibrations in the cross-validation. The values of the models in June for wheat and August for maize are displayed. For better visibility, three year-intervals are merged by adding up the feature importance of those years. Only variables with a feature importance larger than 0.2 are displayed [2-column, color].

This is in agreement with the higher correlations of the predictors to wheat yield anomalies when comparing drought years to all years (Fig. 4). Overall, the models are able to early predict crop yield losses in drought years, but underestimate their magnitude. This is in agreement with other studies showing that crop yield losses in drought years are often underestimated (Kang et al., 2020; van der Velde et al., 2018; Pagani et al., 2017; Bussay et al., 2015).

5.6. Impact of the explanatory variables

The yield anomaly forecasting in this study is based on 18 different explanatory variables, including data from EO, ground observations, reanalysis, and seasonal forecasts. The feature importance shows that datasets of all sources impact the models to some degree (Figs. 10 and 11). Therefore, combining different sources for crop yield forecasting is recommendable. The predictors with the largest impacts on the models are EO and reanalysis-based datasets of the drought indices (ESI and SPEI) and air temperature based on in situ data. Other relevant EO datasets are soil moisture, NDVI (mainly for wheat), and LAI (mainly for maize). While seasonal forecasts of temperature have a high feature importance, seasonal precipitation forecasts have only minor impact on the model, which is explained by the low reliability of seasonal

precipitation forecasts southeastern Europe (Crespi et al., 2021; Gospodinov et al., 2020). Measured precipitation up to the forecast date has low predictive power. This suggests that precipitation does not reliably represent the water in the ground, while soil moisture and combined indices like SPEI, ESI are more representative. We found that the wheat model mainly depends on temperature, whereas water availability (in the form of drought indices and soil moisture) is vital for maize. This supports findings from Kern et al. (2018), who stated that in this region, the temperature in May (LT2) is key for winter wheat yield, while water availability in July and August (LT2 and LT1) is crucial for maize yield.

In severe drought years, the impact of the drought indices is lower than on average (Fig. 10), which shows that in these years soil moisture and SWI better reflect the actual water availability. This supports the findings of Sohrabi et al. (2015), who showed that soil moisture data can help to improve the quantification of droughts over established drought indices. For wheat, the impact of the seasonal temperature forecast is more important in drought years indicating its value for drought forecasting as already pointed out in Hao et al. (2018). It is surprising that the impact of ESI3 is lower in drought years compared to all years, as one would expect a drought indicator to provide key information especially in drought years. Moreover, as single predictor ESI3 showed high correlations with wheat yield anomalies in drought years (Fig. 4). This shows the importance of combining multiple datasets from different sources, instead of only considering the most obvious predictors.

6. Conclusion

In this study, an XGBoost-based crop yield anomaly forecast system using EO, climate and weather data is established for various NUTS3 regions in the Pannonian Basin. Pannonian mean crop yield forecasts show a reasonable performance from two months before harvest for all years considered (2002–2016). We observe a strong dependency on water availability for maize and on temperature for wheat in the two months before harvest. However, the model is not able to reliably differentiate yield anomalies between the NUTS3 regions within individual years and underestimates the negative impacts of severe drought. Future work should focus on how these conditions can be better represented in the forecast models, either by considering other input datasets or by increasing the temporal and spatial resolutions of the predictors. A finer spatial resolution could help to better distinguish the yields between the different regions and even bring crop yield forecasts to a field-level. Above all, more crop yield data at various spatial scales (from field to regional) are needed for improved calibration and validation of such machine learning models.

Funding information

This study has been funded by the European Space Agency (ESA), in the framework of the projects „DryPan: Novel EO data for improved agricultural drought impact forecasting in the Pannonian Basin“ (Contract No. 4000127214/19/I-EF) and “YIPEEO: Yield prediction and estimation using Earth Observation” and by Horizon 2020 in the framework of the project Global Gravity-based Groundwater Product (Grant Agreement No. 870353). LZ was financially supported by the EU project "WATERLINE”, project id CHIST ERA 19 CES 006. MF and MT were additionally supported by the Ministry of Education, Youth and Sports of the Czech Republic for SustES – Adaptation strategies for sustainable ecosystem services and food security under adverse environmental conditions project ref. CZ.02.1.01/0.0/0.0/16_019/0000797. The work of AG was in addition partly funded by the Slovenian Research Agency core funding No. P2-0406, and by research project No. J6-9395.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (http://www.uerra.eu) and the data providers in the ECA&D project (https://www.ecad.eu). In addition, we acknowledge FAO.FAOSTAT. License: CC BY-NC-SA 3.0 IGO. Extracted from: https://www.fao.org/faostat/en/#data/QCL. Date of Access: 01-06-2020; and Eurostat (extracted from https://ec.europa.eu/eurostat/web/agriculture/data/database. Date of Access: 01-06-2020) for providing crop yield data which was used for the quality control of our crop yield data. We would like to thank three anonymous revewiers for the comments on the manuscript.

References

Akhavizadegan, F., Ansarifar, J., Wang, L., Huber, I., Archontoulis, S.V., 2021. A time-dependent parameter estimation framework for crop modeling. Sci. Rep. 11 (1), 1–15. https://doi.org/10.1038/s41598-021-90835-x, 2021 111.

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., Martin, E., 2008. From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations. Hydrol.d Earth Syst. Sci. 12 (6), 1323–1337. https://doi.org/10.5194/hess-12-1323-2008.

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration - Guidlines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, 1st ed. FAO - Food and Agriculture Organisation of the United Nations.

Alsafadi, K., Mohammed, S.A., Ayugi, B., Sharaf, M., Harsányi, E., 2020. Spatial–temporal evolution of drought characteristics over hungary between 1961 and 2010. Pure Appl. Geophys. 177 (8), 3961–3978. https://doi.org/10.1007/s00024-020-02449-5.

Anderson, M.C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J.R., Kustas, W.P., 2011. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. J. Clim. 24 (8), 2025–2044. https://doi.org/10.1175/2010JCLI3812.1.

Bales, R.C., Goulden, M.L., Hunsaker, C.T., Conklin, M.H., Hartsough, P.C., O'Geen, A.T., Hopmans, J.W., Safeeq, M., 2018. Mechanisms controlling the impact of multi-year drought on mountain hydrology. Sci. Rep. 8 (1), 1–8. https://doi.org/10.1038/s41598-017-19007-0, 2018 8:1.

Bandhauer, M., Isotta, F., Lakatos, M., Lussana, C., Båserud, L., Izsák, B., Szentes, O., Tveito, O.E., Frei, C., 2021. Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in European regions. Int. J. Climatol. https://doi.org/10.1002/JOC.7269.

Barlow, K.M., Christy, B.P., O'Leary, G.J., Riffkin, P.A., Nuttall, J.G, 2015. Simulating the impact of extreme heat and frost events on wheat crop production: a review. Field Crops Res. 171, 109–119. https://doi.org/10.1016/j.fcr.2014.11.010. Vol.

Bartošová, L., Fischer, M., Balek, J., Bláhová, M., Kudláčková, L., Chuchma, F., Hlavinka, P., Možný, M., Zahradníček, P., Wall, N., Hayes, M., Hain, C., Anderson, M., Wagner, W., Žalud, Z., Trnka, M., 2022. Validity and reliability of drought reporters in estimating soil water content and drought impacts in central Europe. Agric. For. Meteorol. 315 https://doi.org/10.1016/j.agrformet.2022.108808.

Bognár, P., Ferencz, C., Pásztor, S.Z., Molnár, G., Timár, G., Hamar, D., Lichtenberger, J., Székely, B., Steinbach, P., Ferencz, O.E., 2011. Yield forecasting for wheat and corn in Hungary by satellite remote sensing. Int. J. Remote Sens. 32 (17), 4759–4767. https://doi.org/10.1080/01431161.2010.493566.

Bognár, Péter, Kern, A., Pásztor, S., Lichtenberger, J., Koronczay, D., Ferencz, C, 2017. Yield estimation and forecasting for winter wheat in Hungary using time series of MODIS data. Int. J Remote Sens. 38 (11), 3394–3414. https://doi.org/10.1080/01431161.2017.1295482.

Bussay, A., van der Velde, M., Fumagalli, D., Seguini, L., 2015. Improving operational maize yield forecasting in Hungary. Agric. Syst. 141, 94–106. https://doi.org/10.1016/j.agsy.2015.10.001.

Ceglar, A., Croitoru, A.-E., Cuxart, J., Djurdjevic, V., Güttler, I., Ivancan-Picek, B., Jug, D., Lakatos, M., Weidinger, T., 2018. PannEx: the Pannonian Basin experiment. Clim. Serv. 11, 78–85. https://reader.elsevier.com/reader/sd/pii/S2405880718300165?token=4E073B4381118BFD665C383E893B6C0407AFCEE4119A98C3D9A3DBA425456CB16789A600D11DD32457C3324D63D4D80C.

Ceglar, A., Toreti, A., 2021. Seasonal climate forecast can inform the European agricultural sector well in advance of harvesting. NPJ Clim. Atmosp. Sci. 4 (1), 1–8. https://doi.org/10.1038/s41612-021-00198-3, 2021 4:1.
CGLS. (2020). Copernicus global land operations "Vegetation and Energy". In Algorithm theoretical basis document (p. 19). https://land.copernicus.eu/global/products/lai.
Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.
Chen, Y., Feng, Z., Li, F., Zhou, H., Hakala, T., Karjalainen, M., Hyyppä, J., 2020. Lidar-aided analysis of boreal forest backscatter at Ku band. Int. J. Appl. Earth Observ. Geoinform. 91 (April), 102133 https://doi.org/10.1016/j.jag.2020.102133.
Cornes, R.C., van der Schrier, G., van den Besselaar, E.J.M., Jones, P.D., 2018. An ensemble version of the E-OBS temperature and precipitation data sets. J. Geophys. Res.: Atmosp. 123 (17), 9391–9409. https://doi.org/10.1029/2017JD028200.
Craney, T.A., Surles, J.G., 2002. Model-dependent variance inflation factor cutoff values. Qual. Eng. 14 (3), 391–403. https://doi.org/10.1081/QEN-120001878.
Crespi, A., Petitta, M., Marson, P., Viel, C., Grigis, L., 2021. Verification and bias adjustment of ecmwf seas5 seasonal forecasts over Europe for climate service applications. Climate 9 (12), 181. https://doi.org/10.3390/CLI9120181/S1.
Crocetti, L., Forkel, M., Fischer, M., Jurečka, F., Grlj, A., Salentinig, A., Trnka, M., Anderson, M., Ng, W.T., Kokalj, Ž., Bucur, A., Dorigo, W., 2020. Earth Observation for agricultural drought monitoring in the Pannonian Basin (southeastern Europe): current state and future directions. Reg. Environ. Change 20 (123). https://doi.org/10.1007/s10113-020-01710-w.
Donohue, R.J., Lawes, R.A., Mata, G., Gobbett, D., Ouzman, J., 2018. Towards a national, remote-sensing-based model for predicting field-scale crop yield. Field Crops Res. 227, 79–90. https://doi.org/10.1016/j.fcr.2018.08.005.
Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P.D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y.Y., Miralles, D., Lecomte, P., 2017. ESA CCI soil moisture for improved Earth system understanding: state-of-the art and future directions. Remote Sens. Environ. 203, 185–215. https://doi.org/10.1016/J.RSE.2017.07.001.
Dorigo, W., Preimesberger, W., Moesinger, L., Pasik, A., Scanlon, T., Hahn, S., Van der Schalie, R., Van der Vliet, M., De Jeu, R., Kidd, R., Rodriguez-Fernandez, N., Hirschi, M, 2021. ESA Soil Moisture Climate Change Initiative (Soil_Moisture_cci): ancillary Data used for the ACTIVE, PASSIVE and COMBINED Products, Version 06.1. NERC EDS Centre for Environmental Data Analysis, pp. 12–2022. https://catalogue.ceda.ac.uk/uuid/c3bd175b6ed64020b439eb08ed9c8fc2.
Droutsas, I., Challinor, A.J., Deva, C.R., Wang, E., 2022. Integration of machine learning into process-based modelling to improve simulation of complex crop responses. Silico Plants 4 (2), 1–16. https://doi.org/10.1093/INSILICOPLANTS/DIAC017.
Eck, M.A., Murray, A.R., Ward, A.R., Konrad, C.E., 2020. Influence of growing season temperature and precipitation anomalies on crop yield in the southeastern United States. Agric. For. Meteorol. 291, 108053 https://doi.org/10.1016/J.AGRFORMET.2020.108053.
Eitzinger, J., Thaler, S., Schmid, E., Strauss, F., Ferrise, R., Moriondo, M., Bindi, M., Palosuo, T., Rötter, R., Kersebaum, K.C., Olesen, J.E., Patil, R.H., Şaylan, L., Çaldağ, B., Çaylak, O., 2013. Sensitivities of crop models to extreme weather conditions during flowering period demonstrated for maize and winter wheat in Austria. J. Agric. Sci. 151 (6), 813–835. https://doi.org/10.1017/S0021859612000779.
Feng, P., Wang, B., Liu, D.L., Waters, C., Xiao, D., Shi, L., Yu, Q., 2020. Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. Agric. For. Meteorol. 285-286, 107922 https://doi.org/10.1016/J.AGRFORMET.2020.107922.
Ferraciolli, M.A., Bocca, F.F., Rodrigues, L.H.A., 2019. Neglecting spatial autocorrelation causes underestimation of the error of sugarcane yield models. Comput. Electron. Agric. 161, 233–240. https://doi.org/10.1016/j.compag.2018.09.003.
Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S. U., Jephcott, T.G., Paterson, S.E., Whelan, B.M., Bishop, T.F.A., 2019. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precis. Agric. 20 (5), 1015–1029. https://doi.org/10.1007/s11119-018-09628-4.
Frieler, K., Schauberger, B., Arneth, A., Balkovič, J., Chryssanthacopoulos, J., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., Olin, S., Pugh, T.A.M., Schaphoff, S., Schewe, J., Schmid, E., Warszawski, L., Levermann, A., 2017. Understanding the weather signal in national crop-yield variability. Earth's Fut. 5 (6), 605–616. https://doi.org/10.1002/2016EF000525.
García-León, D., Contreras, S., Hunink, J., 2019. Comparison of meteorological and satellite-based drought indices as yield predictors of Spanish cereals. Agric. Water Manag. 213, 388–396. https://doi.org/10.1016/j.agwat.2018.10.030.
Gómez, D., Salvador, P., Sanz, J., Casanova, J.L., 2019. Potato yield prediction using machine learning techniques and Sentinel 2 data. Remote Sens. (Basel) 11 (15), 1745. https://doi.org/10.3390/rs11151745.
Gospodinov, I., Kazandjiev, V., Georgieva, V., 2020. The potential benefit of the use of seasonal forecast during the agricultural economic year 2019-2020 in Bulgaria. In: Jubilee Scientific International Conference, pp. 64–72. https://doi.org/10.22620/agrisci.2021.30.009.
Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., Dorigo, W., 2019. Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. Earth Syst. Sci. Data 11 (2), 717–739. https://doi.org/10.5194/essd-11-717-2019.
Guarin, J.R., Asseng, S., Martre, P., Bliznyuk, N., 2020. Testing a crop model with extreme low yields from historical district records. Field Crops Res. 249, 107269 https://doi.org/10.1016/j.fcr.2018.03.006.
Hain, C.R., Anderson, M.C., 2017. Estimating morning change in land surface temperature from MODIS day/night observations: applications for surface energy balance modeling. Geophys. Res. Lett. 44 (19), 9723–9733. https://doi.org/10.1002/2017GL074952.
Hao, Z., Singh, V.P., Xia, Y., 2018. Seasonal drought prediction: advances, challenges, and future prospects. Rev. Geophys. 56 (1), 108–141. https://doi.org/10.1002/2016RG000549.
Hernandez-Barrera, S., Rodriguez-Puebla, C., Challinor, A.J., 2016. Effects of diurnal temperature range and drought on wheat yield in Spain. Theor. Appl. Climatol. 129 (1), 503–519. https://doi.org/10.1007/S00704-016-1779-9, 2016 129:1.
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Thépaut, J.-N, 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146 (730), 1999–2049. https://doi.org/10.1002/QJ.3803.
Hlaváčová, M., Klem, K., Smutná, P., Škarpa, P., Hlavinka, P., Novotná, K., Rapantová, B., Trnka, M., 2017. Effect of heat stress at anthesis on yield formation in winter wheat. Plant, Soil Environ. 63 (3), 139–144. https://doi.org/10.17221/73/2017-PSE.
Jakubínský, J., Bláhová, M., Bartošová, L., Steinerová, K., Balek, J., Dížková, P., Semerádová, D., Alexandru, D., Bardarska, G., Bokal, S., Borojević, G., Bucur, A., Kalin, K.C., Barbu, A.C., Debre, B., Đorđević, M., Đurić, I., Mircea, B.F., Gatarić, S., Trnka, M., 2019. Repository of drought event impacts across the Danube catchment countries between 1981 and 2016 using publicly available sources. Acta Univ. Agric. Silvicult. Mendel. Brun. 67 (4), 925–938. https://doi.org/10.11118/actaun201967040925.
Johnson, M.D., Hsieh, W.W., Cannon, A.J., Davidson, A., Bédard, F., 2016. Crop yield forecasting on the Canadian Prairies by remotely sensed vegetation indices and machine learning methods. Agric. For. Meteorol. 218 (219), 74–84. https://doi.org/10.1016/J.AGRFORMET.2015.11.003.
Johnson, S.J., Stockdale, T.N., Ferranti, L., Balmaseda, M.A., Molteni, F., Magnusson, L., Tietsche, S., Decremer, D., Weisheimer, A., Balsamo, G., Keeley, S.P.E., Mogensen, K., Zuo, H., Monge-Sanz, B.M., 2019. SEAS5: the new ECMWF seasonal forecast system. Geosci. Model Dev. 12 (3), 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019.
Jones, M.O., Jones, L.A., Kimball, J.S., McDonald, K.C., 2011. Satellite passive microwave remote sensing for monitoring global land surface phenology. Remote Sens. Environ. 115 (4), 1102–1114. https://doi.org/10.1016/J.RSE.2010.12.015.
Jurečka, F., Fischer, M., Hlavinka, P., Balek, J., Semerádová, D., Bláhová, M., Anderson, M.C., Hain, C., Žalud, Z., Trnka, M., 2021. Potential of water balance and remote sensing-based evapotranspiration models to predict yields of spring barley and winter wheat in the Czech Republic. Agric. Water Manag. 256, 107064 https://doi.org/10.1016/J.AGWAT.2021.107064.
Kang, Y., Ozdogan, M., Zhu, X., Ye, Z., Hain, C., Anderson, M., 2020. Comparative assessment of environmental variables and machine learning algorithms for maize yield prediction in the US Midwest. Environ. Res. Lett. 15 (6), 064005 https://doi.org/10.1088/1748-9326/AB7DF9.
Kern, A., Barcza, Z., Marjanović, H., Árendás, T., Fodor, N., Bónis, P., Bognár, P., Lichtenberger, J., 2018. Statistical modelling of crop yield in Central Europe using climate data and remote sensing vegetation indices. Agric. For. Meteorol. 260 (261), 300–320. https://doi.org/10.1016/j.agrformet.2018.06.009.
Kis, A., Pongrácz, R., Bartholy, J., Gocic, M., Milanovic, M., Trajkovic, S., 2020. Multi-scenario and multi-model ensemble of regional climate change projections for the plain areas of the pannonian basin. Idojaras 124 (2), 157–190. https://doi.org/10.28974/idojaras.2020.2.2.
Kogan, F., Kussul, N., Adamenko, T., Skakun, S., Kravchenko, O., Kryvobok, O., Shelestov, A., Kolotii, A., Kussul, O., Lavrenyuk, A., 2013. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorologicaldata and biophysical models. Int. J. Appl. Earth Observ. Geoinform. 23 (1), 192–203. https://doi.org/10.1016/j.jag.2013.01.002.
Kotsiantis, S.B., 2011. Decision trees: a recent overview. Artif. Intell. Rev. 39 (4), 261–283. https://doi.org/10.1007/S10462-011-9272-4, 2011 39:4.
Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15 (3), 259–263. https://doi.org/10.1127/0941-2948/2006/0130.
Leng, G., Hall, J.W., 2020. Predicting spatial and temporal variability in crop yields: an inter-comparison of machine learning, regression and process-based models. Environ. Res. Lett. 15 (4), 044027 https://doi.org/10.1088/1748-9326/AB7B24.
Li, H., Li, Y., Huang, G., Sun, J., 2021. Probabilistic assessment of crop yield loss to drought time-scales in Xinjiang, China. Int. J. Climatol. 41 (8), 4077–4094. https://doi.org/10.1002/JOC.7059.
Li, Y., Guan, K., Yu, A., Peng, B., Zhao, L., Li, B., Peng, J., 2019. Toward building a transparent statistical model for improving crop yield prediction: modeling rainfed corn in the U.S. Field Crops Res. 234, 55–65. https://doi.org/10.1016/J.FCR.2019.02.005.
Li, J., 2017. Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? PLOS ONE 12 (8), 1–16. https://doi.org/10.1371/JOURNAL.PONE.0183250.
Li, Z., Ding, L., Xu, D., 2022. Exploring the potential role of environmental and multi-source satellite data in crop yield prediction across Northeast China. Sci. Total Environ. 815, 152880 https://doi.org/10.1016/J.SCITOTENV.2021.152880.
Lischeid, G., Webber, H., Sommer, M., Nendel, C., Ewert, F., 2022. Machine learning in crop yield modelling: a powerful tool, but no surrogate for science. Agric. For. Meteorol. 312, 108698 https://doi.org/10.1016/J.AGRFORMET.2021.108698.

Lu, J., Carbone, G.J., Gao, P., 2017. Detrending crop yield data for spatial visualization of drought impacts in the United States, 1895–2014. Agric. For. Meteorol. 237 (238), 196–208. https://doi.org/10.1016/J.AGRFORMET.2017.02.001.

Lukić, T., Lukić, A., Basarin, B., Ponjiger, T.M., Blagojević, D., Mesaroš, M., Milanović, M., Gavrilov, M., Pavić, D., Zorn, M., Komac, B., Miljković, D., Sakulski, D., Babić-Kekez, S., Morar, C., Janićević, S., 2019. Rainfall erosivity and extreme precipitation in the Pannonian basin. Open Geosci. 11 (1), 664–681. https://doi.org/10.1515/geo-2019-0053.

Mathieu, J.A., Aires, F., 2018. Using neural network classifier approach for statistically forecasting extreme corn yield losses in eastern United States. Earth Space Sci. 5 (10), 622–639. https://doi.org/10.1029/2017EA000343.

McElrone, A.J., Choat, B., Gambetta, G.A., Brodersen, C.R., 2013. Water uptake and transport in vascular plants. Nat. Educ. Knowl. 4 (5). https://www.nature.com/scitable/knowledge/library/water-uptake-and-transport-in-vascular-plants-103016037/.

McEvoy, D.J., Huntington, J.L., Hobbins, M.T., Wood, A., Morton, C., Anderson, M., Hain, C., 2016. The evaporative demand drought index. Part II: cONUS-wide assessment against common drought indicators. J. Hydrometeorol. 17 (6), 1763–1779. https://doi.org/10.1175/JHM-D-15-0122.1.

McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two interpretations. Agric. For. Meteorol. 87 (4), 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0.

Moesinger, L., Dorigo, W., De Jeu, R., Van Der Schalie, R., Scanlon, T., Teubner, I., Forkel, M., 2020. The global long-term microwave vegetation optical depth climate archive (VODCA). Earth Syst. Sci.Data 12 (1), 177–196. https://doi.org/10.5194/essd-12-177-2020.

Mohammed, S., Alsafadi, K., Enaruvbe, G.O., Bashir, B., Elbeltagi, A., Széles, A., Alsalman, A., Harsanyi, E., 2022. Assessing the impacts of agricultural drought (SPI/SPEI) on maize and wheat yields across Hungary. Sci. Rep. 12 (1), 1–19. https://doi.org/10.1038/s41598-022-12799-w, 2022 12:1.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D., Piles, M., Rodríguez-Fernández, N., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data Discuss. 1–50. https://doi.org/10.5194/ESSD-2021-82.

Nagy, A., Fehér, J., Tamás, J., 2018. Wheat and maize yield forecasting for the Tisza river catchment using MODIS NDVI time series and reported crop statistics. Comput. Electron. Agric. 151, 41–49. https://doi.org/10.1016/J.COMPAG.2018.05.035.

Nistor, M.M., Cheval, S., Gualtieri, A.F., Dumitrescu, A., Boţan, V.E., Berni, A., Hognogi, G., Irimuş, I.A., Porumb-Ghiurco, C.G., 2017. Crop evapotranspiration assessment under climate change in the Pannonian basin during 1991–2050. Meteorol. Appl. 24 (1), 84–91. https://doi.org/10.1002/met.1607.

Norman, J.M., Jarvis, P.G., 1975. Photosynthesis in Sitka Spruce (Picea sitchensis (Bong.) Carr.): V. Radiation penetration theory and a test case. J. Appl. Ecol. 12 (3), 839–878. https://doi.org/10.2307/2402094.

O'Gorman, P.A., Dwyer, J.G., 2018. Using machine learning to parameterize moist convection: potential for modeling of climate, climate change, and extreme events. J. Adv. Model. Earth Syst. 10 (10), 2548–2563. https://doi.org/10.1029/2018MS001351.

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., Micale, F., 2011. Impacts and adaptation of European crop production systems to climate change. Eur. J. Agron. 34 (2), 96–112. https://doi.org/10.1016/j.eja.2010.11.003. VolIssuepp.

Otkin, J.A., Anderson, M.C., Hain, C., Mladenova, I.E., Basara, J.B., Svoboda, M., 2013. Examining rapid onset drought development using the thermal infrared–based evaporative stress index. J. Hydrometeorol. 14 (4), 1057–1074. https://doi.org/10.1175/JHM-D-12-0144.1.

Pagani, V., Guarneri, T., Busetto, L., Ranghetti, L., Boschetti, M., Movedi, E., Campos-Taberner, M., Garcia-Haro, F.J., Katsantonis, D., Stavrakoudis, D., Ricciardelli, E., Romano, F., Holecz, F., Collivignarelli, F., Granell, C., Casteleyn, S., Confalonieri, R., 2019. A high-resolution, integrated system for rice yield forecasting at district level. Agric. Syst. 168, 181–190. https://doi.org/10.1016/j.agsy.2018.05.007.

Pagani, V., Guarneri, T., Fumagalli, D., Movedi, E., Testi, L., Klein, T., Calanca, P., Villalobos, F., Lopez-Bernal, A., Niemeyer, S., Bellocchi, G., Confalonieri, R., 2017. Improving cereal yield forecasts in Europe – the impact of weather extremes. Eur. J. Agron. 89, 97–106. https://doi.org/10.1016/j.eja.2017.06.010.

Papagiannopoulou, C., Miralles, D.G., Verhoest, N.E.C., Dorigo, W.A., Waegeman, W., 2016. A non-linear Granger causality framework to investigate climate-vegetation dynamics. Geosci. Model Dev. Discuss. 1–24. https://doi.org/10.5194/gmd-2016-266. November.

Papagiannopoulou, C., Miralles, D.G., Dorigo, W.A., Verhoest, N.E.C., Depoorter, M., Waegeman, W., 2017. Vegetation anomalies caused by antecedent precipitation in most of the world. Environ. Res. Lett. 12 (7) https://doi.org/10.1088/1748-9326/aa7145.

Peng, B., Guan, K., Pan, M., Li, Y., 2018. Benefits of seasonal climate prediction and satellite data for forecasting U.S. maize yield. Geophys. Res. Lett. 45 (18), 9662–9671. https://doi.org/10.1029/2018GL079291.

Pinke, Z., Lövei, G.L., 2017. Increasing temperature cuts back crop yields in Hungary over the last 90 years. Glob. Chang. Biol. 23 (12), 5426–5435. https://doi.org/10.1111/GCB.13808.

Piramuthu, S., 2008. Input data for decision trees. Expert Syst. Appl. 34 (2), 1220–1226. https://doi.org/10.1016/J.ESWA.2006.12.030.

Portele, T.C., Lorenz, C., Dibrani, B., Laux, P., Bliefernicht, J., Kunstmann, H., 2021. Seasonal forecasts offer economic benefit for hydrological decision making in semi-arid regions. Sci. Rep. 11 (1), 1–15. https://doi.org/10.1038/s41598-021-89564-y, 2021 11:1.

Potopová, V., Trnka, M., Hamouz, P., Soukup, J., Castraveţ, T., 2020. Statistical modelling of drought-related yield losses using soil moisture-vegetation remote sensing and multiscalar indices in the south-eastern Europe. Agric. Water Manag. 236, 106168 https://doi.org/10.1016/j.agwat.2020.106168.

Potopová, V., Trifan, T., Trnka, M., De Michele, C., Semerádová, D., Fischer, M., Clothier, B., 2023. Copulas modelling of maize yield losses–drought compound events using the multiple remote sensing indices over the Danube River Basin. Agric. Water Manag. 280, 108217.

Rebala, G., Ravi, A., Churiwala, S., 2019. Introduction to machine learning. An Introduction to Machine Learning, 1st ed. Springer, Cham. https://doi.org/10.1016/B978-0-12-815739-8.00001-8.

Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci. Rep. 11 (1), 1–15. https://doi.org/10.1038/s41598-020-80820-1, 2021 11:1.

Sohrabi, M.M., Ryu, J.H., Asce, M., Abatzoglou, J., Tracy, J, 2015. Development of soil moisture drought index to characterize droughts. J. Hydrol. Eng. 20 (11), 04015025 https://doi.org/10.1061/(ASCE)HE.1943-5584.0001213.

Song, L., Jin, J., 2020. Improving CERES-Maize for simulating maize growth and yield under water stress conditions. Eur. J. Agron. 117, 126072 https://doi.org/10.1016/J.EJA.2020.126072.

Teubner, I.E., Forkel, M., Jung, M., Liu, Y.Y., Miralles, D.G., Parinussa, R., van der Schalie, R., Vreugdenhil, M., Schwalm, C.R., Tramontana, G., Camps-Valls, G., Dorigo, W.A., 2018. Assessing the relationship between microwave vegetation optical depth and gross primary production. Int. J.f Appl. Earth Observ. Geoinform. 65, 79–91. https://doi.org/10.1016/J.JAG.2017.10.006.

Trnka, M., Hlavinka, P., Možný, M., Semerádová, D., Štěpánek, P., Balek, J., Bartošová, L., Zahradníček, P., Bláhová, M., Skalák, P., Farda, A., Hayes, M., Svoboda, M., Wagner, W., Eitzinger, J., Fischer, M., Žalud, Z., 2020. Czech Drought Monitor System for monitoring and forecasting agricultural drought and drought impacts. Int. J. Climatol. 40 (14), 5941–5958. https://doi.org/10.1002/joc.6557.

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8 (2), 127–150. https://doi.org/10.1016/0034-4257(79)90013-0.

Udmale, P., Ichikawa, Y., Manandhar, S., Ishidaira, H., Kiem, A.S., 2014. Farmers' perception of drought impacts, local adaptation and administrative mitigation measures in Maharashtra State, India. Int. J. Dis. Risk Reduct. 10 (PA), 250–269. https://doi.org/10.1016/J.IJDRR.2014.09.011.

Upton, G., Cook, I., 2014. A Dictionary of Statistics 3e. Oxford university press. https://doi.org/10.1093/ACREF/9780199679188.001.0001.

Velde, M., Baruth, B., Bussay, A., Ceglar, A., Condado, S.G., Karetsos, S., Lecerf, R., Lopez, R., Maiorano, A., Nisini, L., Seguini, L., Berg, M., 2018. In-season performance of European Union wheat forecasts during extreme impacts. Sci. Rep. 8 (1), 1–10. https://doi.org/10.1038/s41598-018-33688-1, 2018 8:1.

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. J. Clim. 23 (7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17 (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

Vreugdenhil, M., Greimeister-Pfeil, I., Preimesberger, W., Camici, S., Dorigo, W., Enenkel, M., van der Schalie, R., Steele-Dunne, S., Wagner, W., 2022. Microwave remote sensing for agricultural drought monitoring: recent developments and challenges. Front. Water 4 (1045451), 1–21. https://doi.org/10.3389/frwa.2022.1045451.

Wan, Z., 2008. New refinements and validation of the MODIS land-surface temperature/emissivity products. Remote Sens. Environ. 112 (1), 59–74. https://doi.org/10.1016/J.RSE.2006.06.026.

WMO, 2010. Agrometeorlogy of some selected crops. Eds.. In: Brunini, O., Stigter, K. (Eds.), Guide to Agricultural Meteorological Practices. WMO, pp. 10–11. to 10–100. http://www.wamis.org/agm/gamp/GAMP_Chap10.pdf.

Zhang, N., Zhou, X., Kang, M., Hu, B.-G., Heuvelink, E., Marcelis, L.F.M., Zhang, N., 2022. Machine learning versus crop growth models: an ally, not a rival. AoB Plants. https://doi.org/10.1093/AOBPLA/PLAC061.

Zhu, P., Abramoff, R., Makowski, D., Ciais, P., 2021. Uncovering the past and future climate drivers of wheat yield shocks in Europe with machine learning. Earth's Fut. 1–13. https://doi.org/10.1029/2020ef001815.

Zhuo, W., Huang, J., Li, L., Zhang, X., Ma, H., Gao, X., Huang, H., Xu, B., Xiao, X., 2019. Assimilating soil moisture retrieved from Sentinel-1 and Sentinel-2 data into WOFOST model to improve winter wheat yield estimation. Remote Sens. (Basel) 11 (13), 1618. https://doi.org/10.3390/rs11131618.