









Article

Estimating Drought-Induced Crop Yield Losses at the Cadastral Area Level in the Czech Republic

Jan Meitner ^{1,2,*} , Jan Balek ¹ , Monika Bláhová ^{1,2} , Daniela Semerádová ^{1,2}, Petr Hlavinka ^{1,2} ,
Vojtěch Lukas ², František Jurečka ^{1,2}, Zdeněk Žalud ^{1,2} , Karel Klem ¹, Martha C. Anderson ³, Wouter Dorigo ⁴ ,
Milan Fischer ^{1,2}  and Miroslav Trnka ^{1,2} 

- ¹ Global Change Research Institute CAS, Bělidla 986/4a, 603 00 Brno, Czech Republic; jan.balek.83@seznam.cz (J.B.); blahova.m@czechglobe.cz (M.B.); dany357@yahoo.com (D.S.); phlavinka@centrum.cz (P.H.); frantisek.jurecka@centrum.cz (F.J.); zalud@mendelu.cz (Z.Ž.); klem.k@czechglobe.cz (K.K.); fischer.m@czechglobe.cz (M.F.); mirek_trnka@yahoo.com (M.T.)
- ² Department of Agrosystems and Bioclimatology, Faculty of Agriscience, Mendel University in Brno, Zemědělská 1/1665, 613 00 Brno, Czech Republic; vojtech.lukas@mendelu.cz
- ³ USDA, ARS, Hydrology and Remote Sensing Laboratory, 10300 Baltimore Avenue, Beltsville, MD 20705, USA; martha.anderson@usda.gov
- ⁴ Department of Geodesy and Geoinformation, Technische Universität Wien (TU Wien), Wiedner Hauptstrasse 8-10, 1040 Vienna, Austria; wouter.dorigo@tuwien.ac.at
- * Correspondence: meitner.j@czechglobe.cz

Abstract: In the Czech Republic, soil moisture content during the growing season has been decreasing over the past six decades, and drought events have become significantly more frequent. In 2003, 2015, 2018 and 2019, drought affected almost the entire country, with droughts in 2000, 2004, 2007, 2012, 2014 and 2017 having smaller extents but still severe intensities in some regions. The current methods of visiting cadastral areas (approximately 13,000) to allocate compensation funds for the crop yield losses caused by drought or aggregating the losses to district areas (approximately 1000 km²) based on proxy data are both inappropriate. The former due to the required time and resources, the later due to low resolution, which leads to many falsely negative and falsely positive results. Therefore, the study presents a new method to combine ground survey, remotely sensed and model data for determining crop yield losses. The study shows that it is possible to estimate them at the cadastral area level in the Czech Republic and attribute those losses to drought. This can be done with remotely sensed vegetation, water stress and soil moisture conditions with modeled soil moisture anomalies coupled with near-real-time feedback from reporters and with crop status surveys. The newly developed approach allowed the achievement of a proportion of falsely positive errors of less than 10% (e.g., oat 2%, 8%; spring barley 4%, 3%; sugar beets 2%, 21%; and winter wheat 2%, 6% in years 2017, resp. 2018) and allowed for cutting the loss assessment time from eight months in 2017 to eight weeks in 2018.

Keywords: crop yield loss; drought; remote sensing; artificial neural network



Citation: Meitner, J.; Balek, J.; Bláhová, M.; Semerádová, D.; Hlavinka, P.; Lukas, V.; Jurečka, F.; Žalud, Z.; Klem, K.; Anderson, M.C.; et al. Estimating Drought-Induced Crop Yield Losses at the Cadastral Area Level in the Czech Republic. *Agronomy* **2023**, *13*, 1669. <https://doi.org/10.3390/agronomy13071669>

Academic Editor: Louis Kouadio

Received: 19 May 2023

Revised: 15 June 2023

Accepted: 17 June 2023

Published: 21 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Drought has been a major factor affecting crop productivity globally, with high-yielding regions such as the USA [1,2] and Europe [3,4] being no exceptions. The general characteristics of the climate change pattern in Europe can be simplified into a wetting trend in northern Europe and increasing aridity in the south. Anomalous years such as 2018 can largely deviate from this general pattern. For instance, Moravec et al. [5] showed a “water seesaw” pattern between the Mediterranean’s wetter-than-usual and warm conditions and positive crop production anomalies, and Central and Northern Europe’s arid and very warm conditions resulted in negative crop production anomalies across a number of countries in 2018. This means that there is a need to pay attention not only to overall

trends but, even more importantly, to extremes when designing robust adaptation measures. This is especially true in Central Europe, where climate change projections [6,7], as well as observed trends [8,9], indicate that droughts will tend to become more frequent and more intensive.

Despite the fact that agriculture in Central Europe represents a seemingly small part of the economy [10], it is nevertheless the dominant land use, with wheat, barley, winter rapeseed, sugar beets and potatoes being the dominant crops [11]. Since 2000, 6 major droughts within Central Europe have been reported, including the spring drought in 2000, the well-known 2003 drought, 2 more regionally constrained events during 2006–2007 and 2011–2012 [12] and the 2015–2019 drought episode, which can be classified as a 500-year drought [13]. This last five-year drought, particularly in 2015, 2017 and 2018, significantly affected production across the region in terms of the yield deviations of three major field crops, i.e., winter wheat, winter rapeseed and spring barley (Figure 1).

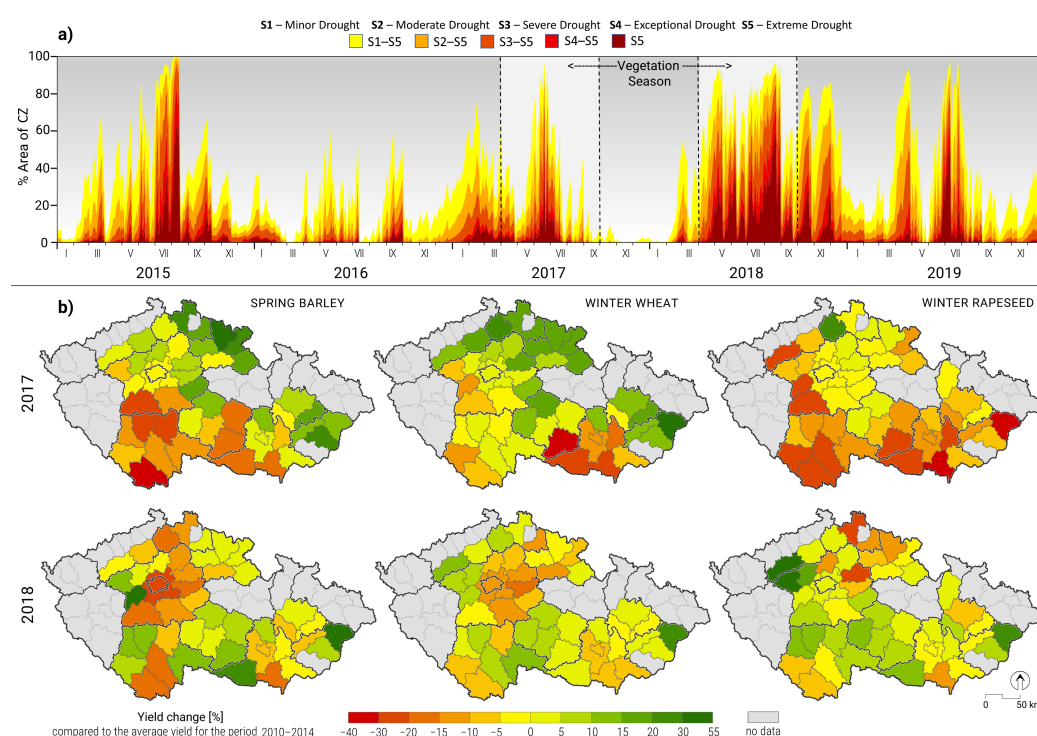


Figure 1. (a) Daily proportion of the Czech Republic (CZ) affected by drought intensity categories [14] between 1 January 2015 and 31 December 2019; (b) yield deviations in NUTS4 regions for winter wheat, spring barley and winter rapeseed during the droughts of 2017 and 2018, expressed as yield anomalies [%] from the 2010–2014 mean.

The occurrence of a major agricultural drought has significant impacts on any agricultural system. The prolonged drought episode of 2015–2019 included three seasons of agricultural droughts that resulted in major crop yield losses. The compounded effect of consecutive droughts in agriculture systems without any drought insurance has led to profound economic problems across the agricultural sector. During 2015, 2017 and 2018, the Ministry of Agriculture of the Czech Republic was confronted with the need to provide farmers with financial subsidies known as “drought compensation” to support the viability of the agricultural sector during years with adverse conditions. However, it has been obvious that the common methods of yield reporting used for the statistical census are too slow and have spatial resolutions that are too low (NUTS2, i.e., approximately 10,000 km² per reporting region) to meet the challenge of providing accurate and spatially explicit estimates and payouts.

In any agricultural drought event, a quick and reasonably accurate yield loss analysis is essential for effectively alleviating socioeconomic drought impacts in the agricultural

sector [15]. The quantification of drought impacts is usually based on (i) simple statistical relationships between yields and ground-based agroclimatic indicators [16], (ii) remote sensing data [17,18] or (iii) crop growth models combined with remotely sensed data [19,20]. While crop growth models may provide complex insights into processes (mostly in a deterministic manner), these models are demanding in terms of their data requirements and are difficult to handle when applied at large scales where many sources of uncertainty are combined [21]. The recent advancements in machine learning provided an opportunity to further analyze the first two approaches and/or their combination.

Quantifying drought and its impact on crops can take advantage of the responsiveness of yield to anomalies in soil moisture, evapotranspiration, spectral crop reflectance, changes in fluorescence, or their combinations. Soil moisture availability to crops or its anomalies with regard to normal conditions can be assessed through ground data-driven systems, such as the Czech Drought Monitor, hereinafter called CzechDM [22,23], or the US Drought Monitor [24], or through the use of remotely sensed soil moisture. Anomalies in evapotranspiration based on land surface temperature remote sensing retrievals are often used as another effective tool for capturing the physiological activities of crops [25–27]. The canopy status can also be assessed on the ground, either during a harvest or by using qualified observers over the course of the season [27]. Conducting assessments over large areas almost implicitly points toward the use of remote sensing approaches, e.g., using microwave technology for soil moisture [28,29]. The spectral properties of certain plant species have been used for crop condition monitoring [30] and yield estimation. Different spectral bands derived from satellite imagery can be used together for the calculation of vegetation indices. Vegetation indices such as the normalized difference vegetation index (NDVI) and the enhanced vegetation index (EVI) are widely used in agriculture due to their ability to track the evolution in crop green biomass and health [20,31]. Crop yield estimation of corn [32,33], barley [34], soybeans [35,36], winter rapeseed [37], sugar beets [38] and wheat [31,39] have frequently taken advantage of the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor onboard the Aqua and Terra satellites and, more recently, of the Sentinel missions operated by the European Space Agency since 2015.

Following the 2015 season, which resulted in major impacts on crop production, a state-of-the-art drought observation and agriculture impact collection system was established in the Czech Republic [22,27]. In 2015, however, the crop loss was only assessed at the NUTS4 level (Figure 1b), resulting in palpable frustration on the side of the farming community, as low spatial resolution led to many false positive and negative errors, especially at the borders of pairs of regions with and without compensation. The newly established approach with higher space resolution could then be tested during the events of 2017 and 2018.

This paper resulted from two successive drought seasons and aims to evaluate several approaches for yield prediction based on remote sensing and machine learning and to provide recommendations on how to effectively provide an independent objective source of data concerning the impacts of droughts on crop yield losses. The key hypothesis is that a machine learning approach based on artificial neural network (ANN) models combining remotely sensed and ground-based data can be used to guide the determination of crop yield losses up to the farm level in near-real time. The presented paper (i) describes the data collected and methods used; (ii) introduces the nature of the 2017 and 2018 drought episodes, including their impacts; and follows with (iii) a description of the applied approaches and the achieved results.

2. Materials and Methods

To estimate the impacts of the 2017 and 2018 droughts on yields, principally, the same workflow was followed for both years (Figure 2).

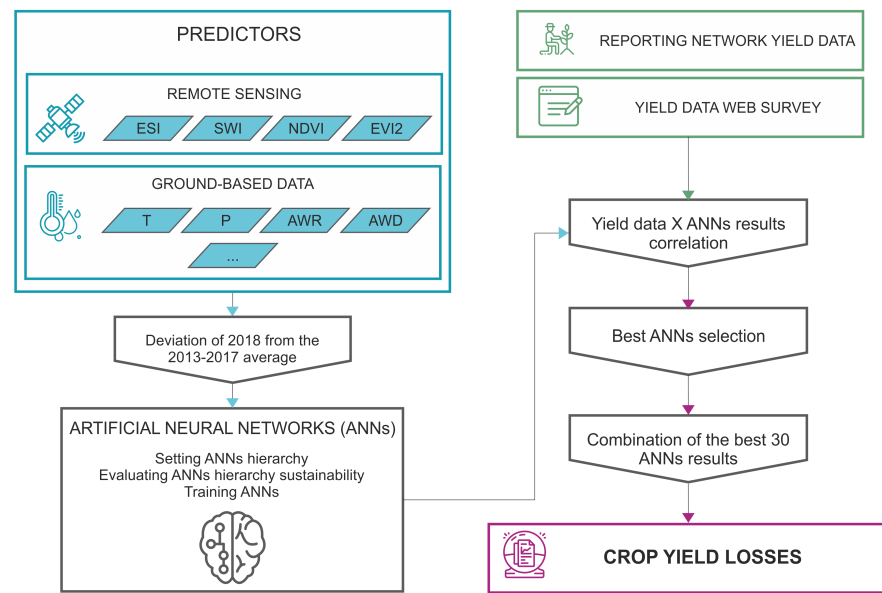


Figure 2. The workflow of the drought impact analysis of yield set for 2018. Artificial neural networks combine predictors and yield data from samples to estimate crop yield losses for the whole cadastral area level.

2.1. Data on Drought-Based Yield Losses

The basic administrative unit of the study was a so-called cadastral area. In total, there are 13,091 cadastral areas in the Czech Republic, with a total area of 78,000 km². The average size of a cadastral area is thus approximately 6 km² or 600 ha, and over 53% of the whole area of the Czech Republic is classified as arable land (Figure 3a). Information about arable land is open to the public in the Land Parcel Identification System (LPIS) [40].

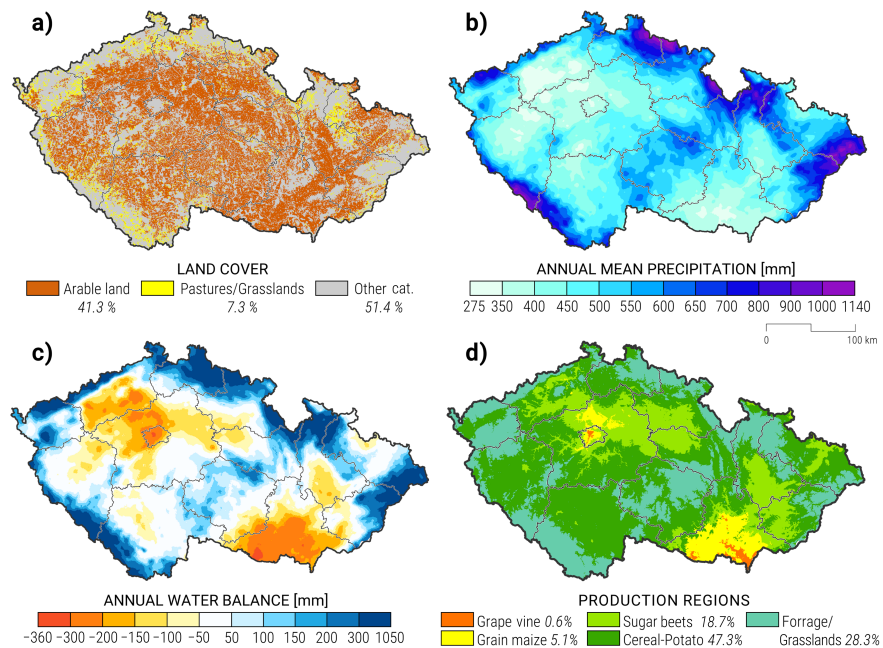


Figure 3. Images of the areas of interest depicting (a) the extent of the arable land and grasslands; (b) the annual mean precipitation; (c) the annual climatic water balance, i.e., the difference between the precipitation and the reference evapotranspiration; and (d) the main cultivars for the period from 1991–2020.

Drought impacts on yields were assessed for 17 different crops in the years of 2017 and 2018 (Table 1 and Table 2, respectively). The drought impact data were not available for all crops in both years, and the included field crops together with grasslands for hay represented over 90% of the agricultural area of the Czech Republic. The two levels of yield decline were considered, i.e., 30 and 50% compared to mean yield in previous five years not counting the highest and lowest yield. When the threshold for the yield decline is being reached at the farm/region, then the state is allowed to financially compensate affected farms based on the national drought response plan notified by the European Commission.

Table 1. Overview of the 2017 drought impact data on yield obtained from the special questionnaire and Czech Drought Monitor (CzechDM) surveys.

Crop	Cadastral Areas Reported (Special Questionnaire)	Usable Records from the Special Questionnaire [%]	Cadastral Areas Reported (CzechDM)	Cadastral Areas with Yield Losses from 30 to 50%	Cadastral Areas with Yield Losses over 50%
Grain maize	443	30.25	105	1104	88
Spring barley	843	30.96	116	947	17
Winter barley	636	24.06	116	423	4
Winter rapeseed	1165	37.25	99	1421	40
Winter wheat	1329	50.04	116	736	41

Table 2. Overview of the 2018 drought impact data obtained with respect to yield from the special questionnaire and Czech Drought Monitor (CzechDM) surveys.

Crop	Cadastral Areas Reported (Special Questionnaire)	Usable Records from the Special Questionnaire [%]	Cadastral Areas Reported (CzechDM)	Cadastral Areas with Yield Losses from 30 to 50%	Cadastral Areas with Yield Losses Over 50%
Alfalfa	1015	57.93	130	2280	586
Clover	1401	49.89	133	3441	992
Grain maize	1243	32.66	297	4332	859
Grasslands	5022	74.07	330	6907	4142
Hops	113	59.29	10	131	42
Oat	1083	28.07	103	909	36
Poppy seeds	921	25.84	42	105	1030
Potatoes	584	44.18	160	1078	147
Silage maize	3139	49.57	300	2017	145
Spring barley	3694	37.28	336	1054	75
Spring wheat	1064	19.92	-	1063	42
Sugar beets	1315	37.11	69	1027	211
Sunflower	358	26.82	49	151	21
Triticale	929	25.51	-	-	-
Winter barley	2267	28.01	361	392	24
Winter rye	542	24.35	357	275	7
Winter wheat	6729	51.67	373	2051	186

In both years, two unique datasets were collected through an intensive data collection. In 2017, drought impact data were collected using weekly questionnaire reports provided by the national reporting network of CzechDM [27]. This database includes expected yield loss information from more than 300 cadastral areas, including only those with at least 3 reports for a given crop during the drought event. To supplement this long-term weekly monitoring data, the Agrarian Chamber of the Czech Republic and the Agricultural Association of the Czech Republic carried out a special drought impact survey. Both institutions collected yield data during 2012–2017 at the cadastral area level through their members using a unified survey form. The final database of reports contained information about winter wheat (1329 reports), spring barley (842 reports), winter barley (636 reports), corn (443 reports) and winter rapeseed (1165 reports). On average, 34.5% of the reports from the special data collection process were used for our modeling purposes. The remaining

reports failed to provide sufficiently long coverage of yield data for the given cadastral areas (~90%), or the data contained errors of various types (~10%).

Based on the 2017 experience, a more thorough system was developed in 2018. As in 2017, the network of CzechDM reporters (3050 reports in total) was used together with the specialized questionnaire in cooperation with the Agricultural Association of the Czech Republic. Thanks to the effort of the major farmers' associations and the Ministry of Agriculture, the data collection process in 2018 had a much wider outreach and obtained information at the cadastral area level for the 2013–2018 period from 7191 cadastral areas. Of all the reports, 6202 cadastral areas were suitable for modeling based on data availability and quality control. The remaining reports failed to include sufficiently long yield data or reported incorrect yield values. Overall, 46.11% of the total number of 31,419 reports were fit for the development of the model. The report with the highest yield loss was selected if there were multiple reports for a given crop and cadastral area.

The cadastral area level yield data (in tons/ha) were screened for (i) missing yield reports in the given year and the five preceding years; (ii) repeating values for multiple cadastral areas, as some farms reported the same values across several cadastral areas they manage; (iii) zeros instead of missing values; (iv) incorrect units; and (v) typing errors. In case the report could not be corrected in collaboration with the reporting farm, the reports were not considered. Only reports with a coverage of at least three out of five years were considered.

The yield deviation (for each cadastral area and each crop) in a given year y was calculated as

$$Dev(Yield, y) = \frac{Yield_y}{AVG_{i=y-5}^{i=y-1} Yield_i} - 1, \quad (1)$$

where $Yield_i$ is yield in year i , we omitted members during years in which the yields were not known and AVG denotes the average. This means that $Dev(Yield, y) > 0$ shows an increase in yield with respect to the previous period and $Dev(Yield, y) < 0$ shows a decrease in yield.

The CzechDM-based data have been active since 2014 and, thanks to collaborating farmers, provides a continuous assessment of drought impact for selected crops. Unlike the specialized surveys in 2017 and 2018, it contains only relative drought impacts on yield, however, these are recorded over the entire season. The collaborating farmers have been trained to provide every week of the growing season their estimation of the crop yield reduction compared to the last three years. The survey includes the following categories: no decrease, decreases of 0–10%, decreases of 10–30%, decreases of 30–40% and decreases over 40%. These relative assessments of yield loss were found quite valuable and have been included in the study. At least three reports from the onset of the drought episode had to be provided for the reporting farm to be used in the study. Crop yield losses reported at the time of the harvest were considered.

Each crop had a different number of cadastral areas with valid data from which the $Dev(Yield, y)$ was calculated. We called each such cadastral area a sample.

2.2. Yield Loss Predictors

On the other hand, we used predictors that were time averaged to a weekly resolution and then spatially averaged to the arable land at the cadastral area level. Let $Pr_i(w)$ define the value of the predictor in year i and week w . Then, the deviation of the predictor for year y (for each cadastral area) for the period from week w_1 to week w_2 was calculated as

$$Dev(Pr, y, w_1, w_2) = \frac{\sum_{w=w_1}^{w_2} Pr_y(w)}{AVG_{i=y-5}^{i=y-1} (\sum_{w=w_1}^{w_2} Pr_i(w))} - 1, \quad (2)$$

where we omitted members during years for which yields were not known due to them being matched with $Dev(Yield, y)$.

The set of predictors is listed in Table 3. The vegetation indices relied on ten-day filters and used the maximum value for each grid cell to eliminate the view obstruction caused by cloud cover. In total, eight remote sensing data predictors were used. The high density of ground-based climate and precipitation stations was used to run the national soil moisture model [14]. In total, 25 agrometeorological characteristics and characteristics based on water balance model simulations were applied in the study.

Table 3. The list of indicators including their time steps and resolutions. The * indicators are based on the remotely sensed data. Soil moisture model indicators apply the 1961–2010 reference period, while for the remotely sensed data, the reference period is 2000–2016. A provider of all indicators was intersucho.cz [23], other alternative providers are in the table.

Acronym of the Indicator	Description	Time Step	Spatial Resolution	Data Provider
AWD	Soil water content anomaly from the reference period in mm for 0–100 cm soil depth	Daily	500 m	
AWD1	Soil water content anomaly from the reference period in mm for 0–40 cm soil depth	Daily	500 m	
AWP	Drought intensity anomaly from the reference period for 0–100 cm soil depth	Daily	500 m	
AWP1	Drought intensity anomaly from the reference period for 0–40 cm soil depth	Daily	500 m	
AWR	Relative soil moisture content as a share of the field capacity in % for 0–100 cm soil depth	Daily	500 m	
AWR1	Relative soil moisture content as a share of the field capacity in % for 0–40 cm soil depth	Daily	500 m	
AWV	Soil moisture content in mm for 0–100 cm soil depth	Daily	500 m	
AWV1	Soil moisture content in mm for 0–40 cm soil depth	Daily	500 m	
DaysAwp_S2+	Number of days with AWP values of 2 or higher per season		500 m	
DaysAwp_S3+	Number of days with AWP values of 3 or higher per season		500 m	
DaysAwp_S4+	Number of days with AWP values of 4 or higher per season		500 m	
DaysAwp1_S2+	Number of days with AWP1 values of 2 or higher per season		500 m	
DaysAwp1_S3+	Number of days with AWP1 values of 3 or higher per season		500 m	
DaysAwp1_S4+	Number of days with AWP1 values of 4 or higher per season		500 m	
DaysAwr_30	Number of days with AWR values of 30 or lower per season		500 m	
DaysAwr_50	Number of days with AWR values of 50 or lower per season		500 m	
DaysAwr1_30	Number of days with AWR1 values of 30 or lower per season		500 m	
DaysAwr1_50	Number of days with AWR1 values of 50 or lower per season		500 m	
DaysHeatDrought	Number of days with AWR < 30% and concurrent heatwaves (periods with average maximal temperature $\geq 30^{\circ}\text{C}$ and daily maximal temperature $\geq 30^{\circ}\text{C}$ for 3+ days in row) per season		500 m	
DaysTmax35	Number of days with maximal temperatures $> 35^{\circ}\text{C}$ per season		500 m	
ET _o	Reference evapotranspiration	Daily	500 m	
ET _a /ET _o	Actual-to-reference evapotranspiration ratio	Daily	500 m	
P	Daily precipitation in mm	Daily	500 m	
P-ET _o	Sum of differences between the sum of daily precipitation and the sum of daily reference evapotranspiration for April–June period	Daily	500 m	
T	Daily average temperature in $^{\circ}\text{C}$	Daily	500 m	
* ESI _{12WK}	12-week accumulated evaporative stress index based on the ALEXI approach	Weekly	3.5 km	USDA/NASA
* ESI _{4WK}	4-week accumulated evaporative stress index based on the ALEXI approach	Weekly	3.5 km	USDA/NASA
* EVI2	MODIS-derived 2-band enhanced vegetation index calculated from surface reflectance bands	Daily	5 km	NASA
* EVI2 _{PAAG}	EVI2 anomaly	Weekly	5 km	

Table 3. Cont.

Acronym of the Indicator	Description	Time Step	Spatial Resolution	Data Provider
* NDVI	MODIS-derived normalized difference vegetation index calculated from surface reflectance bands	Daily	5 km	NASA
* NDVI _{PAAG}	NDVI anomaly	Weekly	5 km	
* SWI _{0–40}	Soil moisture content in % for 0–40 cm soil depth	Weekly	11 km	Copernicus
* SWI _{0–100}	Soil moisture content in % for 0–100 cm soil depth	Weekly	11 km	Copernicus

2.3. Aggregating the Results of Reports on the Cadastral Area

The crop type identification process was performed in cooperation with the State Agricultural Intervention Fund on the tabular data for individual years containing crop species, unique land parcel IDs, grower identification information and crop areas. For the spatial delineation of land parcels, the geospatial data of the LPIS were used as the main administrative units for agricultural land. The LPIS database, provided by the Czech Ministry of Agriculture, is a publicly available data source with shapefile data and a set of attribute information, including unique land parcel IDs [40]. As the next step, the tabular crop type data were paired with the LPIS spatial dataset by the land parcel ID attribute in the Environmental Systems Research Institute ArcGIS environment (Redlands, CA, USA), and a final crop type geodatabase was produced for each year. In cases with land parcels growing more than one crop species (approximately 20% of arable land), the prevailing crop area was used to identify the main crop. As the final step, the land parcel data were aggregated to calculate the total area of the crop types for each cadastral area.

2.4. Development of the Crop Yield Loss Model

To estimate crop yield losses, we used an artificial neural network (ANN). An ANN is a dynamic structure that is used to find masked relations between input and output. For each crop, we set the ANN individually (Table 4). We used a three-layer ANN architecture. The first layer was the input layer composed of the predictors, the second layer was a hidden layer and the third layer was the response variable. In the determination phase of the ANN, we wanted to find the optimal settings of the weights and thresholds of individual neurons to satisfy the ANN-required operations. We set the number of neurons in the first two layers by the number of samples, by the resilient propagation rule and by expert practice. Then, we randomly split the samples into two parts: two-thirds for training the ANN, one-third for validation. We selected suitable predictors (the neurons in the first layer) according to the maximum of the average of the Spearman and Pearson correlation coefficients between yield deviations and predictor deviations. Then, we trained the ANN, starting with random weights. We measured the mean squared error. When the validation error lines and those of the training part were similar, then the ANN was suitable. We had to reduce the number of neurons when they differed. First, we tried to reduce the number of neurons in the hidden layer. We also had to reduce the number of predictors when the previous reduction was insufficient. In this way, the most suitable architecture for the ANN was developed.

The next work phase focused on designing the ANN architecture. Because each ANN performs better in different data subsets, an ensemble could reduce the impact of the poorest performing ANNs (as well as the better ANNs). To obtain reliable outputs, we used an ensemble of 50 different ANNs for a given architecture. They differed in terms of dividing the samples into two parts and randomly initializing the weights of the ANNs. From 50 ANNs, we selected 30 as the best ANNs according to a coefficient of determination validation of the modeled and real sample values. The output of the model was considered the average of the whole ensemble of the best 30 ANN outputs.

Table 4. Settings of the artificial neural networks (ANNs) for the selected crops in 2018, including the number of samples given by the number of different cadastral areas, the hierarchy of the three-level ANN and the predictors. A predictor value equal to 1/0 indicates that predictor was/was not used in the ANN. The complete list is given in Appendix A Tables A1 and A2.

Crop	Winter Wheat	Spring Barley	Grain Maize	Sugar Beets	Potatoes	Poppy Seeds
Samples for the ANN	3839	1711	700	556	417	277
ANN hierarchy	20-20-1	20-15-1	15-7-1	15-6-1	10-7-1	9-5-1
Alt	1	1	1	0	0	
AWD	0	0	0	0	0	1
AWD1	1	1	1	0	1	0
AWP1	1	1	1	0	0	0
AWR1	1	1	1	1	1	0
DaysAWP_S2+	1	1	1	0	0	0
DaysAWP1_S3+	1	1	1	0	0	0
DaysAWR_50	1	1	0	1	1	1
DaysAWR1_30	1	1	0	1	1	0
DaysHeatDrought	1	1	1	1	1	1
DaysTmax35	0	1	0	1	0	0
ESI _{12WK}	1	1	1	1	0	0
ESI _{4WK}	1	1	1	1	1	1
ET _o	0	1	1	1	1	1
ET _a /ET _o	1	1	0	0	0	0
EVI2	1	1	1	0	0	0
EVI2 _{PAAG}	1	1	1	1	0	0
Lat	1	1	0	1	0	0
NDVI	1	1	0	0	0	1
NDVI _{PAAG}	1	1	0	0	0	1
P-ET _o	1	1	1	1	1	1
SWI ₀₋₄₀	1	0	1	1	1	0
SWI ₀₋₁₀₀	0	0	1	1	0	0
T	1	1	1	1	1	1

To evaluate models, it was appropriate to define the following two errors (the overall error and partial error). Let N_{wrong} be the number of all incorrectly estimated cadastral areas, N be the total number of cadastral areas and $N_{falsecomp}$ be the number of cadastral areas that are estimated to receive compensation and actually not eligible to receive compensation; then,

$$Error_{overall} [\%] = 100 \frac{N_{wrong}}{N}, \quad (3)$$

$$Error_{partial} [\%] = 100 \frac{N_{falsecomp}}{N_{wrong}}. \quad (4)$$

3. Results

3.1. Droughts of 2017 and 2018

The drought of 2017 started toward the end of 2016, had two peaks prior to the main growing season (in February and early April of 2017) and then resumed toward the end of May through early August (Figure 1a). The intensity of this drought and impact on vegetation conditions were most pronounced in the southeast and southwest of the country (Figure 4). These two distinct regions were also affected by significant yield reductions (Figure 4). This created almost a “dipole” situation, with the south of the country facing significant yield losses, while the northern regions were much less affected. In contrast, Figures 1 and 4 document that the drought in 2018 was more ubiquitous, starting in February 2018 and continuing until April 2019 (Figures 1 and 4). The effect of this drought on vegetation conditions was profound across the entire country (Figures 1 and 4), with the northern half being affected earlier and more intensively than the southern half; however, the drought of 2017 persisted when the drought of 2018 started. Yield losses were widespread in both years (Figure 1). In particular, 2018 exhibited extreme hydroclimatic

conditions across various locations of the European continent [41], yet the impacts were characterized by considerable spatial [41] and temporal [42] heterogeneity.

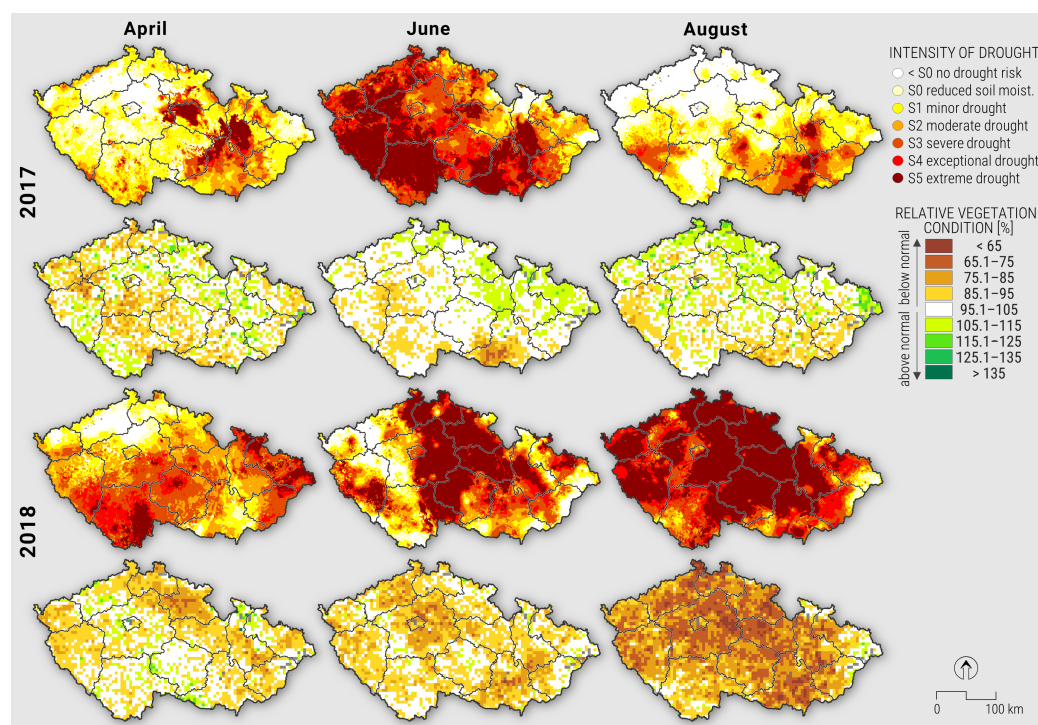


Figure 4. The intensities of the droughts and relative vegetation conditions (based on EVI2) provided for April, June and August of 2017 and 2018.

3.2. Predictors of Crop Loss

Based on correlations, remote sensing data predictors of yield, especially EVI2, performed best. Additionally, predictors based on temperature and data from the SoilClim model, especially AWR1, showed considerable skill in predicting yield losses. The most frequently used predictors in 2017 were AWR1, EVI2 and $EVI2_{PAAG}$, with NDVI, $NDVI_{PAAG}$, AWR and AWR also being used quite often. In 2018, the most frequently used predictors were ESI_{4WK} , T and AWR1, with Awd1, DaysAwr_50, DaysAwr1_30, ETr, EVI2, $EVI2_{PAAG}$, Lat, $P-ET_o$ and SWI_{0-40} also frequently considered.

Since there were significantly fewer samples in 2017 than in 2018, it was necessary to use fewer predictors in the first layer of the ANN. The highest correlations among the individual predictors with yields in 2017 were approximately 0.5 for grain corn and sugar beets in cases involving multiple predictors. For example, for spring barley and winter rapeseed, the maximum correlations were only slightly higher than 0.2. In 2018, the maximum correlations among the crops that were monitored in both years were on average lower than those in 2017. The highest was 0.5 for sunflowers, followed by approximately 0.4 for rye, winter wheat, oats, corn with grain and sugar beets. Other crops had correlations of approximately 0.3, and hops even had correlations below 0.2.

Virtually all crop yield loss models used both ground-based and remotely sensed predictors, except for the potato model, for which satellite data did not improve the model skill.

3.3. Estimating Yield Losses

The results indicate that it was, in general, easier to determine cadastral areas that were entitled to obtain compensation in 2017 than in 2018. This can be clearly seen in the Taylor diagram (Figure 5), where symbols for all crops for 2017 indicate higher correlations and lower errors than their 2018 counterparts. The main reason for this result is the high spatial variability of the drought in 2018 across the whole Czech Republic in June and

August. Similarly, the vegetation condition was worse over the whole Czech Republic in 2018 (Figure 4). In the crucial months of the growing season in 2017, there were especially persistent effects in the southwestern part of the Czech Republic. This result agrees with the poor vegetation status across the region.

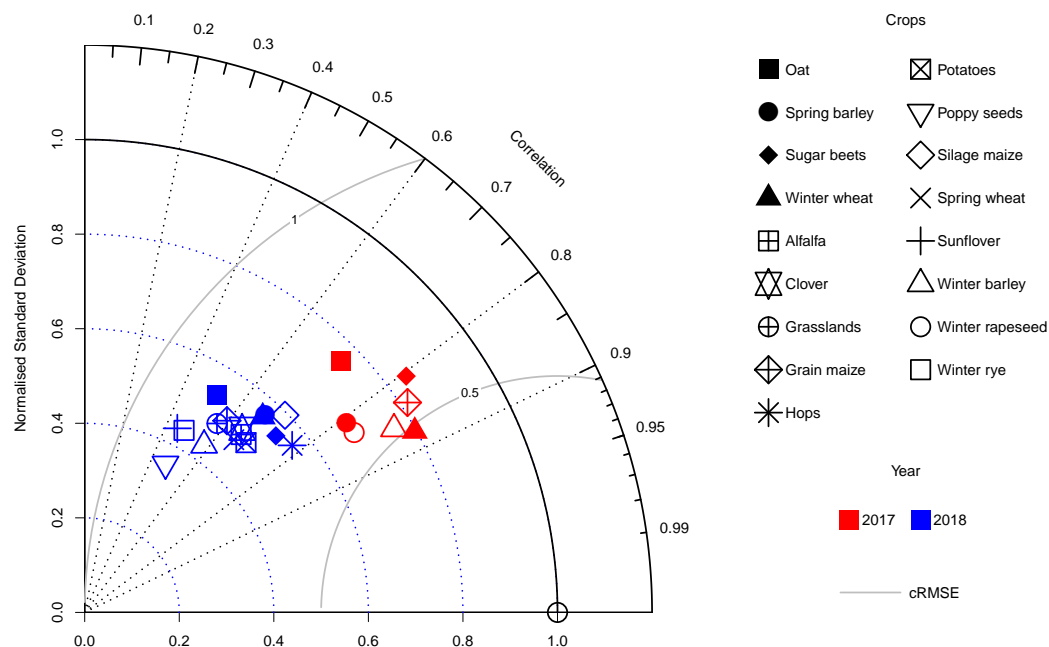


Figure 5. The spatial Taylor diagram of predicted versus observed crop yield losses in 2017 and 2018. The model results for 2017 were more successful, as evidenced by their higher correlations and lower root mean square errors. Solid symbols indicate major crops that were common to both years.

The sample data produced “lighter” tails than the modeled estimates (Figure 6). The only crop yield loss model with a high share of false positive results was the sugar beet model in 2018. The peak of the density in 2018 was left of the -0.3 threshold, meaning that many estimates fell into the compensation category (30% yield loss). At the same time, most of the observed and modeled yield anomalies were below zero (Figure 6); therefore, in most cadastral areas, negative yield anomalies were recorded.

The models underestimated the number of cadastral areas with low and high yield anomalies (Figure 6). This, in general, corresponded to the intended use of the model, as it meant that the models underestimated the scale of the yield decreases by 30 or even 50%, and therefore false positive results did not occur as frequently.

The main aim of the yield loss model was to select as many cadastral areas as possible so we could be almost sure that the yield decrease was at least 30% or even 50% to satisfy the conditions for obtaining compensation. Thus, false positive results, i.e., mistakenly categorizing the cadastral area into the category with a high yield loss, are undesirable. On the other hand, false negative results are less of an issue, as an alternative (but more administratively demanding) method is available for claiming compensation.

Given this condition, the overall error (Equation (3)) and partial error (Equation (4)), which are ratios of boxes in Figure 7, were defined. To compare the errors among different crops and seasons, it is not possible to track only one of these errors. At the very least, we must look at both of these errors simultaneously because each of them alone does not give a sufficient picture of the situation. Moreover, the percentage of cadastral areas with genuine claims for compensation is also relevant information.

If we look at errors (Figure 7a), the model in 2017 for sugar beets came out best, with an overall error of 7% and a partial error of 23%, which corresponds to higher correlations of the input predictors to crop yield losses and low percentage of cadastral areas eligible for compensation. Winter wheat had an overall error of 12% and a partial error of 16%.

Furthermore, the results obtained for winter barley and oats were very similar (overall errors of 14% and 16%, partial errors of 16% and 12%). Then, winter rapeseed (an overall error of 15%, a partial error of 23%) and spring barley performed even worse, with an overall error of 22% and a partial error of 17%; these two crops also had the lowest correlations with the input predictors. The poorest performing model in 2017 was for grain maize (an overall error of 13%, a partial error of 50%).

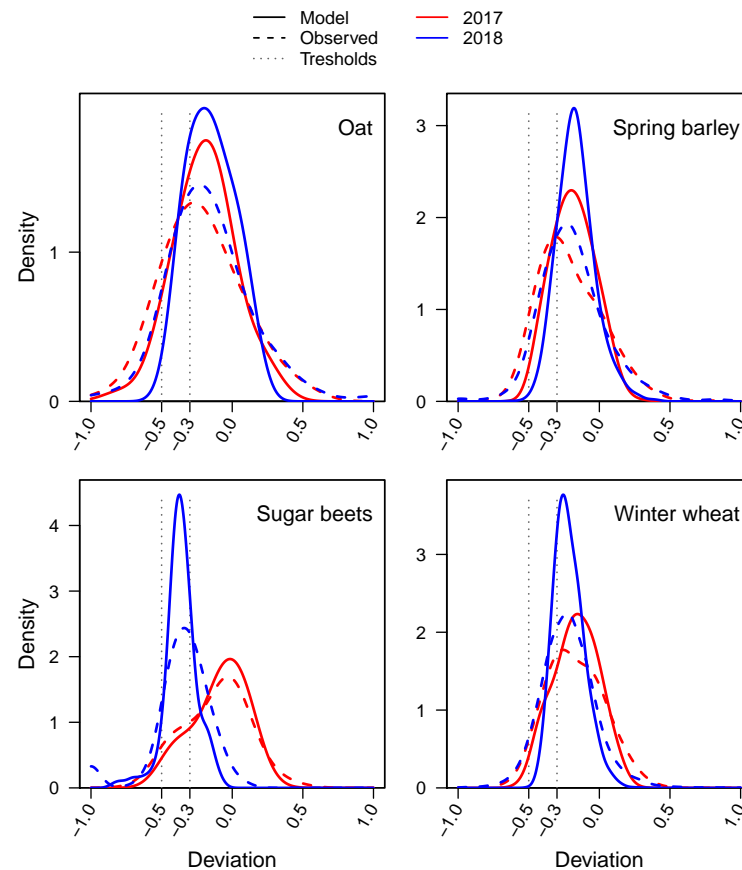


Figure 6. Kernel density estimation series with the bin size equal to 0.1 times the deviations of the estimated and observed yields of oats, spring barley, sugar beets and winter wheat in 2017 and 2018. The vertical dashed lines depict thresholds of two degrees of compensation (30 and 50% yield loss).

For 2018, more models were developed but with different accuracies compared to 2017, with the model performance primarily being affected by the extent of the data used for model development. For example, poppy seeds produced the smallest overall error among all processed crops with 17%, but its partial error was 100%. This is because the correlations of predictors to crop yield losses were quite low, the percentage of cadastral areas eligible for compensation were very high and we had the second-smallest number of samples for poppy seeds. Similarly, for hops, which had the fewest samples and the lowest correlations, the overall error was only 18%, but the partial error was 86%. On the other hand, we considered the best three models to be those of winter barley, winter rye and spring barley, which all had an overall error of approximately 25% and a partial error of less than 12%. Acceptable results could be found for winter wheat, silage maize, spring wheat, oats and sunflowers.

An overview of the errors for each crop is presented in Figure 7b. Compared to Figure 7a, we see that performing weighting at the cadastral area level did not provide any significant improvement of the model predictions. However, in both years, the area weighting actually reduced the part corresponding to the compensation claim for all crops but winter wheat.

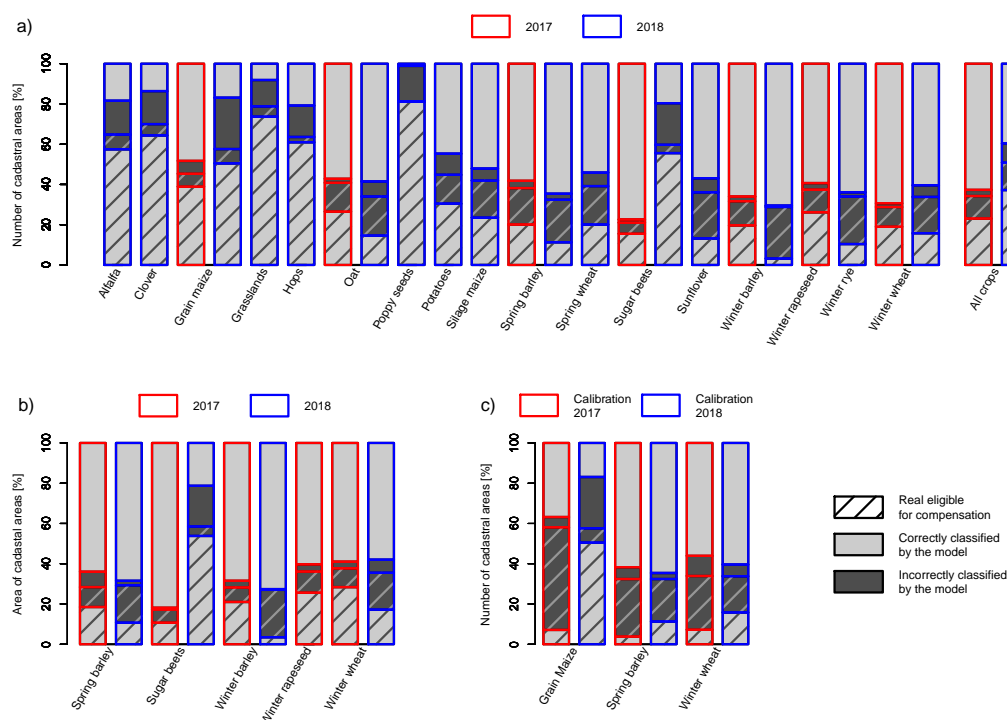


Figure 7. Scaled bar plots for both 2017 and 2018 showing the distributions of the successes of the model for all crops and the total distribution over all crops together (a) as a percentage of the number of cadastral areas; (b) as the percentage of the area of cadastral areas and (c) as a comparison between the distribution of the success of the model for 2018 with two different calibration datasets from 2017 and 2018. The results show that only the calibration dataset from the current year should be used. Overall error is the dark-grey box. False positive results are in the hatched area in the dark-grey box, and the partial error is the ratio of hatched dark-grey to full dark grey area.

Figure 7c compares the errors obtained using the 2017 calibration for the 2018 estimates with the correct 2018 calibration for the 2018 estimates. The models with the 2017 calibration failed, and the overall error increased to 78–88% for the three selected crops: grain maize, spring barley and winter wheat. Thus, we did not correctly detect almost any of the cadastral areas entitled to compensation. For spring barley and winter wheat, the models were total failures, as they induced large partial errors when using the 2017 calibration. This implies that modeling cannot replace yield loss data collection at the cadastral area level, but can only supplement ongoing survey efforts—effectively serving to improve spatial detail in the loss assessment. With this approach, loss models for each crop will need to be recalibrated for each drought-affected season, as a generally calibrated crop loss model would be highly inaccurate.

For a spatial overview of the estimated situation compared to special questionnaire web survey outputs in the cadaster area level for both years 2017 and 2018, maps for individual crops were produced, and a spatial comparison between years 2017 and 2018 of reported crop yield losses over 30% where the highest loss over all crops for a given cadastral area was considered (Figure 8). Drought in 2017 had an impact on this general crop yield loss over 30% by copying of the annual water balance pattern (Figure 3c). Drought in 2018 affected mostly the complement of the growing area compared to in 2017 (Figure 4). Therefore, the crop yield losses over 30% were over the whole growing area of the Czech Republic in 2018.

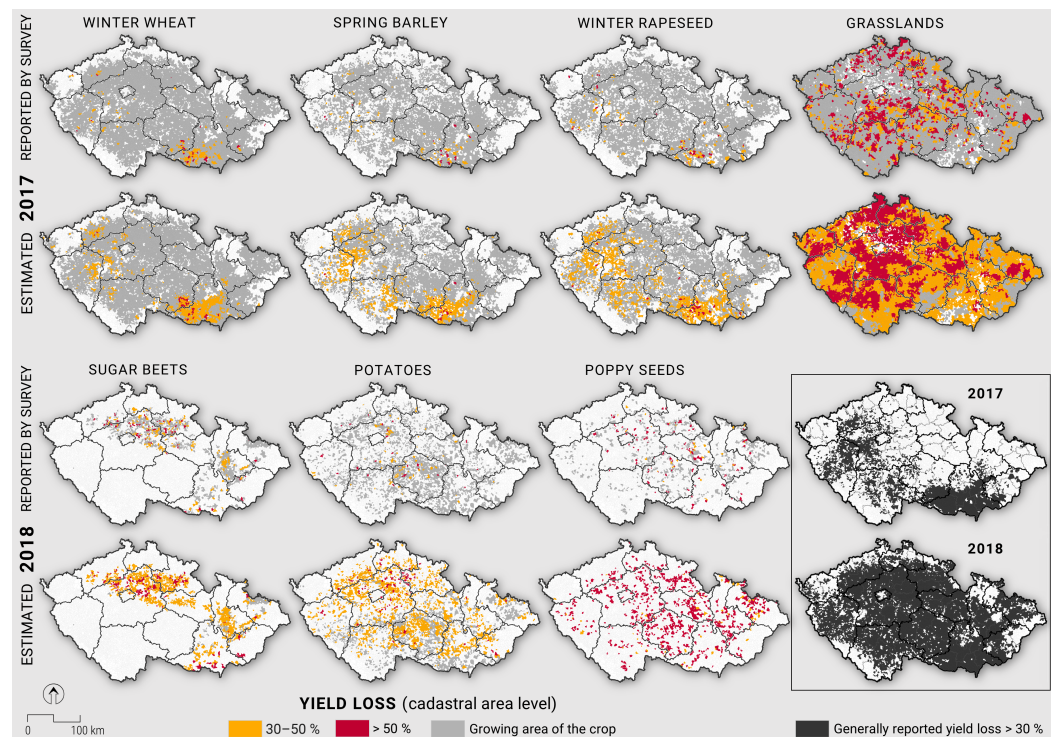


Figure 8. Final maps of the cadastral areas established for allocating grants in 2017 for winter wheat, spring barley and winter rapeseed and in 2018 for grasslands, sugar beets, potatoes and poppy seeds. There is a comparison between the upper maps based on the special questionnaire web survey and the lower maps based on estimations. In the bottom right, there are total yield losses reported by the special questionnaire web survey (without the Czech Drought Monitor) as greater than 30% for the years 2017 and 2018.

3.4. Applicability of the Method

The response to the 2015 drought involved compensation that was aggregated at the district (NUTS4) level. This leads to considerable discontent within the farming community, as it resulted in many false positive results (and payments), while many impacted farmers were left without any support [43] (Václav Hlaváček, Ph.D.—vice president of the Czech Agrarian Chamber).

The compensation method introduced in 2017 provided a considerable improvement in spatial resolution, from 78 districts to 13,078 cadastral areas, but required a far more detailed yield loss survey. This required significantly more time, and as a result, it took almost four months to prepare the data and an additional eight months to obtain approval from the European Commission and the national government and proceed with the drought payment compensation process.

However, once the method was set up and tested in 2017, the subsequent drought of 2018 was evaluated much faster. The government approved compensation for the 2018 drought in August 2018, crop yield loss data collection was conducted immediately afterward and, in mid-September, the cadastral area-based data were transferred to the Ministry of Agriculture and to the agency responsible for the compensation process, cutting the required time period from a year to two months. Therefore, the study also shows that this time lag can be easily avoided if proper methodologies and data are ready to be used on short notice, as was the case with the Czech drought monitoring system in early 2018.

4. Discussion

The presented ANN-based models exhibit practical applicability that enables the allocation of drought compensation funds immediately after harvest without a large delay. This is facilitated through the combination of predictors derived from both satellite imagery

and ground observations. ANNs allow for using a variety of predictors. At the same time, the approach allows for collecting and combining yield loss data from cadastral areas with different sources, i.e., not only from the emergency data collection process performed after a drought episode. It is very useful to include a gradually growing network of agricultural respondents that regularly report on drought status. It is therefore very appropriate to use such a yield loss model as the first criterion for compensation. Nevertheless, true losses should still be quantified in case the ANN method fails to determine whether the given cadastral area is eligible for compensation.

In both 2017 and 2018, the compensation decisions were not directly based on the outputs of the above-described models. For the farm to qualify for support, there were two options: a farmer could prove their damage (a) through certified yield reporting, which is administratively complicated both on the side of the farmer and at the Ministry of Agriculture or (b) through a simplified scheme. The simplified scheme clustered cadastral areas into three groups based on the above-described crop-specific models: (i) yield losses of more than 50%; (ii) yield losses between 50 and 30%; and (iii) yield losses less than 30%. The cadastral areas in the first two groups were eligible for compensation if the cadastral areas suffered from severe agricultural drought for at least three weeks during the key part of the growing season of the given crop and the vegetation condition (Figure 4) was worse than normal during at least part of the growing season of the studied crop, or if they had reliable continuous reports of at least 30% yield decreases from the certified reporter. In this way, only cadastral areas affected by drought and those with very likely drought-induced crop yield losses exceeding 30% were considered for compensation.

We have shown that the most efficient way to improve the model would be to improve the quality of the yield and crop loss data. The results confirm that the quality of the collected data is paramount for the tightness of the model fit with the given set of predictors and the overall robustness of the final model. The most efficient way to improve the model accuracy is to collect more accurate yield loss data, as more predictors can then be included in the first ANN layer. Another factor affecting the results was the distribution of the number of cadastral areas with compensation claims in the sample relative to their unknown true distribution. In general, the yield losses reported from the web survey showed relatively poor accuracy, which is shown by the difference between the submitted number of reports and the total number of report contacts.

Clearly, the motivation of survey respondents is much higher in drought-affected cadastral areas than in unaffected areas. This leads to difficulties in constructing any objective modeling framework that is dependent on high-quality inputs for model calibration. Other problems came from the web survey in cases when the participants filled in their damage values per farm, which represented multiple cadastral areas. An improved sample size can rapidly increase the prediction quality achieved for crop yields [44]. The random forest model for the prediction of wheat, barley and winter rapeseed had a prediction accuracy at a root mean square error (RMSE) level ranging from 360 to 420 kg/ha, and it improved greatly as the sample size increased. The data collection process can be aided by using the methods of Kang and Özdoğan [45], which would allow the county-level yield statistics to be downscaled through the Markov Chain Monte Carlo algorithm on leaf area index time series from Landsat. This approach could also provide an independent method for verifying crop loss data if properly used. Thus, we see the potential for improving the calculations if we were able to obtain more continuous questionnaires and ensure that they were being completed more accurately. A larger number of continuous reports would ensure that the same percentage distribution of cadastral areas between those eligible for compensation and those not eligible for compensation is used in the samples and outputs of the models, thus approximating the actual distribution. The method of calculating the estimates is already so complex that adding further conditions would probably not lead to a significant improvement.

Apart from focusing on collecting more representative and higher-quality yield data, further enlargement of the model inputs could be considered. For example, Guan et al. [46]

combined multiple sources of remotely sensed satellite data for large-scale crop monitoring and yield estimation for the U.S. Corn Belt. They included the EVI and ALEXI-based evapotranspiration, as in our study, but also estimated gross primary production based on the Global Ozone Monitoring Experiment-2 solar-induced fluorescence (SIF-GPP), QuikSCAT Ku-band radar backscatter and AMSR-E X-band passive microwave vegetation optical depth and used the partial least squares regression method for dimensionality reduction. The results of a validation conducted on the United States Department of Agriculture county-level crop yield statistics showed that most of the satellite-derived metrics shared common information related to above-ground crop biomass. However, Ku-band backscatter, thermal-based ALEXI-ET and X-band vegetation optical depth provide unique information on environmental stresses that improve the overall skill of crop yield prediction. Moreover, the additional use of ancillary climate data (e.g., precipitation and temperature) improved the model performance, in part because the crop reproduction stage related to the harvest index is highly sensitive to environmental stresses, but they are not fully captured by satellite data. The results presented for the Czech Republic have confirmed these findings as well.

Clearly, improvement would also be possible to achieve by using predictors with higher resolutions than those of MODIS, such as Landsat [47] or Sentinel-2 [48] imagery. As noted by Anderson et al. [49], a robust interpretation of the remote sensing results as a key crop yield estimator requires the integration of crop growth and pest models, which should correctly reflect air temperature and moisture differences in critical growing stages. An illustration of the potential improvement achieved by using higher-spatial-resolution satellite data comes from the work of Gaso et al. [50]. The authors estimated the crop yield of winter wheat by incorporating time series of vegetation indices from Landsat 7 and 8. The result of the simple regression method achieved an RMSE of 966 kg/ha, and the pixel size of the data (30 m) enabled the identification of within-field areas with yield potential. Gaso et al. [51] incorporated leaf area index retrieval from Sentinel-2 into their crop model to predict the within-field crop yield of soybeans. The model accuracy presented by the relative RMSE ranged from 28 to 51%, which confirmed the capability of the methodology to represent observed spatial yield patterns. Similarly, Hunt et al. [52] combined freely available Sentinel-2 data with environmental data (e.g., meteorological, topographical and soil moisture data) to estimate the within-field wheat yield variability in a single year via a random forest regression model. The validation conducted on harvester yield data achieved an RMSE of 660 kg/ha under the presence of within-field crop yield variability at a 10 m pixel size. Although ultra-high resolution images have been shown to improve estimates [53], they are generally only applicable to smaller areas and may not be feasible for use across the entire Czech Republic.

In our study, ANNs were applied, but further improvements could be achieved through the use of other machine learning algorithms and statistical methods for crop yield prediction. The most commonly used models are random forests, neural networks, linear regression models and gradient boosting trees, but convolutional neural networks, long short-term memory models and deep neural networks are also the most preferred deep learning algorithms [54]. The parallel use of multiple methods could lead to the development of multimodel ensembles. Franz et al. [55] combined the spatial information derived from vegetation conditions, soil water content and topography via the random forest method to predict maize crop yield at the field level. For their study, the historical yield maps were the best predictors, followed by crop condition, soil water content and elevation data, depending on the site.

A somewhat controversial point could involve shifting the threshold of the estimation boundary that decides the category to which we classify a given cadastral area—does it/does it not qualify for compensation. The data show that it is possible to slightly increase the overall model error while simultaneously reducing the partial error. This would lead to more farmers having to prove potential compensation claims in an administratively burdening way, but we would reduce the number of situations in which farmers who

should not have been entitled to compensation actually receive it; i.e., we would reduce an important partial error.

All of the above options should be considered when developing a methodology for evaluating the upcoming drought events that are predicted to affect large portions of the agricultural area, not only in the Czech Republic [7] but also globally [6].

5. Conclusions

This study demonstrates the potential of machine learning in combination with remote sensing and ground observation data to assess yield declines during a year with a severe drought. The results indicate that it is necessary to evaluate each such year separately, and for each dry year, it is necessary to obtain a new calibration set and tune the model to the event, as we did for 2017 and 2018. We have shown that the declines vary by crop depending on the course of the growing season. ANN-based models have great potential for practical applications, such as enabling the government to allocate drought compensation immediately after a harvest without a large time lag. It is therefore very appropriate to use such a yield loss model as a first criterion for allocating compensation.

Author Contributions: Conceptualization, J.M., P.H., Z.Z. and M.T.; methodology, J.M., J.B., M.B., D.S., P.H., Z.Z. and M.T.; software, J.M., J.B., M.B. and D.S.; validation, J.M., J.B., M.B. and D.S.; formal analysis, J.M., J.B., M.B. and D.S.; investigation, M.B., D.S., V.L. and Z.Z.; resources, Z.Z. and M.T.; data curation, J.M., J.B., M.B., D.S. and V.L.; writing—original draft preparation, J.M., J.B., M.B., D.S., V.L., F.J., K.K. and M.T.; writing—review and editing, P.H., M.C.A., W.D. and M.F.; visualization, J.M., M.B. and D.S.; supervision, Z.Z. and M.T.; project administration, M.T. and W.D.; funding acquisition, M.T. and W.D. All authors have read and agreed to the published version of the manuscript.

Funding: The study was conducted with the support of the SustES-Adaptation strategies for sustainable ecosystem services and food security under adverse environmental conditions (CZ.02.1.01/0.0/0.0/16_019/0000797). The work on the manuscript was additionally supported by the European Space Agency (ESA) project “YIPEEO: Yield prediction and estimation using Earth observation” (contract no. 4000141154/23/I-EF) and “DryPan: Novel EO data for improved agricultural drought impact forecasting in the Pannonian Basin” (contract nr: 4000127214/19/I-EF).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. Special questionnaire data are from The Ministry of Agriculture of the Czech Republic and are not open to public. Predictors data and reporting network data are from the cited Intersucho project and raw data are not open to public at the website.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The settings of the artificial neural networks (ANNs) for all crops in 2017, including the number of samples given by the number of different cadastral areas, the three-level ANN hierarchy and the predictors. A predictor value equal to 1/0 indicates that predictor was/was not used in the ANN.

Crop	Grain Maize	Oats	Spring Barley	Sugar Beets	Winter Barley	Winter Rapeseed	Winter Wheat
Samples for the ANN	231	59	368	114	262	528	799
ANN hierarchy	7-5-1	3-3-1	10-5-1	5-3-1	7-5-1	10-9-1	15-10-1
AWD1	0	0	0	0	0	0	1
AWR	1	0	1	0	1	1	1
AWR1	1	1	1	1	1	1	1
AWV	1	0	1	0	1	1	1
AWV1	0	0	1	0	0	1	1

Table A1. *Cont.*

Crop	Grain Maize	Oats	Spring Barley	Sugar Beets	Winter Barley	Winter Rapeseed	Winter Wheat
daysAWP_S4+	0	0	1	0	0	1	1
daysAWP1_S2+	0	0	0	0	0	0	1
daysAWP1_S4+	0	0	1	0	0	1	1
daysAWR1_30	0	0	0	0	0	0	1
EVI2	1	1	1	1	1	1	1
EVI2 _{PAAG}	1	1	1	1	1	1	1
NDVI	1	0	1	1	1	1	1
NDVI _{PAAG}	1	0	1	1	1	1	1
P	0	0	0	0	0	0	1
T	0	0	0	0	0	0	1

Table A2. The settings of the artificial neural networks (ANNs) used for all crops in 2018, including the number of samples given by the number of different cadastral areas, the three-level ANN hierarchy and the predictors. A predictor value equal to 1/0 indicates that predictor was/was not used in the ANN.

Crop	Alfalfa	Clover	Grain Maize	Grasslands	Hops	Oats	Poppy Seeds	Potatoes	Silage Maize	Spring Barley	Spring Wheat	Sugar Beets	Sunflowers	Winter Barley	Winter Rye	Winter Wheat
Samples for the ANN	713	826	700	4014	77	396	277	417	1854	1711	209	556	144	992	484	3839
ANN hierarchy	15	15	15	20	6	10	9	10	20	20	7	15	7	15	15	20
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	7	10	7	20	2	7	5	7	15	15	5	6	3	10	5	20
	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Alt	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1
AWD	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
AWD1	0	1	1	1	1	1	0	1	1	1	0	0	1	1	0	1
AWP	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AWP1	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	1
AWR	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
AWR1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1
daysAwp_S2+	1	0	1	0	0	0	0	0	1	1	1	0	0	0	0	1
daysAwp_S3+	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
daysAwp1_S2+	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
daysAwp1_S3+	0	1	1	0	0	0	0	0	1	1	1	0	1	1	0	0
daysAwr_50	1	1	0	1	0	0	1	1	1	1	0	1	0	1	1	1
daysAwr1_30	0	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1
daysAwr1_50	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
daysHeatDrought	1	0	1	1	0	0	1	1	1	0	0	1	0	0	0	1
daysTmax35	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0
ESI _{12WK}	1	0	1	1	0	0	0	0	1	1	0	1	0	0	1	1
ESI _{4WK}	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1
ET _o	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
ET _a /ET _o	0	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1
EVI2	1	1	1	1	0	1	0	0	1	1	0	0	0	1	1	1
EVI2 _{PAAG}	1	1	1	1	0	1	0	0	1	1	0	1	0	1	1	1
Lat	1	1	0	1	0	1	0	0	1	1	0	1	0	1	1	1
NDVI	0	1	0	0	0	0	1	0	1	1	1	0	0	1	1	1
NDVI _{PAAG}	0	0	0	0	1	0	1	0	1	1	0	0	1	1	1	1
P	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
P-ET _o	1	1	1	1	0	0	1	1	1	1	0	1	0	0	1	1
SWI ₀₋₄₀	0	1	1	1	0	1	0	1	1	0	0	1	1	1	0	1
SWI ₀₋₁₀₀	0	1	1	1	0	0	0	0	1	0	0	1	0	0	0	0
T	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1

References

1. Boyer, J.S.; Byrne, P.; Cassman, K.G.; Cooper, M.; Delmer, D.; Greene, T.; Gruis, F.; Habben, J.; Hausmann, N.; Kenny, N.; et al. The US drought of 2012 in perspective: A call to action. *Glob. Food Secur.* **2013**, *2*, 139–143. [CrossRef]
2. Dannenberg, M.P.; Yan, D.; Barnes, M.L.; Smith, W.K.; Johnston, M.R.; Scott, R.L.; Biederman, J.A.; Knowles, J.F.; Wang, X.; Duman, T.; et al. Exceptional heat and atmospheric dryness amplified losses of primary production during the 2020 US Southwest hot drought. *Glob. Chang. Biol.* **2022**, *28*, 4794–4806. [CrossRef]
3. Trnka, M.; Olesen, J.E.; Kersebaum, K.C.; Rötter, R.P.; Brázdil, R.; Eitzinger, J.; Jansen, S.; Skjelvåg, A.O.; Peltonen-Sainio, P.; Hlavinka, P.; et al. Changing regional weather crop yield relationships across Europe between 1901 and 2012. *Clim. Res.* **2016**, *70*, 195–214. [CrossRef]
4. Blauhut, V.; Stoelzle, M.; Ahopelto, L.; Brunner, M.I.; Teutschbein, C.; Wendt, D.E.; Akstinis, V.; Bakke, S.J.; Barker, L.J.; Bartošová, L.; et al. Lessons from the 2018–2019 European droughts: A collective need for unifying drought risk management. *Nat. Hazards Earth Syst. Sci.* **2022**, *22*, 2201–2217. [CrossRef]
5. Moravec, V.; Markonis, Y.; Trnka, M.; Hanel, M. Extreme Hydroclimatic Events Compromise Adaptation Planning in Agriculture Based on Long-term Trends. *Sci. Total Environ.* **2023**, submitted.
6. Trnka, M.; Feng, S.; Semenov, M.A.; Olesen, J.E.; Kersebaum, K.C.; Rötter, R.P.; Semerádová, D.; Klem, K.; Huang, W.; Ruiz-Ramos, M.; et al. Mitigation efforts will not fully alleviate the increase in water scarcity occurrence probability in wheat-producing areas. *Sci. Adv.* **2019**, *5*, eaau2406. [CrossRef] [PubMed]
7. Štěpánek, P.; Zahradníček, P.; Farda, A.; Skalák, P.; Trnka, M.; Meitner, J.; Rajdl, K. Projection of drought-inducing climate conditions in the Czech Republic according to Euro-CORDEX models. *Clim. Res.* **2016**, *70*, 179–193. [CrossRef]
8. Lhotka, O.; Trnka, M.; Kyselý, J.; Markonis, Y.; Balek, J.; Možný, M. Atmospheric circulation as a factor contributing to increasing drought severity in central Europe. *J. Geophys. Res. Atmos.* **2020**, *125*, e2019JD032269. [CrossRef]
9. Jaagus, J.; Aasa, A.; Aniskevich, S.; Boincean, B.; Bojariu, R.; Briede, A.; Danilovich, I.; Castro, F.D.; Dumitrescu, A.; Labuda, M.; et al. Long-term changes in drought indices in eastern and central Europe. *Int. J. Climatol.* **2022**, *42*, 225–249. [CrossRef]
10. The World Bank; Agriculture, Forestry, and Fishing, Value Added (% of GDP)-Czechia. Available online: <https://data.worldbank.org/indicator/NV.AGR.TOTL.ZS?locations=CZ> (accessed on 26 April 2023).
11. Papadimitriou, L.; Trnka, M.; Harrison, P.; Holman, I. Cross-sectoral and trans-national interactions in national-scale climate change impacts assessment—the case of the Czech Republic. *Reg. Environ. Chang.* **2019**, *19*, 2453–2464. [CrossRef]
12. Zahradníček, P.; Trnka, M.; Brázdil, R.; Možný, M.; Štěpánek, P.; Hlavinka, P.; Žalud, Z.; Malý, A.; Semerádová, D.; Dobrovolný, P.; et al. The extreme drought episode of August 2011–May 2012 in the Czech Republic. *Int. J. Climatol.* **2015**, *35*, 3335–3352. [CrossRef]
13. Büntgen, U.; Urban, O.; Krusic, P.J.; Rybníček, M.; Kolář, T.; Kyncl, T.; Ač, A.; Koňasová, E.; Čáslavský, J.; Esper, J.; et al. Recent European drought extremes beyond Common Era background variability. *Nat. Geosci.* **2021**, *14*, 190–196. [CrossRef]
14. Trnka, M.; Možný, M.; Jurečka, F.; Balek, J.; Semerádová, D.; Hlavinka, P.; Štěpánek, P.; Farda, A.; Skalák, P.; Cíenciala, E.; et al. Observed and estimated consequences of climate change for the fire weather regime in the moist-temperate climate of the Czech Republic. *Agric. For. Meteorol.* **2021**, *310*, 108583. [CrossRef]
15. Ceglar, A.; Toreti, A.; Prodhomme, C.; Zampieri, M.; Turco, M.; Doblas-Reyes, F.J. Land-surface initialisation improves seasonal climate prediction skill for maize yield forecast. *Sci. Rep.* **2018**, *8*, 1322. [CrossRef]
16. Kogan, F.; Kussul, N.; Adamenko, T.; Skakun, S.; Kravchenko, O.; Kryvobok, O.; Shelestov, A.; Kolotii, A.; Kussul, O.; Lavrenyuk, A. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 192–203. [CrossRef]
17. Son, N.T.; Chen, C.F.; Chen, C.R.; Minh, V.Q.; Trung, N.H. A comparative analysis of multitemporal MODIS EVI and NDVI data for large-scale rice yield estimation. *Agric. For. Meteorol.* **2014**, *197*, 52–64. [CrossRef]
18. Büechli, E.; Fischer, M.; Crocetti, L.; Trnka, M.; Grlj, A.; Zappa, L.; Wouter, D. Crop yield anomaly forecasting in the Pannonian Basin using gradient boosting and its performance in years of severe drought. *Agric. For. Meteorol.* **2023**, submitted.
19. De Wit, A. *Regional Crop Yield Forecasting Using Probabilistic Crop Growth Modelling and Remote Sensing Data Assimilation*; Wageningen University and Research; ProQuest Dissertations Publishing: Ann Arbor, MI, USA, 2007.
20. Dorigo, W.A.; Zurita-Milla, R.; de Wit, A.J.; Brazile, J.; Singh, R.; Schaepman, M.E. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 165–193. [CrossRef]
21. 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. A high-resolution, integrated system for rice yield forecasting at district level. *Agric. Syst.* **2019**, *168*, 181–190. [CrossRef]
22. Trnka, M.; Hlavinka, P.; Možný, M.; Semerádová, D.; Štěpánek, P.; Balek, J.; Bartošová, L.; Zahradníček, P.; Bláhová, M.; Skalák, P.; et al. Czech Drought Monitor System for monitoring and forecasting agricultural drought and drought impacts. *Int. J. Climatol.* **2020**, *40*, 5941–5958. [CrossRef]
23. Intersucho. Available online: <https://www.intersucho.cz/> (accessed on 26 April 2023).
24. Svoboda, M.; LeComte, D.; Hayes, M.; Heim, R.; Gleason, K.; Angel, J.; Rippey, B.; Tinker, R.; Palecki, M.; Stooksbury, D.; et al. The drought monitor. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 1181–1190. [CrossRef]

25. Anderson, M.C.; Hain, C.R.; Jurečka, F.; Trnka, M.; Hlavinka, P.; Dulaney, W.; Otkin, J.A.; Johnson, D.; Gao, F. Relationships between the evaporative stress index and winter wheat and spring barley yield anomalies in the Czech Republic. *Clim. Res.* **2016**, *70*, 215–230. [CrossRef]
26. Jurečka, F.; Fischer, M.; Hlavinka, P.; Balek, J.; Semerádová, D.; Bláhová, M.; Anderson, M.C.; Hain, C.; Žalud, Z.; Trnka, M. 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.* **2021**, *256*, 107064. [CrossRef]
27. Bartošová, L.; Fischer, M.; Balek, J.; Bláhová, M.; Kudláčková, L.; Chuchma, F.; Hlavinka, P.; Možný, M.; Zahradníček, P.; Wall, N.; et al. Validity and reliability of drought reporters in estimating soil water content and drought impacts in central Europe. *Agric. For. Meteorol.* **2022**, *315*, 108808. [CrossRef]
28. Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207. [CrossRef]
29. Crocetti, L.; Forkel, M.; Fischer, M.; Jurečka, F.; Grlj, A.; Salentinig, A.; Trnka, M.; Anderson, M.; Ng, W.T.; Kokalj, Ž.; Bucur, A. Earth Observation for agricultural drought monitoring in the Pannonian Basin (southeastern Europe): Current state and future directions. *Reg. Environ. Chang.* **2020**, *20*, 123. [CrossRef]
30. Wardlow, B.D.; Anderson, M.C.; Verdin, J.P. *Remote Sensing of Drought: Innovative Monitoring Approaches*; CRC Press: Boca Raton, FL, USA, 2012.
31. Becker-Reshef, I.; Vermote, E.; Lindeman, M.; Justice, C. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sens. Environ.* **2010**, *114*, 1312–1323. [CrossRef]
32. Hayes, M.J.; Decker, W.L. Using NOAA AVHRR data to estimate maize production in the United States Corn Belt. *Remote Sens.* **1996**, *17*, 3189–3200. [CrossRef]
33. Jackson, T.J.; Chen, D.; Cosh, M.; Li, F.; Anderson, M.; Walthall, C.; Doriaswamy, P.; Hunt, E.R. Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sens. Environ.* **2004**, *92*, 475–482. [CrossRef]
34. Kastens, J.H.; Kastens, T.L.; Kastens, D.L.; Price, K.P.; Martinko, E.A.; Lee, R.Y. Image masking for crop yield forecasting using AVHRR NDVI time series imagery. *Remote Sens. Environ.* **2005**, *99*, 341–356. [CrossRef]
35. Doraiswamy, P.C.; Sinclair, T.R.; Hollinger, S.; Akhmedov, B.; Stern, A.; Prueger, J. Application of MODIS derived parameters for regional crop yield assessment. *Remote Sens. Environ.* **2005**, *97*, 192–202. [CrossRef]
36. Johnson, D.M. An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* **2014**, *141*, 116–128. [CrossRef]
37. Johnson, D.M. A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 65–81. [CrossRef]
38. Ferencz, C.; Bogner, P.; Lichtenberger, J.; Hamar, D.; Tarcsai, G.; Timár, G.; Molnár, G.; Pásztor, S.Z.; Steinbach, P.; Székely, B.; et al. Crop yield estimation by satellite remote sensing. *Int. J. Remote Sens.* **2004**, *25*, 4113–4149. [CrossRef]
39. Labus, M.P.; Nielsen, G.A.; Lawrence, R.L.; Engel, R.; Long, D.S. Wheat yield estimates using multi-temporal NDVI satellite imagery. *Int. J. Remote Sens.* **2002**, *23*, 4169–4180. [CrossRef]
40. The Ministry of Agriculture of the Czech Republic; Public Land Registry-LPIS. Available online: <https://eagri.cz/public/app/lpisext/lpis/verejny2/plpis/> (accessed on 26 April 2023).
41. Moravec, V.; Markonis, Y.; Rakovec, O.; Svoboda, M.; Trnka, M.; Kumar, R.; Hanel, M. Europe under multi-year droughts: How severe was the 2014–2018 drought period? *Environ. Res. Lett.* **2021**, *16*, 034062. [CrossRef]
42. Rakovec, O.; Samaniego, L.; Hari, V.; Markonis, Y.; Moravec, V.; Thober, S.; Hanel, M.; Kumar, R. The 2018–2020 Multi-year drought sets a new benchmark in Europe. *Earth's Future* **2022**, *10*, e2021EF002394. [CrossRef]
43. Hlaváček, V. (Czech Agrarian Chamber, Czech Republic). Personal communication, 2022.
44. Filippi, P.; Jones, E.J.; Wimalathunge, N.S.; Somarathna, P.D.; Pozza, L.E.; Ugbaje, S.U.; Jephcott, T.G.; Paterson, S.E.; Whelan, B.M.; Bishop, T.F. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precis. Agric.* **2019**, *20*, 1015–1029. [CrossRef]
45. Kang, Y.; Özdoğan, M. Field-level crop yield mapping with Landsat using a hierarchical data assimilation approach. *Remote Sens. Environ.* **2019**, *228*, 144–163. [CrossRef]
46. Guan, K.; Wu, J.; Kimball, J.S.; Anderson, M.C.; Frohling, S.; Li, B.; Hain, C.R.; Lobell, D.B. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sens. Environ.* **2017**, *199*, 333–349. [CrossRef]
47. Doraiswamy, P.C.; Hatfield, J.L.; Jackson, T.J.; Akhmedov, B.; Prueger, J.; Stern, A. Crop condition and yield simulations using Landsat and MODIS. *Remote Sens. Environ.* **2004**, *92*, 548–559. [CrossRef]
48. Inglada, J.; Vincent, A.; Arias, M.; Marais-Sicre, C. Improved early crop type identification by joint use of high temporal resolution SAR and optical image time series. *Remote Sens.* **2016**, *8*, 362. [CrossRef]
49. Anderson, M.C.; Neale, C.M.U.; Li, F.; Norman, J.M.; Kustas, W.P.; Jayanthi, H.; Chavez, J.O.S.E. Upscaling ground observations of vegetation water content, canopy height, and leaf area index during SMEX02 using aircraft and Landsat imagery. *Remote Sens. Environ.* **2004**, *92*, 447–464. [CrossRef]
50. Gaso, D.V.; Berger, A.G.; Ciganda, V.S. Predicting wheat grain yield and spatial variability at field scale using a simple regression or a crop model in conjunction with Landsat images. *Comput. Electron. Agric.* **2019**, *159*, 75–83. [CrossRef]

51. Gaso, D.V.; de Wit, A.; Berger, A.G.; Kooistra, L. Predicting within-field soybean yield variability by coupling Sentinel-2 leaf area index with a crop growth model. *Agric. For. Meteorol.* **2021**, *308*, 108553. [[CrossRef](#)]
52. Hunt, M.L.; Blackburn, G.A.; Carrasco, L.; Redhead, J.W.; Rowl, C.S. High resolution wheat yield mapping using Sentinel-2. *Remote Sens. Environ.* **2019**, *233*, 111410. [[CrossRef](#)]
53. Sharifi, A. Estimation of biophysical parameters in wheat crops in Golestan province using ultra-high resolution images. *Remote Sens. Lett.* **2018**, *9*, 559–568. [[CrossRef](#)]
54. Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [[CrossRef](#)]
55. Franz, T.E.; Pokal, S.; Gibson, J.P.; Zhou, Y.; Gholizadeh, H.; Tenorio, F.A.; Rudnick, D.; Heeren, D.; McCabe, M.; Ziliani, M.; et al. The role of topography, soil, and remotely sensed vegetation condition towards predicting crop yield. *Field Crop. Res.* **2020**, *252*, 107788. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.