

TURKISH JOURNAL OF REMOTE SENSING
(Türkiye Uzaktan Algılama Dergisi)
e-ISSN 2687-4997
https://dergipark.org.tr/en/pub/tuzal

Wheat yield prediction for sustainable food security using Sentinel-2 and machine learning

Neslişah Civelek[1], Levent Genç[*2], Özgün Akçay[3]

[1] *Çanakkale Onsekiz Mart University, Department of Geographic Information Technologies, Çanakkale, Türkiye;* neslisahcivelek8@gmail.com
[2] *Çanakkale Onsekiz Mart University, Faculty of Architecture and Design, Department of Urban and Regional Planning, Çanakkale, Türkiye;* leventgc@comu.edu.tr
[3] *Çanakkale Onsekiz Mart University, Faculty of Engineering, Department of Geomatic Engineering, Çanakkale, Türkiye;* akcay@comu.edu.tr

Check for updates

Article History:
Received: 7 August 2025
Revised: 11 September 2025
Accepted: 29 September 2025
Published: 25 March 2026

Copyright: © 2026 by the authors.
This article is an open access article distributed under terms and conditions of the Creative Commons Attribution (CC BY-SA) license. https://creativecommons.org/licenses/by-sa/4.0/

Abstract: Accurate pre-harvest wheat yield estimation is critically important for sustainable agricultural policies. In this study, 13 vegetation indices derived from Sentinel-2A/B images (10 from the literature and 3 derived for this study), temperature, precipitation, and sunshine duration data from the ERA-5 dataset, and ground measurements were used. Pearson correlation analysis was applied to determine the most suitable period for yield estimation, and the most effective factors were identified using the Decision Tree (DT) model. Considering these factors, yield estimation was performed using the Multiple Linear Regression (MLR) model. The results showed that the factors most affecting wheat yield were NDVI, WDRVI, RDVI, and Red Edge 1 in April; Derived Index 3 in December; and LAI, average temperature, and total precipitation in March. The heading-flowering period has been determined as the most suitable time for yield estimation. There is a high degree of agreement between the actual and predicted yields ($R = 0.92$, $R^2 = 0.86$).

Keywords: Machine learning; decision tree; ERA-5; Sentinel-2; winter wheat yield estimation

Citation: Civelek, R., Genç, L., & Akçay, Ö. (2026). Wheat yield prediction for sustainable food security using Sentinel-2 and machine learning. *Turk. J. Remote Sens.,* 8, 1760511. https://doi.org/10.51489/tuzal.1760511

1. Introduction

Research underscores the need to adopt innovative technologies and approaches to ascertain and enhance the spatial distribution of winter wheat output and yield per unit area (Zhen et al., 2020; Ma et al., 2022; Jamali et al., 2024; Ashfaq et al., 2024; Guo et al., 2024; Du et al., 2025). From this perspective, the integration of Remote Sensing (RS) and Machine Learning (ML) will significantly enhance the development of global sustainable agriculture policies by enabling rapid and precision predictions, thereby increasing economic resilience and food security. RS technologies are vital tools for observing the Earth's surface and can assist digital agriculture and sustainable food security at different spatial, radiometric, and temporal resolutions (Toscano et al., 2019; Pandey & Pandey, 2023; Choudhary et al., 2022; Khurshid et al., 2024).

In the last 50 years, NASA, the European Space Agency (ESA), and the private sector have provided satellite images with sufficient resolution for the monitoring and analysis of agricultural crops in fields and regions (Rouse, 1973; Lee et al., 2010; Araus & Cairns, 2014; Holzman & Rivas, 2016; Kingra et al., 2016; Sibanda et al., 2019; Weiss et al., 2020; Cheng et al., 2022; Singh et al., 2022; Vidican et al., 2023). The evaluation of the physiological and physical characteristics of plants over time, along with environmental variables, is made possible through the use of Vegetation Indices (VI) derived from satellite images (Thenkabail et al., 2000; Oppelt & Mauser, 2004; Glenn et al., 2008; Roberts et al., 2018; Saric et al., 2022;

*Corresponding Author; Levent Genç /leventgc@comu.edu.tr

De Almeida et al., 2023; Zheng et al., 2023; Haque et al., 2024; Panumonwatee et al., 2025; Xu et al., 2025).

By utilizing VIs and correlating them with factors that influence plant wheat yield, it becomes possible to forecast yields for the purpose of ensuring sustainable food security. (Kogan & Kogan, 2019; Shmelev et al., 2021; Joshi et al., 2023; Zhou et al., 2023; Devkota et al., 2024; Lu et al., 2024; Karongo et al., 2025; Raza et al., 2025). Normalized Difference Vegetation Index (NDVI) (Rouse, 1973), Green Normalized Difference Vegetation Index (GNDVI) (Gitelson et al., 1996), and various other reliable VIs have been used for yield prediction studies. The NDVI, GNDVI, and RDVI indices used in the study provide information about plant health and density (Adeniyi et al., 2020). NDVI looks at the overall health of plants on a scale from (-1) to (1), where healthy and thick plants have high positive numbers, and unhealthy or thin plants have lower numbers based on how dense they are (Civelek et al., 2023). GNDVI is more suitable for detecting chlorophyll content and nitrogen stress due to the use of the green band (Jamali et al., 2024). RDVI, on the other hand, is the normalized version of NDVI, increasing sensitivity in low-density vegetation and reducing saturation in high-density areas (Saad El Imanni et al., 2022). EVI is better at fixing soil reflections, giving better information about chlorophyll and water content, and studying thick vegetation that is affected by the atmosphere; in contrast, MSAVI was created to lessen soil effects in areas with sparse vegetation (Cheng et al., 2022; Liu et al., 2022). WDRVI, on the other hand, is similar in structure to NDVI and reduces the saturation problem by being optimized according to different vegetation density levels with a weighting coefficient (Naqvi et al., 2019). The Clrededge, RE-1, RE-3, LAI, DI-1, DI-2, and DI-3 indices created from the red edge bands of Sentinel-2A/B provide ideal data for the precise measurement of plant physiology, leaf chlorophyll content, and structural parameters (Gitelson et al., 2004; Afrasiabian et al., 2021). It offers high sensitivity in determining phenological processes (Faqe Ibrahim et al., 2023). While creating the formula for the LAI index, the NDVI formula was used, and the LAI-NDVI relationship was established at the global level (Waters et al., 2002; Saad El Imanni et al., 2022).

The application of diverse, high-volume data, such as RS, through ML algorithms enables reliable and comprehensive evaluations across numerous scales and geographies (Brenton et al., 2022). The complexity of such data necessitates adept preprocessing and management of multidimensional datasets (Goldberg et al., 2021; Liu et al., 2022). Open-access platforms like Google Earth Engine (GEE) have been utilized to analyse this type of data, capitalizing on their large data banks and formidable computational capabilities (Shelestov et al., 2017; Liu, 2020; Tian et al., 2019; Cao et al., 2021; Pan et al., 2021; Chowdhury et al., 2023; Singha & Swain, 2023; Zhang et al., 2023; Arora, 2025).

To ascertain the optimal timing for the prediction of winter wheat yield, researchers have integrated VIs derived from satellite imagery and biomass data, such as the Leaf Area Index (LAI), into the Random Forest (RF) algorithm for yield estimation. It was determined that the optimal timing for the anticipated winter wheat production was during the heading-flowering phase in April (Ma et al., 2022; Newete et al., 2024). Furthermore, climate data with RS data is an effective component for agricultural monitoring and plant yield prediction, as it directly influences plant phenology and biological processes (Thieme et al., 2020; Bouras et al., 2023). Thus, the amalgamation of multispectral RS imaging with meteorological data has become crucial in enhancing the accuracy of yield predictions (Dobermann & Ping, 2004; Prasad et al., 2007; Skakun et al., 2017; Zhu et al., 2021; Chen & Tao, 2022; Haseeb et al., 2025; Hou et al., 2025; Raza et al., 2025). In studies analysing the effectiveness of crop yield predictions using RS data, climate data, and a combination of both through ML, it has been stated that the integration of climate and remote sensing datasets produces superior prediction outcomes (Peng et al., 2024; Ashfaq et al., 2024). However, analyzing climatic data gathered from meteorological stations has been a hard task, particularly in rural areas.

The ERA-5 Land Daily Aggregated (LDA) climate dataset, derived from satellite data by the European Centre for Medium-Range Weather Forecasts (ECMWF), offers complimentary access to daily meteorological data, including temperature, humidity, and pressure, at a resolution of 9 km (Muñoz-Sabater et al., 2021; Bojanowski et al., 2022; EEDC, 2025). Studies have been done by researchers is notable results using ERA-5 LDA and VIs for agricultural plant yield prediction. Bojanowski et al. (2022) reported that they were able to predict plant yields using ERA-5 images based meteorological data along with the XGBoost method. Segarra et al. (2020) indicated in their analyses employing Multiple Linear Regression (MLR) with VIs, alongside climatic variables (precipitation, temperature, and Growing Degree Days) and topographic data (slope, elevation, aspect), that environmental factors contributed to 11%-20% of yield variability, with the heading-blooming period identified as the most dependable phenological predictor.

The ERA-5 meteorological data used to study provides an important input for winter wheat yield prediction. In the research by Zhao et al. (2023) in China, it was found that predicting winter wheat yields could be done accurately and precisely ($R2 = 0.81$; $p < 0.01$) at low cost using satellite data (Sentinel-2 and Landsat 8) and ERA-5 meteorological data. Similarly, Bora et al. (2024) found that the GDD ($R^2 = 0.95$) and PTU (Photothermal Units) ($R^2 = 0.95$) data derived from the ERA-5 dataset exhibited high prediction accuracy compared to the HTU (Heliotermal Units) ($R^2 = 0.32$) data obtained from meteorological stations.

We need to tackle the problem of connecting winter wheat yield predictions to Sustainable Development Goals (SDGs) by using RS-based VIs and ERA-5 climate data with ML. This strategy is cost-effective and efficiently delivers valuable information that can enhance food security. The SDGs particularly SDG-1 'No Poverty,' SDG-2 'Zero Hunger,' and SDG-13 'Climate Action,' focus on prioritizing global initiatives to eliminate poverty, ensure food security, and create agricultural systems that can withstand climate challenges (Scown & Nicholas, 2020; AIMS, 2024). A substantial opportunity for individuals engaged in agricultural decision-making aligned with the SDGs to improve food security through the utilization of RS and ML methodologies, for early yield prediction purposes (Segarra et al., 2020; Goldberg et al., 2021; Ma et al., 2022; Graf et al., 2023; Peng et al., 2024). This emphasis is vital for the sustainable cultivation of agricultural crops and the proficient management of crises, particularly in improving the impacts of disruptions to food supply chains, natural disasters, climate change, pandemics, conflicts, and political turmoil (Ogundeji, 2022; Shams Esfandabadi et al., 2022; Zhou et al., 2023; UNESCO, 2024; AIMS, 2024; FAO, 2024; GEOGLAM, 2024; GRFC, 2024). To meet global food demand by 2050, food output must rise by 60% in industrialized nations and 100% in developing countries, including Türkiye (Raza et al., 2025; GRFC, 2024). This circumstance highlights the essential need for intervention and the insufficiency of existing measures for food security (Baldos & Hertel, 2014; Kumar, 2016; Thieme et al., 2020; Zhou et al., 2023). Based on the literature there are a few studies using multispectral RS data-based VIs and ERA-5 climate data together with ML algorithm to predict winter wheat yield. No study to use regression decision tree (RDT) algorithm using VIs and ERA-5 data together to predict winter wheat yield and link the result SDGs.

The objective of this research is to make a prediction regarding winter wheat production by utilizing VIs from Sentinel-2, meteorological data from ERA-5 satellite, and ground data. The prediction will be accomplished through the utilization of ML, RDT and MLR analyses. The secondary objective of this study is to assess the ways in which accurate winter wheat yield predictions are tied to the specific SDGs that are associated with the issue of global food security in the North Aegean region of Türkiye.

2. Material and Metod

2.1. Study area

The study area encompasses 18 villages (Kumkale, Halileli, Dümrek, Tevfikiye, Gökçalı, Çıplak, Civler, Erenköy, Kalafat, Akçapınar, Pınarbaşı, Taştepe, Mahmudiye, Kumburun, Yeniköy, Üvecik, Çamoba, and Bozköy) in the Central and Ezine districts of Çanakkale province. As a result of field visits conducted in November 2023, 220 winter wheat plots have been identified for ground data collection (Figure 1). The province's yearly average temperature is 15.2 degrees Celsius (°C) and surpasses 35 °C multiple times during the summer months. Most of the precipitation occurred between autumn and winter, with an average annual total of 623.2 mm. Given the meteorological attributes noted in the region, the cultivation of winter wheat, barley, and tomatoes is prominent in the province. Çanakkale is one of the important winter wheat production regions, possessing suitable climatic characteristics for winter wheat cultivation and sufficient water resources, which makes it an ideal location for the study.

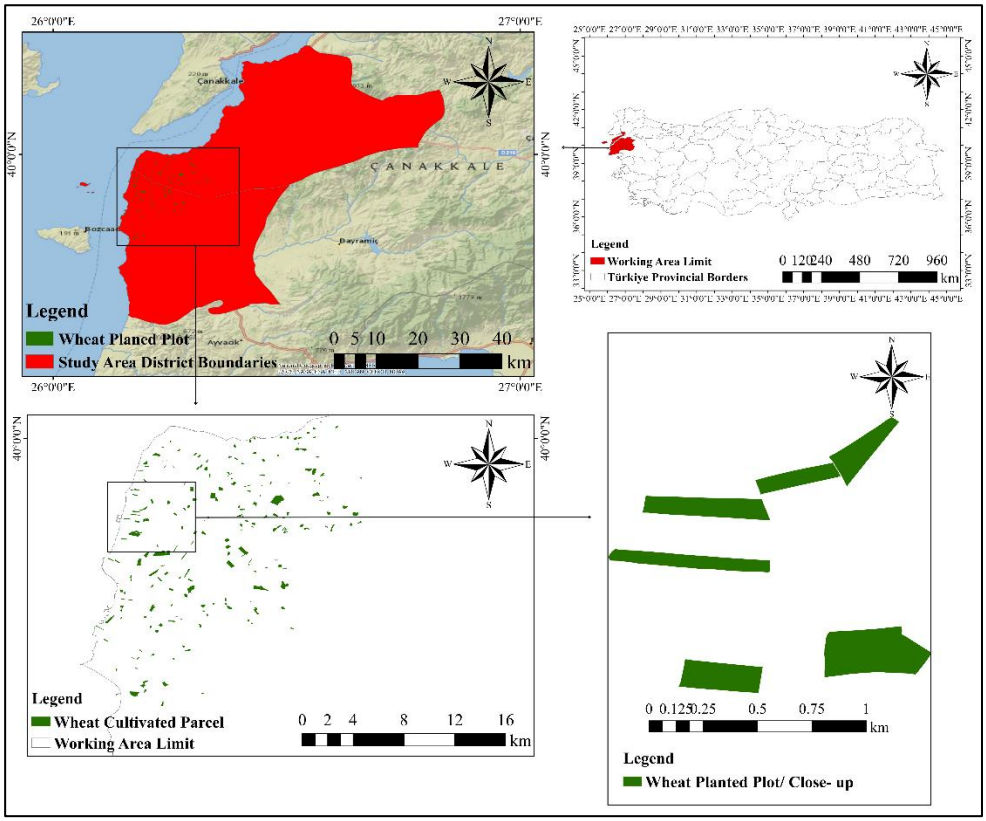


Figure 1. Study area

2.2. Data and pre-processing

In this study, 13 vegetation indices (10 from the literature and 3 derived for this study) obtained from Sentinel-2A/B satellite images, meteorological data (temperature, precipitation, and sunshine duration) from the ERA-5 LDA dataset, and field measurements were used to estimate winter wheat yield. All data were sampled at monthly and spatial resolution across 220 wheat plots in the study area. For data processing, GEE, RapidMiner, ArcGIS, and Excel software were used; GEE was used for data access and VI calculations, RapidMiner for data analysis, ArcGIS for spatial analysis, and Excel for data management. Satellite imagery was accessed through GEE, and no preprocessing was applied other than a 20% cloud filter and 10-meter scaling. The 13 VIs used are listed with their calculation formulas in Table 1. Meteorological variables were accessed from the ERA-5 LDA dataset through GEE, and time series features were extracted along with plant indices,

meteorological data, and land data. The heading-flowering period was determined using field observations and literature references, and the most suitable period for yield estimation was identified through Pearson correlation analysis.

Yield estimation was performed using a two-stage modeling approach. First, variable importance analysis was conducted using the Decision Tree operator in RapidMiner, with hyperparameters set as criterion = least_square, maximum depth = 5, minimum gain = 0.01, and minimum leaf size = 2. Note that RapidMiner's Decision Tree operator automatically implements a regression tree algorithm, similar to standard regression tree methods. Subsequently, an MLR model was built using the most effective variables; interaction terms were not added, and multicollinearity was considered using min tolerance = 0.05 and ridge = 1.0E-8. The dataset was split into training (60%) and testing (40%) sets, and cross-validation was implemented using the default Holdout method in RapidMiner. Model performance was evaluated using the t-statistic and p-value, along with the R and R^2 metrics. All these steps are shown in Figure 2 to ensure reproducibility and transparency. This study is the first, to our knowledge, to use a sequential DT-MLR approach for winter wheat yield estimation, enabling both accurate predictions and interpretable variable effects.

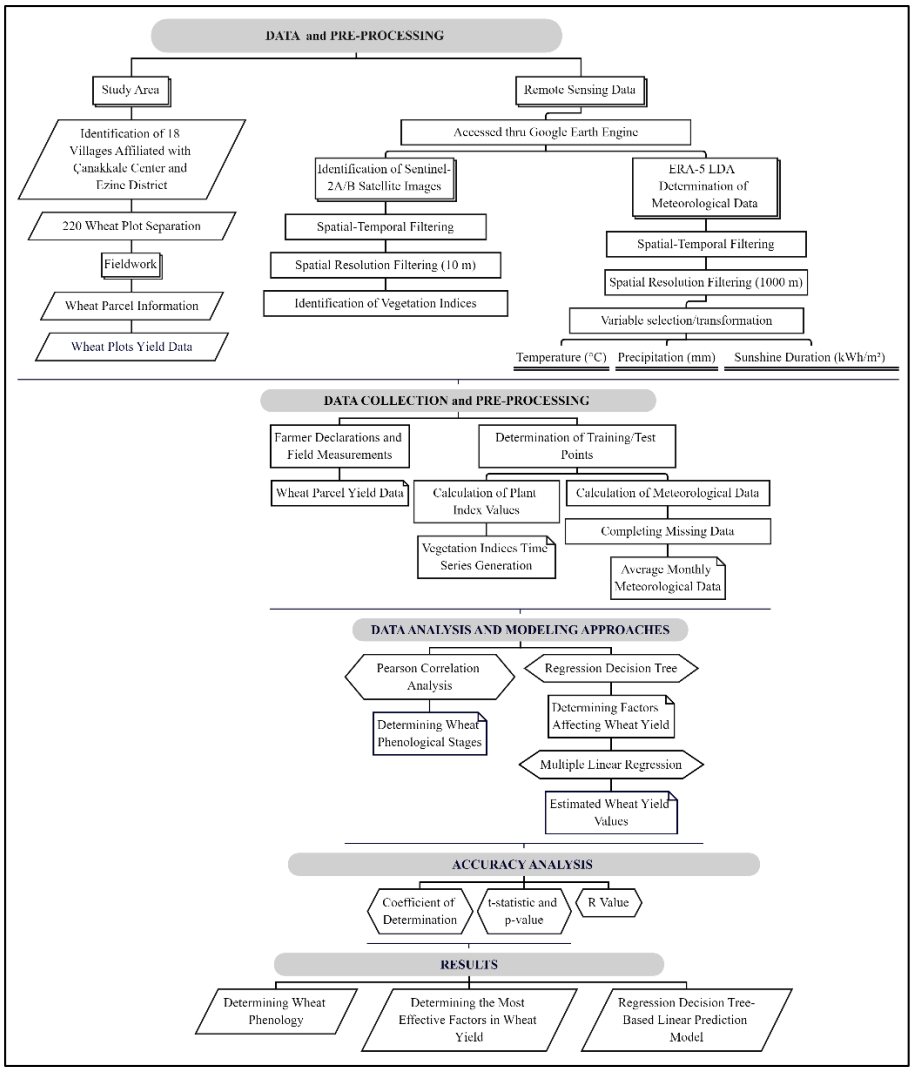


Figure 2. Flow chart of the study

2.3. Field data

Field visits were conducted during the planted-harvest period (December to June) for the 2023/2024 season to observe the phenological development of the 220-winter wheat-planted plots designated for this study (Figure 3). At the same time, we collected yield data

from authorized individual parcels and farmers using interviews. During our field trips, access was gained access to 40 plots, which represents 18.18% of the total 220 plots cultivated with winter wheat, although we harvested crops from only 9 of those plots. The prevalent winter wheat varieties in the region are Lucilla and Masaccio. The research region has an average annual precipitation of 625 mm; yet certain parcels were irrigated throughout the growing season. Furthermore, 25 kg of diammonium phosphate (DAP) and 25 kg of urea fertilizer were administered per decare (da) in each plot, with harvesting planned for June 2024. One-meter segments were systematically collected from nine distinct plots to evaluate the yield metrics of the winter wheat fields. The quantity of ears per cultivated area varies between 400 and 800. In the laboratory, we conducted a triadic count of the stems and grains in each ear. After each count, we randomly selected and measured 1000 units. It was computed the yield figures for the winter wheat plots after the weighing and replicated this procedure for nine plots.

We calculated the yield values by selecting a 1 square meter sample area within the parcel. In this calculation, the spike density per square meter, the average number of grains per spike, and the weight of 1000 grains were considered, and the obtained values were scaled to per decare. A sample calculation is presented below: The density of ears per square meter is 410, with each ear containing 25 grains. The average weight of 10 ears, based on a 1000-grain weight (across 3 repetitions) of 40 grams, yields 472 kg/da (as declared by the farmer: 450 kg/da). It was noted that there was a strong correlation between the computed yields and those reported by the farmers and authorized officials. Consequently, no additional sampling was required for the remaining plots, and the data was included in the database directly based on farmer declarations, as a high degree of agreement was observed between the calculated yields and farmer declarations. It was noted a strong correlation between the computed yields and those reported by the farmers and authorized officials. Consequently, yield data were incorporated with the remaining 211 wheat plots into the database.

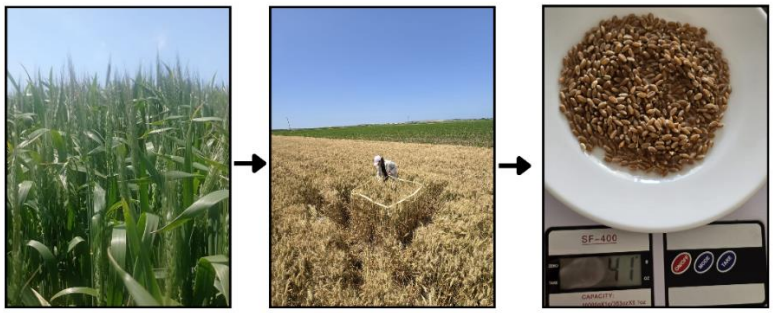


Figure 3. Field visit images and wheat plots yield values calculation process

Identification of Training/ Test Points: Training/test points were determined in vector format for 220 wheat plots, each measuring 10 m x 10 m (one pixel). The GPS coordinates and identification numbers (ID) of each point (Point *i, i* = 1, 2, 3…, 220) were defined, and a feature collection was created in the GEE software environment (Figure 4). The purpose of the feature collection is to calculate the spectral values of each point over a time series of satellite images and to perform spatial-temporal analysis. Training and test points help analyse the spectral values for each pixel, and the high resolution of satellite images makes the results very accurate, like what you would get from field measurements (Raza et al., 2025).

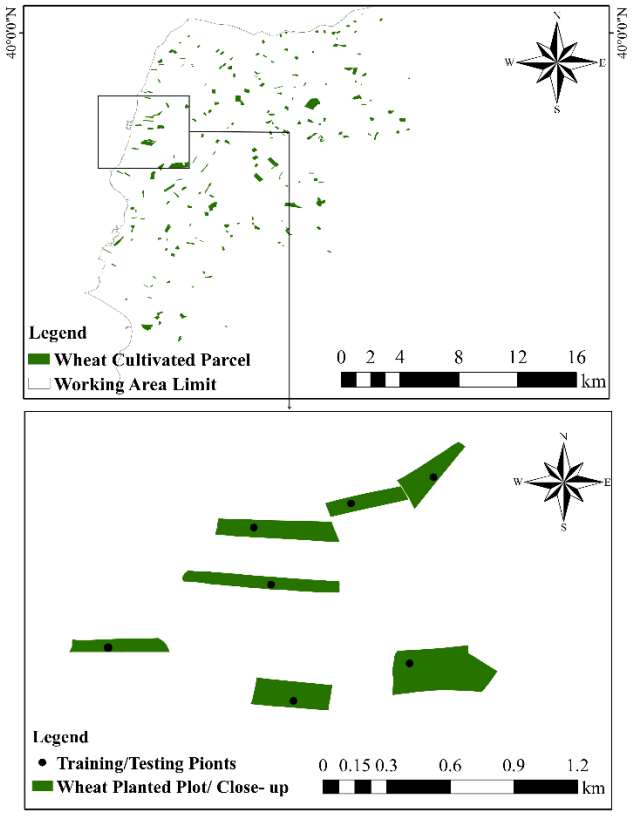


Figure 4. Location and close-up of training-test points

2.4. Satellte based data and preprocessing

In the study, Sentinel-2A/B satellite data covering the period from December 2023 to May 2024 were used. These data were accessed from the GEE cloud platform (Ayub et al., 2022). Cloud filtering and scaling preprocessing were applied to ensure the reliability and accuracy of the data. To get accurate surface reflectance values, a cloud filter was applied to keep cloud cover below 20%, and the satellite bands B2 (blue), B3 (green), B4 (red), B5, B6, B7 (red edge bands), B8 (NIR), B8A, B11, and B12 (shortwave infrared) were adjusted to a resolution of 10 m for analyzing spectral information. Considering the phenological process of winter wheat and the availability of images, a time series was created from satellite images dated 12.28.2023, 01.05.2024, 02.06.2024, 03.27.2024, 04.09.2024, and 05.21.2024.

Vegetation Indices: In this study, based on national and international literature reviews, the studies were classified considering criteria such as the purpose of using plant indices, the accuracy values of the studies, and commonly used band combinations. As a result of evaluating the literature and conducting a bibliometric analysis of agricultural production studies in the Web of Science (WOS) database, the most suitable indices for the study's purpose were determined. Current research indicates that indices derived from combining spectral bands with mathematical calculations can provide spectral information for estimating crop data (Zhen et al., 2020). In this context, three different VIs were derived in the study to optimize the reflectance values of the red edge bands of the Sentinel-2A/B satellite. In the study, NDVI (Wang et al., 2015), GNDVI (Jamali et al., 2024), RDVI (Xing et al., 2020), EVI (Saad El Imanni et al., 2022), MSAVI (Segarra et al., 2020), WDRVI (Naqvi et al., 2019), Clrededge (Gitelson et al., 2003), RE-1, RE-3 indices (Afrasiabian et al., 2021), LAI (Gumma et al., 2022), and the DI-1, DI-2, and DI-3 indices derived from the red edge (B5, B6, B7) bands were calculated using the formulas provided in Table 1.

Table 1. Plant index formulas used in the study

Vegetation Index	Formula	Reference
NDVI	((B8 - B4))/ (B8 + B4)	Rouse (1973)
GNDVI	(B8−B3) / (B8 +B3)	Gitelson et al. 1996)
RDVI	((B8 - B4))/ √(B8 + B4)	Roujean et al. (1995)
EVI	2.5 x (B8-B4))/(B8+(6 x B4)-(7.5xB2)+1)	Huete et al. (2002)
MSAVI	(2 x B8 +1-√(2 x B8 +1)2 -8(B8-B4))/2	Qi et al. (1994)
WDRVI	(0.1 x B8-B4)/(0.1 x B8+B4)	Gitelson (2004)
Clrededge	$\left(\frac{B8}{B5}\right)$-1	Gitelson et al. (2003)
RE-1	$\frac{B8 - B5}{B8 + B5}$	Delegido et al. (2013)
RE-3	$\frac{B8 - B7}{B8 + B7}$	Delegido et al. (2013)
LAI	(3.618x($\frac{B8-B4}{B8+B4}$)-0.118)	Myneni & Williams (1994)
DI -1	$\frac{((B5+B6+B7)-(B4))}{((B5+B6+B7)+(B4))}$	This study
DI -2	((B5+B7)-(*B4*))/((B5+B7)+B4	This study
DI -3	((B6+B7)-(B4))/((B6+B7)+B4	This study

The VI values were calculated monthly for the December–May period, corresponding to 220 different spatial points used for training and testing, and the analysis of VI values in the phenology process was conducted (Table 2). The minimum and maximum VI values for each month in the dataset are given in Table 3.

Table 2. Sample representation of the dataset of monthly vegetation index values of training/test points

Vegetation Index$_i$ (i=1,2…220)	December	January	Februaru	March	April	May
NDVI$_1$	0.087	0.163	0.348	0.537	0.708	0.241
NDVI$_{220}$	0.076	0.084	0.245	0.536	0.731	0.303
GNDVI$_1$	0.159	0.196	0.344	0.484	0.64	0.329
GNDVI$_{220}$	0.156	0.171	0.278	0.488	0.65	0.364
GNDVI$_1$	0.159	0.196	0.344	0.484	0.64	0.329
GNDVI$_{220}$	0.156	0.171	0.278	0.488	0.65	0.364
RDVI$_1$	0.063	0.131	0.31	0.462	0.62	0.159
RDVI$_{220}$	0.042	0.032	0.208	0.458	0.648	0.222
EVI$_1$	0.276	0.708	1.183	2.379	2.738	0.625
EVI$_{220}$	0.229	0.255	0.721	2.193	2.732	0.784
MSAVI$_1$	0.159	0.28	0.516	0.698	0.829	0.389
MSAVI$_{220}$	0.141	0.155	0.393	0.698	0.844	0.465
WDRVI$_1$	-0.787	-0.756	-0.657	-0.502	-0.262	-0.719
WDRVI$_{220}$	-0.791	-0.788	-0.717	-0.502	-0.217	-0.685
Clrededge$_1$	0.189	0.389	1.068	2.315	1.85	0.636
Clrededge$_{220}$	0.165	0.183	0.648	2.313	5.429	0.87
RE-1$_1$	0.082	0.151	0.341	0.526	0.688	0.203
RE-1$_{220}$	0.056	0.056	0.229	0.52	0.717	0.27
RE-3$_1$	0.063	0.131	0.31	0.462	0.62	0.159
RE-3$_{220}$	0.042	0.032	0.208	0.458	0.648	0.222

Table 2. Cont.

Vegetation Index$_i$ (i=1,2…220)	December	January	Februaru	March	April	May
LAI$_1$	0.697	0.833	1.283	1.99	2.967	1.1
LAI$_{220}$	0.681	0.693	1.008	1.988	3.129	1.155
DI -1$_1$	0.546	0.584	0.682	0.76	0.837	0.608
DI -1$_{220}$	0.53	0.528	0.626	0.757	0.851	0.645
DI -2$_1$	0.388	0.429	0.544	0.643	0.751	0.462
DI -2$_{220}$	0.369	0.369	0.476	0.64	0.772	0.507
DI-3$_1$	0.396	0.453	0.595	0.712	0.812	0.486
DI-3$_{220}$	0.376	0.372	0.515	0.708	0.83	0.536

Table 3. Minimum and maximum VI values for each month in the dataset

Vegetation Index	December	January	Februaru	March	April	May
NDVImin-max	0.06–0.60	0.06–0.65	0.07–0.67	0.39–0.56	0.53–0.75	0.10–0.58
GNDVImin-max	0.121–0.538	0.107–0.608	0.028–0.617	0.363–0.523	0.509–0.698	0.185–0.537
RDVImin-max	0.014–0.564	0.023–0.602	0.028–0.617	0.336–0.499	0.464–0.694	0.06–0.504
EVImin-max	0.172–2.781	0.23–3.2	0.216–2.62	2.003–2.938	2.111–2.906	0.351–2.17
WDRVImin-max	-0.797 –-0.411	-0.795 –-0.324	-0.795 –-0.301	-0.626 –-0.469	-0.503 –-0.175	-0.779 –-0.448
Clrededgemin-max	0.128–3.172	0.141–4.106	0.153–4.376	2.007–2.998	1.85–6.024	0.241–2.811
RE-1min-max	0.021–0.606	0.04–0.66	0.037–0.672	0.386–0.559	0.524–0.742	0.091–0.583
RE-3min-max	0.014–0.564	0.023–0.602	0.336–0.499	0.464–0.694	0.06–0.504	0.06–0.504
LAImin-max	0.656–2.379	0.664–2.731	0.673–2.17	1.001–1.997	1.986–3.3	0.732–2.224
DI-1min-max	0.514–0.804	0.516–0.825	0.523–0.833	0.694–0.777	0.759–0.867	0.55–0.784
DI-2min-max	0.348–0.698	0.349–0.731	0.362–0.741	0.557–0.666	0.64–0.791	0.396–0.681
DI-3min-max	0.352–0.769	0.363–0.798	0.362–0.806	0.622–0.734	0.712–0.849	0.399–0.743

2.5. Meteorological data-ERA-LDA

The ‘temperature sum' (°C), ' total_precipitation_sum ' (mm), and 'surface_solar_radiation_downwards_sum' (kWh/m^2) related to sunlight duration were sourced from the ERA5-LDA dataset. The daily mean values of these climatic variables were calculated and subsequently converted into monthly mean values on images. When determining the date intervals for calculating meteorological parameters, the dates of the Sentinel-2A/B satellite images were used as a reference, with each interval set to 30 days prior to the image date. These dates: 11.28.2023-12.28.2023, 12.05.2023-01.05.2024, 01.06.2024-02.06.2024, 02.27.2024-03.27.2024, 03.09.2024-04.09.2024, and 04.21.2024-05.21.2024. Satellite images have been resampled at a 1 km scale. The obtained values have been calculated for areas with a pixel size of 10 m x 10 m This situation reveals the spatial differences of meteorological variables in detail at different scales (Bojanowski et al., 2022). However, missing data were encountered at the training/test points due to the high resolution and coverage of the ERA5-LDA dataset. This situation has created the necessity to complete the missing data to ensure the continuity and reliability of the analyses. Therefore, we used the data imputation method, considering the overall average (Equation 1). The missing data ($X_{\text{missing data}}$) for each meteorological variable in the dataset using the overall average value(s) within the relevant period (Can, 2024). An example representation of the meteorological data set is provided in Table 4.

$$X_{missing\ data} = X = \overline{X} + \frac{1}{n}\sum_{i=1}^{n} X_i \quad (1)$$

Table 4. Monthly average values of temperature (°C), precipitation (mm) and sunshine duration (kWh/m²) of training/test points

Meteorological variable$_i$ (i=1,2…220)	December	January	February	March	April	May
Temperature$_1$	12.18	8.826	7.33	11.36	13.49	16.94
…	…	…	…	…	…	…
Temperature$_{220}$	11.77	11.11	6.89	11.06	13.26	16.84
Precipitation$_1$	100.77	32.61	72.16	58.79	42.15	20.03
…	…	…	…	…	…	…
Precipitation$_{220}$	139.85	95.36	73.01	60.61	45.37	22.93
Sunshine Duration$_1$	5980.6	6100.9	8244.9	1584.42	18938.1	22425.98
…	…	…	…	…	…	…
Sunshine Duration$_{220}$	5937.1	6068.5	8243.8	15521.3	18938.1	22369.8

2.6. Data analysis

2.6.1. Pearson correlation analysis

To determine the optimal timing for estimating winter wheat yield, the relationship between monthly temperature, total rainfall, and sunshine duration (from December to May) and actual yield values was examined by analyzing 220 data points using the Pearson Correlation Coefficient (R). The calculation was performed in accordance with Equation 2. In the given formula, Yi, represents the observed (actual) values, $\hat{Y}$i, represents the predicted values, $\bar{Y}$, represents the average of the observed values, $\bar{\hat{Y}}$, represents the average of the predicted values, and n is the number of observations.

$$R = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2}} \quad (2)$$

2.6.2. Regression decision trees

The RDT method was used to determine the factors affecting winter wheat yield. The Least Squares Deviation Method (LSDM) was used as the splitting criterion to determine the split points of the model. LSDM is a method aimed at minimizing the total squared errors to select the most suitable partition point within the RDT. was used (Equations 3 and 4). According to the given formula, C is the splitting criterion, Rj are the subsets formed as a result of the split, y_i is the true value of the data point, $\widehat{y_{\mathrm{Rj}}}$is the average predicted value in the subset, and $\sum i \epsilon \mathrm{Rj}$ represents the sum over all data points in the subset.

$$C = \sum_{i=1}^{2} \sum_{i \epsilon Rj} (y_i - \widehat{y_{Rj}})^2 \quad (3)$$

$$\widehat{Y_{Rj}} = \frac{1}{|Rj|} \sum_{i \epsilon Rj} y_i \quad (4)$$

2.6.3. Multiple linear regression

MLR, among ML methods, is the most fundamental supervised learning method. The regression method was used to calculate the impact coefficients of the factors obtained through MLR management and to make predictions. Equation 5 guided the calculation. In Equation (15), $\hat{Y}$, represents the estimated yield value, βo the constant coefficient, β_i, the regression coefficients of the independent variables, and X_i, the values of the independent variables.

$$\widehat{Y} = \beta o + \sum_{i=1}^{n} (\beta_i X_i) \quad (5)$$

Although both RDT and MLR are regression methods, they are used sequentially due to their different advantages. RDT can naturally capture non-linear relationships and interactions between variables and is robust to outliers; however, it is not possible to directly

interpret the direction and magnitude of the variables' effects on the prediction. MLR, on the other hand, allows for the interpretation of the linear effects of independent variables using β coefficients. Therefore, while complex relationships are determined with RDT, the effects of significant variables can be explained with MLR; thus, the sequential use of the two methods makes the modeling both powerful and interpretable.

2.6.4. Assessment of model performance and validation

Various evaluation metrics were used to determine the reliability and effectiveness of the methods employed in the study. These include statistical criteria such as the Coefficient of Determination (R^2), t-statistics, and p-value (Sharma et al., 2013; Chai et al., 2014; Sykas et al., 2022). Equation 6 calculates R^2, a measure of how accurate the RDT is. This measures how much of the change in the dependent variable is explained by the independent variables in the regression models and how well the model works (Zhang, 2011). According to Equation 6, $\hat{y}_i$represents the predicted value for observation *i*, and y_i, represents the actual value of the observation.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2} \quad (6)$$

We applied T-statistics and p-value to evaluate the level of statistical significance (Kul, 2014; Akbulut, 2022). These two statistical measures were evaluated together to understand and interpret the results of the hypothesis test (Akbulut, 2022). These two statistical measures are complementary tools for interpreting the results of a hypothesis test. The t-statistic expresses the effect of the tested variable in the sample, while the p-value determines the significance level corresponding to this statistic (Kul, 2014). Additionally, the p-value is calculated based on the t-statistic results (Akbulut, 2022). Therefore, to ensure that the evaluation in the study is conducted accurately and comprehensively, both criteria have been interpreted together. According to the p-value results, if $p < 0.05$, the hypothesis was considered significant (Quintano et al., 2024). Therefore, to calculate the p-value and determine the significance level of the results, the t-statistic (t) and the t-distribution were first calculated (Equation 7) (Kul, 2014). According to Equation (7), $\overline{X}$, represents the sample mean; μ, the population mean; n, the sample size; s, the standard deviation; and $SE\left(\frac{s}{\sqrt{n}}\right)$, the standard error (Kul, 2014). From the t and degrees of freedom (df) derived from Equation (7), the cumulative distribution function (F) of the t-distribution and the p-value were calculated (Equation 8). F (t, df), the cumulative probability corresponding to the given t-value and degrees of freedom, |t|, is the absolute value of the t-statistic.

$$t = \frac{\bar{X} + \mu}{SE\left(\frac{s}{\sqrt{n}}\right)} \quad (7)$$

$$p = 2 \, x \, (1\text{-}F \, (|t|), \, df) \quad (8)$$

2.7. Modelling scenarios

Identification of Factors Affecting Winter Wheat Yield Using DT: In the process of winter wheat phenology, a RDT model was created to determine the VIs and meteorological variables affecting yield values. The VIs values and meteorological variables, calculated monthly from the datasets used in the model, were determined as independent variables. We added the actual yield values as a numerical attribute to the relevant dataset and designated them as the dependent variable. A RDT model was created in accordance with the working principle.

Winter Wheat Yield Prediction Model: The outputs of the RDT model, which identifies the factors affecting winter wheat yield, were used as data input in the model for

predicting winter wheat yield. In this case, MLR analysis was conducted to calculate the "βi" of the factors affecting yield on yield. Later, the significance of the "βi" of the independent variables on the dependent variable was evaluated according to the p-value criteria ($p < 0.05$). Statistically insignificant variables ($p > 0.05$) were retained in the yield prediction model because they interacted with significant variables and contributed to the model's prediction performance. Using the obtained "βi" values, the predicted winter wheat yield values were calculated using the MLR method.

3. Result

3.1. Winter wheat phenology

The winter wheat phenology observed in the research area has been analyzed in four primary stages: December to January and the initial week of February encompass the seeding-germination phase; the final three weeks of February and March pertain to tillering-jointing; April and the first week of May represent the heading-blooming period; and May to June signify the maturity-harvesting stages. The harvest phase was incorporated into the phenological calendar due to its observation during field trips in the initial three weeks of June. The results derived from RS data encompass the period from December to May. Figure 5 presents the phenological process and the phenology calendar of winter wheat (Mashonganyika et al., 2021). The technique was determined by analyzing the outcomes of the R analysis about VIs values, meteorological data, and winter wheat yield data.

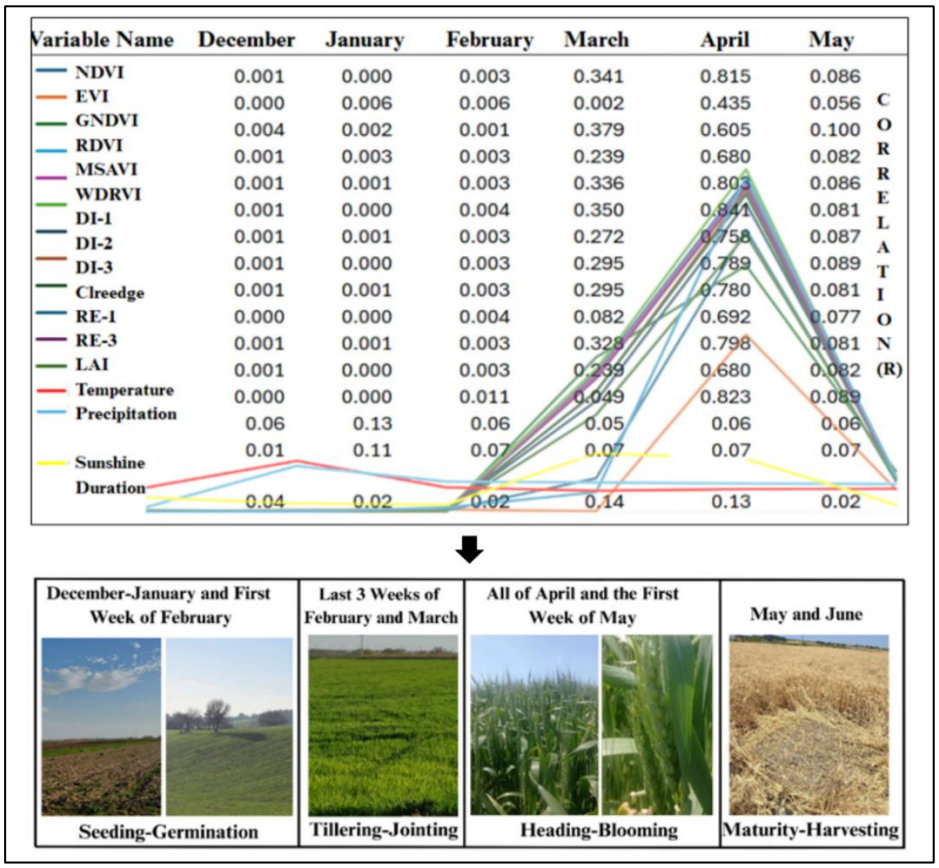


Figure 5. Field images for determining the phenological process and the phenology calendar of winter wheat

The phenological process of winter wheat begins with the seeding-germination phase in December and January and the first week of February (Saad El Imanni et al., 2022). During the seeding-germination stage, the green plant cover has not yet fully developed. Therefore,

the plant index values and their correlation with yield are at the lowest level. At this stage, temperature and precipitation levels are at a certain level and are more effective compared to other variables. In the second stage, called the tillering-jointing phase, which lasts from the last weeks of February to March, plant growth and the amount of plant material produced are important for how much they can yield (Tian et al., 2019). At this stage, we observe a positive rise in the index values (Peng et al., 2024). At this stage, while the values of the variable temperature generally remain constant, there is a gradual decrease in the amount of precipitation; conversely, there is an increase in sunshine duration. The third stage, the heading-blooming phase, is a strong indicator of final yield results because it is influenced by grain filling and directly affects grain number and size (Amin et al., 2022). As a result, the index values reach their peak, positively influencing the yield. At this stage, the amount of rainfall decreases, and the duration of sunshine is the most effective variable. Therefore, the most suitable time for estimating the winter wheat yield is the heading-blooming period. During this period, the plant experiences physiological stress and reaches its peak in terms of spectral reflection. The final two weeks of May are known as the maturity-harvesting phase. The leaves of the winter wheat plant turn straw-coloured, and the grains become very hard (Zhang et al., 2023). This situation indicates that winter wheat is losing green biomass and maturing (Da Silva et al., 2020). After the harvest, the index values are at their lowest level. While temperature and rainfall maintain a stable relationship with winter wheat yield, the relationship between sunshine duration and yield has gradually begun to diminish.

3.2. Impact of vegetation indices and meteorological data on winter wheat yield

To determine the factors affecting winter wheat yield, when the RDT model was examined, the effective factors identified were 'Average Temperature (°C) - December,' 'Average Total Precipitation (mm) - December,' 'NDVI - April Values,' 'WDRVI - April Values,' 'RDVI - April Values,' 'RE-1 - April Values,' 'DI-3 - December Values,' and 'LAI - March Values. 'Model performance was calculated as R = 0.89 and R^2 = 0.80. These accuracy values indicated that the model could identify the factors affecting wheat yield with high accuracy and reliability. Due to the excessive number of branches after the root node, two separate sub-decision trees were created, one for the right branch and one for the left branch, using the first branch as a reference (Figure 7, Figure 8).

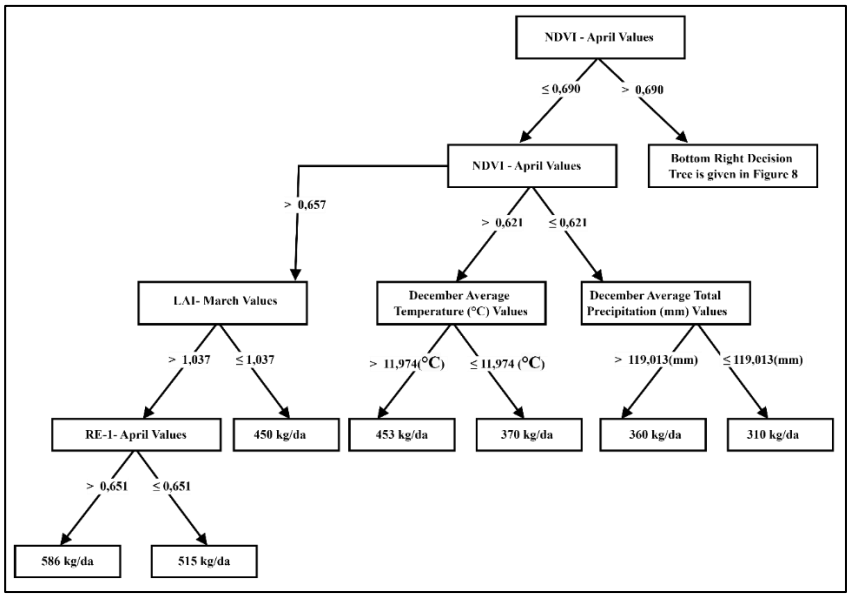


Figure 7. Bottom right decision tree of the decision tree model created to determine the factors affecting winter wheat yield

When looking at the left side of the tree after the first branching from the root node (Figure 7), March LAI values were checked for when April NDVI is greater than 0.657; if LAI is greater than 1.037, the April RED-1 values were used; if RE-1 is greater than 0.651, the yield is 586 kg/da. If RE-1 is 0.651 or less, the yield is estimated to be 515 kg/da. If March LAI is 1.037 or less, the yield is estimated to be 450 kg/da. In the case of RE-1 ≤ 0.651, the yield was estimated to be 515 kg/da. In the case of March with LAI ≤ 1.037, the yield was estimated to be 450 kg/da. In April, when NDVI > 0.621, the average temperature of December was evaluated. If the temperature was > 11.974 °C, the yield was estimated to be 453 kg/da, while if the temperature was ≤ 11.974 °C, the yield was estimated to be 370 kg/da. In the case of NDVI ≤ 0.621, the total rainfall amount for December was considered. If precipitation > 119.013 mm, the yield was estimated at 453 kg/da; if precipitation ≤ 119.013 mm, the yield was estimated at 310 kg/da.

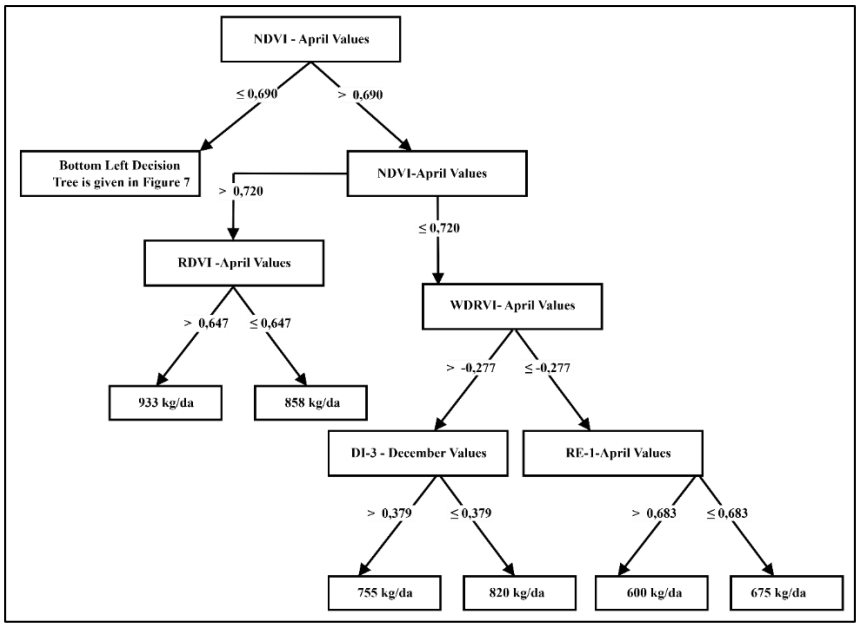


Figure 8. Bottom left decision tree of the decision tree model created to determine the factors affecting winter wheat yield

When examining the sub-decision tree created for the right side after the first branching from the root node (Figure 8), if the NDVI values for April were greater than 0.690, the model was subjected to detailed examination. In this case, when NDVI>0.720, the April RDVI values were evaluated; if RDVI>0.647, the yield was estimated to be 933 kg/da, and if RDVI ≤ 0.647, the yield was estimated to be 858 kg/da. In the case of NDVI≤ 0.720, the WDRVI values for April were considered. If WDRVI > -0.277, the December TI-3 values were examined; if DI-3 > 0.379, the yield was estimated to be 755 kg/da, and if DI-3 ≤ 0.379, the yield was estimated to be 820 kg/da. In the case of WDRVI ≤-0.277, the RE-1 values for April were evaluated; if RE-1 > 0.683, the yield was estimated to be 600 kg/da, and if RE-1 ≤ 0.683, the yield was estimated to be 675 kg/da.

The effects of the factors on winter wheat yield were measured using the MLR method, and their importance was assessed (Table 5). When the analysis results in Table 5 are examined, the average temperature variable for the interval has a positive and significant effect $\beta i = 62.38$, $p < 0.05$. The total rainfall amount in December is negatively impactful but not statistically significant $\beta i = -0.10$, $p > 0.05$. The NDVI variable for the month of April shows a negative and highly statistically significant effect $\beta i = -6967.62$, $p < 0.05$. With $\beta i = 7151.3$, $p < 0.05$, the WDRVI variable for April is positive and statistically significant. The variables

for April RDVI βi = 273.75, RE-1 βi = -592.22, December DI-3 βi = 15.68, and March LAI βi = -0.19 were not found to be statistically significant p > 0.05. The intercept coefficient of the model is statistically significant at a very high level βi = 7037.11, p < 0.05.

Table 5. Coefficient values and significance status obtained from the regression decision tree

Factors affecting winter wheat yield	Regression Coefficient Value (βi)	T-Statistics	P-Value	Significance Status
December Average Temperature (°C)	62.38292024	2.837	0.00499	*
December Average Precipitation (mm)	-0.109491885	-0.464	0.643	NS
NDVI-April Values	-6967.627675	-4.685	5.02E-06	***
WDRVI- April Values	7151.35613	7.956	1.07E-13	***
RDVI- April Values	273.7540227	0.874	0.383	NS
RE-1- April Values	-592.2274243	-0.773	0.441	NS
DI-3- December Values	15.68577656	0.352	0.725	NS
LAI-March Values	-0.192473172	-0.013	0.989	NS
Intersection Coefficient	7037.117483	5.612	6.24E-08	***

If 0.01≤ p, it is statistically significant (*), if 0.001≤ p, it is highly statistically significant (**), if p<0.001, it is very highly statistically significant (***), if 0.05≤ p<0.10, there is borderline significance or a tendency towards significance (+), if p>0.10, there is no statistical significance (NS: not significant) (Akbulut, 2022).

3.3. Winter wheat yield prediction

Considering the effect coefficients of the independent variables determined by the regression method using decision trees, the estimated yield values of the winter wheat plots were predicted with high accuracy. The relationship between the actual yield values and the predicted yield values is shown in Figure 9. When the graph was examined, it was observed that there is a forceful linear relationship between the actual yield and the predicted yield values (R = 0.92). At the same time, the calculated R^2 = 0.86 value showed that the predicted yield values explained the actual yield values by 86%. The findings obtained proved that the yield prediction model showed strong and reliable results.

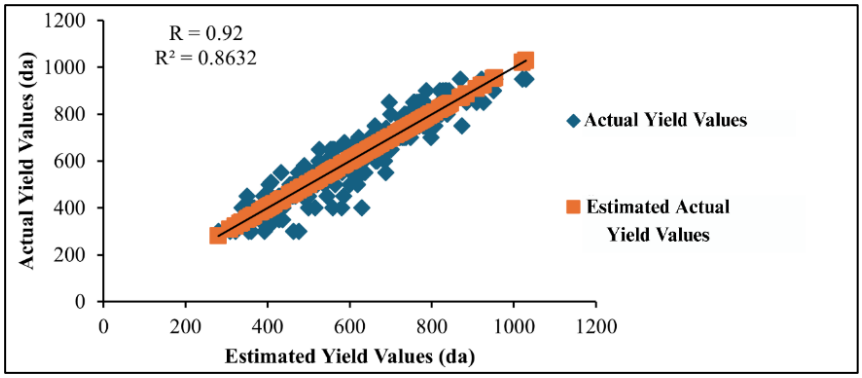


Figure 9. The graph shows the linear relationship between the actual yield values (da) and the estimated yield values (da) and the R and R^2 values in the upper left corner of the graph

4. Discussion

The main objective of this study is to determine the factors affecting winter wheat yield and the most suitable time for prediction using RS, meteorological parameters, and terrestrial data with ML techniques. In semi-arid regions, determining the appropriate time interval for training models is crucial for accurately predicting winter wheat yield (Cheng et al., 2022). Plant indices obtained from satellite images, influenced by meteorological parameters, exhibit varying performances of winter wheat at different phenological stages (Xiao et al., 2024). The effectiveness of VI in predicting crop yield depends on the development status of the vegetation and its chlorophyll content (Raza et al., 2025). In particular, the low reflection of plants in the early and late growth stages can make all VI useless during these times, which lowers the accuracy of yield predictions. (Haseeb et al., 2025). Therefore, this study has focused on making yield predictions by considering the entire crop season and the peak phenological stages. Thus, the study analyzed the variability

in how VI contributes to yield prediction under different phenological stages and changing environmental conditions, ensuring a more accurate determination of the phenological process of winter wheat. As emphasized in the study by Newete et al. (2024), plant index values are low at the beginning of the phenological process, increase as the process continues, and then decrease again after the heading-blooming stages. So, it has been found that the VI from the Sentinel-2A/B satellite directly affects how we identify the phenological stages of winter wheat. This approach is also consistent with other studies that reveal the relationship between VI and winter wheat yield. For example, Faqe Ibrahim et al. (2023) conducted a study on yield prediction and found a significant relationship between VI and yield (R^2 = 0.77). Similarly, Bouras et al. (2023) stated that the integration of VI into yield prediction models increased the accuracy of the predictions by 70%. Meteorological data significantly influence the continuous development process of winter wheat from sowing to harvest. In this context, the linear relationships between winter wheat yield, VI, and meteorological variables were simultaneously evaluated in this study.

In wheat yield prediction, researchers have observed that increased sunshine duration during the heading-flowering period significantly affects plant development and yield, in addition to identifying average temperature and total precipitation during the sowing-germination period as effective factors. These findings indicate that ERA-5 climate data provide reliable results at different phenological stages (Muñoz-Sabater et al., 2021; Zhao et al., 2023). The study shows that ERA-5 data is very accurate and reliable for predicting winter wheat yield and is a useful source for calculating weather-related factors. The obtained results are consistent with the studies in the literature. Zhao et al. (2023) showed that combining ERA-5 data with satellite observations is a good way to accurately predict wheat yield, achieving a high accuracy score (R^2: 0.81) and a significant result ($p < 0.01$). Bora et al. (2024) used heat units to predict the heading-flowering period of wheat to evaluate the applicability of ERA-5 climate data. In the study, the climate data from the ERA-5 dataset (R^2: 0.95) was much better at predicting outcomes than the Heliothermal Unit (HTU) (R^2: 0.32). In this context, the findings obtained in the study once again demonstrate the effectiveness of ERA-5 data in predicting wheat yield. In the early stages of phenology, meteorological variables such as temperature and precipitation were more effective, but this effect diminished as the process progressed, and RS data became more efficient. The joint evaluation of VI and meteorological data has revealed significant relationships with winter wheat yield. Therefore, we concluded that evaluating only RS data or only meteorological data would be insufficient in determining the phenological process of winter wheat. The findings suggest that an integrated analysis of both data sources is necessary, as they jointly affect winter wheat yield. Similarly, Ashfaq et al. (2024), in their study predicting winter wheat yield, found that the model using both climate and RS data (R^2 = 0.78-0.88) provided higher accuracy and reliability compared to the model using only climate data (R^2 = 0.65-0.732) or the model using only RS data (R^2 = 0.49-0.70). Similarly, in the study by Jin et al. (2022), they looked at different combinations of 13 variables (like indices, soil properties, and weather data) and found that adding RS data improved the model's prediction accuracy by 92.8%. The use of multiple datasets is also a critical point in determining the most suitable time for yield prediction by better capturing the interactions between plant growth dynamics and VI throughout the phenological stages of winter wheat. As emphasized by Du et al. (2025), the month of April (heading-blooming) has emerged as the most suitable and crucial period.

Factors affecting winter wheat yield were analyzed using the RDT model (R = 0.89 and R^2 = 0.80). It was found that the NDVI, WDRVI, RDVI, and RE-1 values from April, the DI-3 values from December, and the LAI values from March are the most important factors for predicting winter wheat yield. At the same time, it has been determined that the temperature and precipitation variables observed in December play a critical role in low yield predictions. As a result of the modeling conducted by considering the impact coefficients of the factors

affecting winter wheat yield, the estimated yield values were obtained with high accuracy values (R = 0.92, R^2 = 0.86). The obtained results have allowed for a clearer and more reliable identification of the dynamics affecting winter wheat growth processes and yield by evaluating RS data alongside meteorological variables such as temperature and precipitation. The findings obtained are also supported by other studies in literature. In their study, Manjunath et al. (2002) used the MLR method and looked at the results from the meteorological model (R^2=0.22-0.91), the spectral model (R^2=0.80), and the spectral-meteorological model (R^2=0.21-0.98), highlighting that using both VI and meteorological variables together improves prediction accuracy. In Segarra et al. (2020), they examined the impact of RS and meteorological data on winter wheat yield prediction using the MLR and RF methods. In their study, they stated that the MLR model (R^2 = 0.83) explained 11%–20% of the regional yield differences due to meteorological variables and that these variables played a critical role in data prediction. Peng et al. (2024), using the Recurrent Neural Network (RNN) method in studies predicting winter wheat yield, achieved R^2 = 0.503 with only RS data, R^2 = 0.560 with only meteorological data, and R^2 = 0.662 when combining RS and meteorological data. In their studies, they emphasized that meteorological data should be used in winter wheat yield prediction and that using it together with RS data increases the prediction accuracy by 0.16%.

This study has also shown that the measurements taken from the red edge bands of Sentinel-2A/B are dependable and useful for tracking winter wheat growth stages and predicting yields. These findings are consistent with the results of the studies by Xing et al. (2020) and Xiao et al. (2024), supporting the critical role of indices obtained from red edge bands. Xing et al. (2020) stated that the Triangular Vegetation Index (TTVI) they produced from red edge bands showed satisfactory performance for winter wheat yield prediction (R^2=0.62). In the same way, Xiao et al. (2024) created different models to predict winter wheat yield and highlighted that red edge bands performed particularly well in the ACNN (Attention Mechanism Convolutional Neural Network) model, achieving a high accuracy rate of R^2 = 0.87. This study has shown that the temporal changes of winter wheat can be observed using RS and meteorological data, and contrary to Raza et al. (2025), reliable predictions can be made using DT-based ML algorithms. Raza et al. (2025) analyzed the winter wheat phenology process using VI and meteorological variables to determine the most suitable time for winter wheat yield prediction and to evaluate the effectiveness of meteorological data. In the aforementioned study, the RF, BT (Bagging Tree), GB (Gradient Boosting), and DT models were compared. Raza et al. (2025) emphasized that the heading-blooming period is the most suitable time for yield prediction and highlighted the necessity of meteorological variables in their study. They noted that the RF model stood out with the highest accuracy value of R^2 = 0.82 and stated that the RDT models, exhibiting the lowest R^2 values, were the least successful in explaining yield variance. Otherwise, Thayanandeswari et al. (2024) compared the effectiveness of SVM, DT, and RF methods in yield prediction and noted that the RDT-RF algorithm demonstrated the best performance and more accurate results with a 98% accuracy rate. In this study, the strong correlation results between the actual and predicted yield values (R^2 = 0.92, R^2 = 0.85) have proven the effectiveness and reliability of RDT models in predicting winter wheat yield, as noted by Thayanandeswari et al. (2024). In this context, the study has demonstrated that the combined use of RS data and ML technologies has high potential in spatial yield prediction. Various studies in literature support the findings of this research. For example, Li et al. (2024) and Khechba et al. (2025) have examined environmental factors to increase winter wheat wheat yield by using RS and meteorological data with ML methods. Li et al. (2024) stated that the integration of VI, meteorological data, and phenological information resulted in an 8% improvement in winter wheat yield prediction. Khechba et al. (2025) highlight that using weather data with XGBoost and RF models has made the models perform better (R^2 = 0.81, R^2 = 0.60) and greatly improved how well MLR methods can predict outcomes.

Even if the study results are consistent, there are some limitations that need to be addressed in future studies. One of these limitations is that more precise modeling of climatic factors such as heat, frost, and water stress could enhance prediction accuracy. However, the analysis of soil characteristics and the inclusion of factors such as fertilizer applications in the model can provide significant contributions to yield prediction models. Field studies that look at things like changes in data, how data is prepared, people's economic choices, and real-life conditions that differ from the model should be reviewed to check how accurate and applicable the model is based on the area's geography and climate. In conclusion, the integration of meteorological data with RS into DT-based ML algorithms provides a more holistic understanding of the process for predicting winter wheat yield and offers strong foresight for agricultural decision-making processes. In future studies, it will be beneficial to integrate environmental variables such as climate data and soil properties, test the model with longer-term data, and generalize the model to areas with different scales and conditions. Additionally, the model can be further improved with advanced deep learning techniques such as more sophisticated ML methods, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM).

Considering these situations in future studies will increase both the scientific validity and practical applicability of the model. Thus, the accuracy of yield forecasts increases and contributes to the more efficient use of resources in the agricultural sector. In this context, it is believed that the outputs of the study are strongly aligned with SDG 1, SDG 2, and SDG 13 goals. The model, created to meet SDG 1 and SDG 2 goals, offers the possibility of more accurately predicting winter wheat yields using digital agriculture and making agricultural production more predictable. This situation is important for ensuring that decision-makers make the right choices, facilitating planned agricultural production, and increasing the global potential of winter wheat (Fu et al., 2025). Thus, it contributes to poverty reduction by supporting food supply security (Raza et al., 2025). Additionally, it promotes the efficient use of agricultural land and provides a significant tool in combating hunger by enabling the development of strategies aligned with global crises (Liu et al., 2021). In line with SDG 13, the consequences of climate change are highlighted, and producers are helped to better adapt to changing climate conditions. This study shows that its predictive outputs are directly related to policymaking and farm-level decision-making processes. The projected yield increases, and resource use efficiency offer concrete application potential for agricultural management and strategic planning.

5. Conclusion

The accurate and timely prediction of winter wheat yield is critical for the formulation of agricultural policies, ensuring food security, and supporting sustainable development goals. Therefore, this study has successfully predicted winter wheat yield using RDT-based ML models by evaluating RS and meteorological data together. In determining the phenology of winter wheat, a significant relationship has been identified between VI, meteorological data, and yield values. Especially, the heading-blooming period has emerged as the stage with the highest correlation between index values and meteorological variables with yield. Therefore, the heading-blooming period has been determined as the most suitable time for yield estimation. In predicting winter wheat yield, the best factors from weather data and RS data were found using a RDT model, which had a correlation of R = 0.89 and R^2 = 0.80. These factors are the NDVI-WDRVI-RDVI-RE1-April Values, DI-3-December Values, LAI-March Values, Average Temperature (°C)-December, and Average Total Precipitation (mm)-December meteorological variables. Particularly, we have determined that the index values measured during the winter wheat phenological stages have strong predictive ability. At the same time, it has been observed that the temperature and precipitation variables recorded during the winter months play a critical role in predicting low yields. The estimated yield values of winter wheat for future periods have been calculated at R= 0.92 and R^2 =0.86 values.

This situation demonstrates that the developed model is highly reliable and can be effectively used in yield predictions. These precise methods provide a decision support mechanism for optimal yield in winter wheat farming and effective management of agricultural lands. Additionally, this study provides an important foundation for future scientific research addressing similar topics. In this context, the developed model lays the groundwork for the development of policies aimed at increasing agricultural production efficiency and conserving natural resources by directly contributing to sustainable goals, such as SDG 1, SDG 2, and SDG 13.

Acknowledgments:

This study was produced from a section of the first author, Neslişah Civelek's, master's thesis. Neslişah Civelek's thesis advisor is Levent Genç, and Özgün Akçay is among the members of the thesis jury.

QuillBot version 36.11.4 was used only for English language editing. All data interpretation and results are entirely the responsibility of the authors.

Author Contributions:
N. Civelek: Data curation, Software, Writing—original Draft, Writing—review & editing

L. Genç: Conceptualization, Data curation, Software, Writing—original Draft, Writing—review & editing.

Ö. Akçay: Writing—Review & Editing.

Funding: This study received no external funding.

Research and publication ethics statement: In the study, the authors declare that there is no violation of research and publication ethics and that the study does not require ethics committee approval.

Conflicts of Interest: The authors declare that there is no conflict of interest related to this study.

Data Availability: The data presented in this study are available on request from the corresponding author.

References

Adeniyi, O. D., Szabo, A., Tamás, J., & Nagy, A. (2020). Winter wheat yield forecasting based on Landsat NDVI and SAVI time series. https://doi.org/10.20944/preprints202007.0065.v1

Afrasiabian, Y., Noory, H., Mokhtari, A., Nikoo, M. R., Pourshakouri, F., & Haghighatmehr, P. (2021). Effects of spatial, temporal, and spectral resolutions on the estimation of winter wheat and barley leaf area index using multi- and hyper-spectral data (Case study: Karaj, Iran). *Precision Agriculture*, 22(3), 660–688. https://doi.org/10.1007/s11119-020-09749-9

AIMS. (2024). Agricultural market information system. Retrieved May 1, 2025, from https://www.amis-outlook.org/home

Akbulut, Ö. (2022). Current approaches to reporting statistical significance in scientific research: Errors and truths. *International Journal of Eastern Mediterranean Agricultural Research*, 5(1), 1-19.

Amin, E., Belda, S., Pipia, L., Szantoi, Z., El Baroudy, A., Moreno, J., & Verrelst, J. (2022). Multi-season phenology mapping of Nile Delta croplands using time series of Sentinel-2 and Landsat 8 Green LAI. *Remote Sensing*, 14(8). https://doi.org/10.3390/rs14081812

Araus, J. L., & Cairns, J. E. (2014). Field high-throughput phenotyping: The new crop breeding frontier. *Trends in plant science*, 19(1), 52-61. https://doi.org/10.1016/j.tplants.2013.09.008

Arora, N. (2025). Role of AI and IoT in agricultural applications using google earth engine. Elsevier. https://doi.org/10.1016/B978-0-443-27372-8.00026-X

Ashfaq, M., Khan, I., Alzahrani, A., Tariq, M. U., Khan, H., & Ghani, A. (2024). Accurate winter wheat yield prediction using machine learning and climate-NDVI data fusion. *IEEE Access*, 12, 40947–40961. https://doi.org/10.1109/ACCESS.2024.3376735

Ayub, M., Khan, N. A., & Haider, R. Z. (2022). Winter wheat crop field and yield prediction using Remote Sensing and Machine Learning. Proceeding Book of 2nd IEEE International Conference on Artificial Intelligence, ICAI, Islamabad, Pakistan, 158–164.

Baldos, U. L. C., & Hertel, T. W. (2014). Global food security in 2050: The role of agricultural productivity and climate change. *Australian Journal of Agricultural and Resource Economics,* 58(4), 554–570. https://doi.org/10.1111/1467-8489.12048

Bojanowski, J. S., Sikora, S., Musiał, J. P., Woźniak, E., Dąbrowska-Zielińska, K., Slesiński, P., Milewski, T., & Łączyński, A. (2022). Integration of Sentinel-3 and MODIS vegetation indices with ERA-5 agro-meteorological indicators for operational crop yield forecasting. *Remote Sensing*, 14(5). https://doi.org/10.3390/rs14051238

Bora, S., Majumder, A., Pal, R. K., & Kingra, P. K. (2024). Use of ERA5-L reanalysis datasets to derive heat units and predict the maturity period of winter wheat crop in central Punjab. *Journal of Agrometeorology*, 26(4), 419–424. https://doi.org/10.54386/jam.v26i4.2623

Bouras, E. houssaine, Olsson, P. O., Thapa, S., Díaz, J. M., Albertsson, J., & Eklundh, L. (2023). Winter wheat yield estimation at high spatial resolution through the assimilation of Sentinel-2 data into a crop growth model. *Remote Sensing*, 15(18). https://doi.org/10.3390/rs15184425

Brenton, P., Chemutai, V., & Pangestu, M. (2022). Trade and food security in a climate change-impacted world. *Agricultural Economics*, 53(4), 580–591. https://doi.org/10.1111/agec.12727

Can, U. (2024). Solving the missing data problem. YBS Encyclopedia, 12(1).

Cao, J., Zhang, Z., Luo, Y., Zhang, L., Zhang, J., Li, Z., & Tao, F. (2021). Winter wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine. *European Journal of Agronomy*, 123. https://doi.org/10.1016/j.eja.2020.126204

Chai, T., & Draxler, R. R. (2014). Root means square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chen, Y., & Tao, F. (2022). Potential of remote sensing data-crop model assimilation and seasonal weather forecasts for early-season crop yield forecasting over a large area. *Field Crops Research*, 276, 108398. https://doi.org/10.1016/j.fcr.2021.108398

Cheng, E., Zhang, B., Peng, D., Zhong, L., Yu, L., Liu, Y., Xiao, C., Li, C., Li, X., Chen, Y., Ye, H., Wang, H., Yu, R., Hu, J., & Yang, S. (2022). Winter wheat yield estimation using remote sensing data based on machine learning approaches. *Frontiers in Plant Science*, 13. https://doi.org/10.3389/fpls.2022.1090970

Choudhary, K., Shi, W., Dong, Y., & Paringer, R. (2022). Random Forest for rice yield mapping and prediction using Sentinel-2 data with google earth engine. *Advances in Space Research*, 70(8), 2443-2457. https://doi.org/10.1016/j.asr.2022.06.073

Chowdhury, A., Ghosh, S., & Holmatov, B. (2023). Earth observation-based rice mapping studies in Vietnamese Mekong delta compared to global context: a bibliometric analysis. *Sustainability*, 16(1), 189. https://doi.org/10.3390/su16010189

Civelek, N., Inalpulat, M., & Genç, L. (2023). Evaluation of earthquake impacts on land use and land cover (LULC) using Google Earth Engine (GEE), Sentinel-2 imageries, and machine learning: Case study of Antakya. *Journal of Anatolian Environmental and Animal Sciences*. 8(4), 642-650. https://doi.org/10.35229/jaes.1349826

Da Silva, V. S., Salami, G., da Silva, M. I. O., Silva, E. A., Monteiro Junior, J. J., & Alba, E. (2020). Methodological evaluation of vegetation indexes in land use and land cover (LULC) classification. *Geology, Ecology, and Landscapes*, 4(2), 159–169. https://doi.org/10.1080/24749508.2019.1608409

De Almeida, G. S., Rizzo, R., Amorim, M. T. A., Dos Santos, N. V., Rosas, J. T. F., Campos, L. R., ... & Demattê, J. A. (2023). Monitoring soil–plant interactions and maize yield by satellite vegetation indexes, soil electrical conductivity and management zones. *Precision Agriculture*, 24(4), 1380-1400. https://doi.org/10.1007/s11119-023-09994-8

Delegido, J., Verrelst, J., Meza, C.M., Rivera, J. P., Alonso, L., & Moreno, J. (2013). A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy*, 46, 42-52. https://doi.org/10.1016/j.eja.2012.12.001

Devkota, K. P., Bouasria, A., Devkota, M., & Nangia, V. (2024). Predicting winter wheat yield gap and its determinants combining remote sensing, machine learning, and survey approaches in rainfed Mediterranean regions of Morocco. *European Journal of Agronomy*, 158, 127195. https://doi.org/10.1016/j.eja.2024.127195

Dobermann, A., & Ping, J. L. (2004). Geostatistical integration of yield monitor data and remote sensing improves yield maps. *Agronomy journal*, 96(1), 285-297. https://doi.org/10.2134/agronj2004.2850

Du, X., Zhu, J., Xu, J., Li, Q., Tao, Z., Zhang, Y., Wang, H., & Hu, H. (2025). Remote sensing-based winter winter wheat yield estimation integrating machine learning and crop growth multi-scenario simulations. *International Journal of Digital Earth*, 18(1). https://doi.org/10.1080/17538947.2024.2443470

EEDC. (2025). Era5-Land: ECMWF climate reanalysis. Google Earth Engine. Retrieved May 1, 2025, from https://developers.google.com/earthengine/datasets/catalog/ECMWF_ERA5_LAND_DAILY_AGGR#citations

FAO. (2024). Food and agriculture organization of the United Nations-Türkiye. Retrieved May 1, 2025, from https://www.fao.org/turkiye/tr/

Faqe Ibrahim, G. R., Rasul, A., & Abdullah, H. (2023). Sentinel-2 accurately estimated winter wheat yield in a semi-arid region compared with Landsat 8. *International Journal of Remote Sensing*, 44(13), 4115–4136. https://doi.org/10.1080/01431161.2023.2232542

Fu, H., Lu, J., Li, J., Zou, W., Tang, X., Ning, X., Sun, Y. (2025). Winter wheat yield predictions using satellite remote sensing data and deep learning models. *Agronomy*, 15, 205. https://doi.org/10.3390/agronomy15010205

GEOGLAM. (2024). Group on earth observations global agricultural monitoring. Retrieved May 1, 2025, from https://earthobservations.org/geoglam.php

Gitelson, A. A. (2004). Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *Journal of plant physiology,* 161(2), 165-173. https://doi.org/10.1078/0176-1617-01176

Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3), 271-282. https://doi.org/10.1078/0176-1617-00887

Gitelson, A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3), 289-298. https://doi.org/10.1016/S0034-4257(96)00072-7

Glenn, E. P., Huete, A. R., Nagler, P. L., & Nelson, S. G. (2008). Relationship between remotely sensed vegetation indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape. *Sensors*, 8(4), 2136-2160. https://doi.org/10.3390/s8042136

Goldberg, K., Herrmann, I., Hochberg, U., & Rozenstein, O. (2021). Generating up-to-date crop maps optimized for Sentinel-2 imagery in İsrael. *Remote Sensing*, 13(17). https://doi.org/10.3390/rs13173488

Graf, L. V., Merz, Q. N., Walter, A., & Aasen, H. (2023). Insights from field phenotyping improve satellite remote sensing based in-season estimation of winter winter wheat growth and phenology. *Remote Sensing of Environment*, 299. https://doi.org/10.1016/j.rse.2023.113860

GRFC. (2024). Global report on food crises. Retrieved May 1, 2025, from https://www.fsinplatform.org/report/global-report-food-crises-2024/

Gumma, M. K., Kadiyala, M. D. M., Panjala, P., Ray, S. S., Akuraju, V. R., Dubey, S., Smith, A. P., Das, R., & Whitbread, A. M. (2022). Assimilation of Remote Sensing data into crop growth model for yield estimation: A case study from India. *Journal of the Indian Society of Remote Sensing*, 50(2), 257–270. https://doi.org/10.1007/s12524-021-01341-6

Guo, X., Zhang, P., & Yue, Y. (2024). Prediction of global winter wheat cultivation distribution under climate change and socioeconomic development. *Science of The Total Environment*, 919, 170481. https://doi.org/10.1016/j.scitotenv.2024.170481

Haque, M. A., Reza, M. N., Ali, M., Karim, M. R., Ahmed, S., Lee, K. D., Khang, Y.H., & Chung, S. O. (2024). Effects of environmental conditions on vegetation indices from multispectral images: A review. *Korean Journal of Remote Sensing*, 40(4), 319-341. https://doi.org/10.7780/kjrs.2024.40.4.1

Haseeb, M., Tahir, Z., Mahmood, S. A., & Tariq, A. (2025). Winter winter wheat yield prediction using linear and nonlinear machine learning algorithms based on climatological and remote sensing data. *Information Processing in Agriculture*. https://doi.org/10.1016/j.inpa.2025.02.004

Holzman, M. E., & Rivas, R. E. (2016). Early maize yield forecasting from remotely sensed temperature/vegetation index measurements. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(1), 507-519. https://doi.org/10.1109/JSTARS.2015.2504262

Hou, X., Zhang, J., Luo, X., Zeng, S., Lu, Y., Wei, Q., Liu, J., Feng, W., & Li, Q. (2025). Peanut yield prediction using remote sensing and machine learning approaches based on phenological characteristics. *Computers and Electronics in Agriculture*, 232, 110084. https://doi.org/10.1016/j.compag.2025.110084

Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Elsevier.

Jamali, M., Bakhshandeh, E., Yeganeh, B., & Özdoğan, M. (2024). Development of machine learning models for estimating winter wheat biophysical variables using satellite-based vegetation indices. *Advances in Space Research*, 73(1), 498–513. https://doi.org/10.1016/j.asr.2023.10.004

Jin, N., Tao, B., Ren, W., He, L., Zhang, D., Wang, D., & Yu, Q. (2022). Assimilating remote sensing data into a crop model improves winter winter wheat yield estimation based on regional irrigation data. *Agricultural Water Management*, 266. https://doi.org/10.1016/j.agwat.2022.107583

Joshi, A., Pradhan, B., Chakraborty, S., & Behera, M. D. (2023). Winter winter wheat yield prediction in the conterminous United States using solar-induced chlorophyll fluorescence data and XGBoost and random forest algorithm. *Ecological Informatics*, 77, 102194. https://doi.org/10.1016/j.ecoinf.2023.102194

Karongo, J., Mwaniki, J. I., Ndiritu, J., & Mokaya, V. (2025). Sorghum yield prediction based on remote sensing and machine learning in conflict affected South Sudan. *Scientific Reports*, 15(1), 4469. https://doi.org/10.1038/s41598-025-89030-z

Khechba, K., Belgiu, M., Laamrani, A., Stein, A., Amazirh, A., & Chehbouni, A. (2025). The impact of spatiotemporal variability of environmental conditions on winter wheat yield forecasting using remote sensing data and machine learning. *International Journal of Applied Earth Observation and Geoinformation*, 136. https://doi.org/10.1016/j.jag.2025.104367

Khurshid, A., Yaseen, I., Mushtaq, F., & Ahmed, L. (2024). A potential role of geoinformatics for food security and climate change: An introduction. In Geoinformatics, 37-58. Apple Academic Press.

Kingra, P. K., Majumder, D., & Singh, S. P. (2016). Application of remote sensing and Gis in agriculture and natural resource management under changing climatic conditions. *Agricultural Research Journal*, 53(3). https://doi.org/10.5958/2395-146X.2016.00058.2

Kogan, F., & Kogan, F. (2019). Vegetation Health-Based Modeling Crop Yield and Food Security Prediction. Remote sensing for food security, 115-162. Springer.

Kul, S. (2014). Interpretation of statistical results: what is p value and confidence interval? *Plevra Bulteni*, 8(1), 11–13. https://doi.org/10.5152/pb.2014.003

Kumar, M. (2016). Impact of climate change on crop yield and role of model for achieving food security. *Environmental Monitoring and Assessment,* 188(8). https://doi.org/10.1007/s10661-016-5472-3

Lee, W. S., Alchanatis, V., Yang, C., Hirafuji, M., Moshou, D., & Li, C. (2010). Sensing technologies for precision specialty crop production. *Computers and electronics in agriculture*, 74(1), 2-33. https://doi.org/10.1016/j.compag.2010.08.005

Li, S., Huang, J., Xiao, G., Huang, H., Sun, Z., & Li, X. (2024). Improved winter winter wheat yield estimation by combining remote sensing data, machine learning, and phenological metrics. *Remote Sensing, 16*(17), 3217. https://doi.org/10.3390/rs16173217

Liu, L., Xiao, X., Qin, Y., Wang, J., Xu, X., Hu, Y., & Qiao, Z. (2020). Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine. *Remote Sensing of Environment*, 239, 111624. https://doi.org/10.1016/j.rse.2019.111624

Liu, S., Peng, D., Zhang, B., Chen, Z., Yu, L., Chen, J., Pan, Y., Zheng, S., Hu, J., Lou, Z., Chen, Y., & Yang, S. (2022). The accuracy of winter winter wheat identification at different growth stages using remote sensing. *Remote Sensing*, 14(4). https://doi.org/10.3390/rs14040893

Liu, W., Chen, Y., He, X., Mao, Ping., Tian, H. (2021). Is current research on how climate change impacts global food security really objective?. *Foods*, 10, 2342. https://doi.org/10.3390/ foods10102342

Lu, J., Li, J., Fu, H., Tang, X., Liu, Z., Chen, H., Sun, Y., & Ning, X. (2024). Deep learning for multi-source data-driven crop yield prediction in northeast China. *Agriculture*, 14(6), 794. https://doi.org/10.3390/agriculture14060794

Ma, C., Liu, M., Ding, F., Li, C., Cui, Y., Chen, W., & Wang, Y. (2022). Winter wheat growth monitoring and yield estimation based on remote sensing data assimilation into the SAFY crop growth model. *Scientific Reports*, 12(1). https://doi.org/10.1038/s41598-022-09535-9

Manjunath, K. R., Potdar, M. B., & Purohit, N. L. (2002). Large area operational winter wheat yield model development and validation based on spectral and meteorological data. *International Journal of Remote Sensing*, 23(15), 3023–3038. https://doi.org/10.1080/01431160110104692

Mashonganyika, F., Mugiyo, H., Svotwa, E., & Kutywayo, D. (2021). Mapping of winter winter wheat using Sentinel-2 NDVI data: A case of Mashonaland central provincse in Zimbabwe. *Frontiers in Climate*, 3. https://doi.org/10.3389/fclim.2021.715837

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. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Myneni, R. B., & Williams, D. L. (1994). On the relationship between FAPAR and NDVI. *Remote Sensing of Environment*, 49(3), 200-211. https://doi.org/10.1016/0034-4257(94)90016-7

Naqvi, S. M. Z. A., Tahir, M. N., Shah, G. A., Sattar, R. S., & Awais, M. (2019). Remote estimation of winter wheat yield based on vegetation indices derived from time series data of Landsat 8 imagery. *Applied Ecology and Environmental Research*, 17(2), 3909–3925. https://doi.org/10.15666/aeer/1702_39093925

Newete, S. W., Abutaleb, K., Chirima, G. J., Dabrowska-Zielinska, K., & Gurdak, R. (2024). Phenology-based winter winter wheat classification for crop growth monitoring using multi-temporal Sentinel-2 satellite data. *Egyptian Journal of Remote Sensing and Space Science*, 27(4), 695–704. https://doi.org/10.1016/j.ejrs.2024.10.001

Ogundeji, A. A. (2022). Adaptation to climate change and impact on smallholder farmers' food security in south Africa. *Agriculture*, 12(5). https://doi.org/10.3390/agriculture12050589

Oppelt, N., & Mauser, W. (2004). Hyperspectral monitoring of physiological parameters of winter wheat during a vegetation period using AVIS data. *International Journal of Remote Sensing*, 25(1), 145-159. https://doi.org/10.1080/0143116031000115300

Pan, L., Xia, H., Yang, J., Niu, W., Wang, R., Song, H., ... & Qin, Y. (2021). Mapping cropping intensity in Huaihe basin using phenology algorithm, all Sentinel-2 and Landsat images in Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation,* 102, 102376. https://doi.org/10.1016/j.jag.2021.102376

Pandey, P. C., & Pandey, M. (2023). Highlighting the role of agriculture and geospatial technology in food security and sustainable development goals. *Sustainable Development*, 31(5), 3175-3195. https://doi.org/10.1002/sd.2600

Panumonwatee, G., Choosumrong, S., Pampasit, S., Premprasit, R., Nemoto, T., & Raghavan, V. (2025). Machine learning technique for carbon sequestration estimation of mango orchards area using Sentinel-2 Data. *Carbon Research*, 4(1), 33. https://doi.org/10.1007/s44246-025-00201-z

Peng, D., Cheng, E., Feng, X., Hu, J., Lou, Z., Zhang, H., Zhao, B., Lv, Y., Peng, H., & Zhang, B. (2024). A Deep–learning network for winter wheat yield prediction combining weather forecasts and Remote Sensing data. *Remote Sensing*, 16(19). https://doi.org/10.3390/rs16193613

Prasad, A. K., Singh, R. P., Tare, V., & Kafatos, M. (2007). Use of vegetation index and meteorological parameters for the prediction of crop yield in India. *International Journal of Remote Sensing*, 28(23), 5207-5235. https://doi.org/10.1080/01431160601105843

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119-126. https://doi.org/10.1016/0034-4257(94)90134-1

Quintano, C., Fernández-Manso, A., Fernández-Guisuraga, J. M., & Roberts, D. A. (2024). Improving fire severity analysis in mediterranean environments: A comparative study of eeMETRIC and SSEBop Landsat-based evapotranspiration models. *Remote Sensing*, 16(2). https://doi.org/10.3390/rs16020361

Raza, A., Shahid, M. A., Zaman, M., Miao, Y., Huang, Y., Safdar, M., Maqbool S., & Muhammad, N. E. (2025). Improving winter wheat yield prediction with multi-source remote sensing data and machine learning in arid regions. *Remote Sensing*, 17(5), 774. https://doi.org/10.3390/rs17050774

Roberts, D. A., Roth, K. L., Wetherley, E. B., Meerdink, S. K., & Perroy, R. L. (2018). Hyperspectral vegetation indices. In Hyperspectral indices and image classifications for agriculture and vegetation. CRC press.

Roujean, J. L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sensing of Environment, 51(3), 375-384. https://doi.org/10.1016/0034-4257(94)00114-3

Rouse, J.W., Haas, R.H., Schell, J.A., D.W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. Proceedings Book of Third Earth Resources Technology Satellite Symposium, Washington, D. C., USA, 309-317.

Saad El Imanni, H., El Harti, A., & El Iysaouy, L. (2022). Winter wheat yield estimation using remote sensing indices derived from Sentinel-2 time series and Google Earth Engine in a highly fragmented and heterogeneous agricultural region. *Agronomy*, 12(11). https://doi.org/10.3390/agronomy12112853

Saric, R., Nguyen, V. D., Burge, T., Berkowitz, O., Tritle, M., Whelan, J., ... & Custovic, E. (2022). Applications of hyperspectral imaging in plant phenotyping. *Trends in plant science*, 27(3), 301-315. https://doi.org/10.1016/j.tplants.2021.12.003

Scown, M. W., & Nicholas, K. A. (2020). European agricultural policy requires a stronger performance framework to achieve the Sustainable Development Goals. *Global Sustainability*, 3. https://doi.org/10.1017/sus.2020.5

Segarra, J., González-Torralba, J., Aranjuelo, Í., Araus, J. L., & Kefauver, S. C. (2020). Estimating winter wheat grain yield using Sentinel-2 imagery and exploring topographic features and rainfall effects on winter wheat performance in Navarre, Spain. *Remote Sensing*, 12(14). https://doi.org/10.3390/rs12142278

Shams Esfandabadi, H., Ghamary Asl, M., Shams Esfandabadi, Z., Gautam, S., & Ranjbari, M. (2022). Drought assessment in paddy rice fields using remote sensing technology towards achieving food security and SDG2. *British Food Journal*, 124(12), 4219–4233. https://doi.org/10.1108/BFJ-08-2021-0872

Sharma, R., Ghosh, A., & Joshi, P. K. (2013). Decision tree approach for classification of remotely sensed satellite data using open source support. *Journal of Earth System Science*, 122(5), 1237-1247. https://doi.org/10.1007/s12040-013-0339-2

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google Earth Engine platform for big data processing: Classification of multi-temporal satellite imagery for crop mapping. *Frontiers in Earth Science*, 5(17). https://doi.org/10.3389/feart.2017.00017

Shmelev, S. E., Salnikov, V., Turulina, G., Polyakova, S., Tazhibayeva, T., Schnitzler, T., & Shmeleva, I. A. (2021). Climate change and food security: the impact of some key variables on winter wheat yield in Kazakhstan. *Sustainability*, 13(15), 8583. https://doi.org/10.3390/su13158583

Sibanda, M., Mutanga, O., Dube, T., Odindi, J., & Mafongoya, P. L. (2019). The utility of the upcoming HyspIRI's simulated spectral settings in detecting maize gray leafy spot in relation to Sentinel-2 MSI, VENμS, and Landsat 8 OLI sensors. *Agronomy*, 9(12), 846. https://doi.org/10.3390/agronomy9120846

Singh, A., Mehrotra, R., Rajput, V. D., Dmitriev, P., Singh, A. K., Kumar, P., ... & Singh, A. K. (2022). Geoinformatics, artificial intelligence, sensor technology, big data: emerging modern tools for sustainable agriculture. *Sustainable Agriculture Systems and Technologie*s, 295-313. https://doi.org/10.1002/9781119808565.ch14

Singha, C., & Swain, K. C. (2023). Rice crop growth monitoring with sentinel 1 SAR data using machine learning models in google earth engine cloud. *Remote Sensing Applications: Society and Environment*, 32, 101029. https://doi.org/10.1016/j.rsase.2023.101029

Skakun, S., Vermote, E., Roger, J.-C., & Franch, B. (2017). Combined use of Landsat-8 and Sentinel-2A images for winter crop mapping and winter winter wheat yield assessment at regional scale. *AIMS Geosciences*, 3(2), 163–186. https://doi.org/10.3934/geosci.2017.2.163

Sykas, D., Sdraka, M., Zografakis, D., & Papoutsis, I. (2022). A Sentinel-2 multiyear, multicountry benchmark dataset for crop classification and segmentation with deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3323–3339. https://doi.org/10.1109/JSTARS.2022.3164771

Thayanandeswari, C. S. S., Jaya, T., Ahamed, S. N., Bommy, B., Kumar, G. B., & Shyjith, M. B. (2024). A machine learning approach for crop yield prediction using weather condition. Proceedings of the 2024 International Conference on Advancement in Renewable Energy and Intelligent Systems, AREIS 2024. https://doi.org/10.1109/AREIS62559.2024.10893606

Thenkabail, P. S., Smith, R. B., & De Pauw, E. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote sensing of Environment*, 71(2), 158-182. https://doi.org/10.1016/S0034-4257(99)00067-X

Thieme, A., Yadav, S., Oddo, P. C., Fitz, J. M., McCartney, S., King, L. A., Keppler, J., McCarty, G. W., & Hively, W. D. (2020). Using NASA Earth observations and Google Earth Engine to map winter cover crop conservation performance in the Chesapeake Bay watershed. *Remote Sensing of Environment*, 248. https://doi.org/10.1016/j.rse.2020.111943

Tian, H., Meng, M., Wu, M., & Niu, Z. (2019). Mapping spring canola and spring winter wheat using Radarsat-2 and Landsat-8 images with Google Earth Engine. *Current Science*, 116(2), 291-298.

Toscano, P., Castrignanò, A., Di Gennaro, S. F., Vonella, A. V., Ventrella, D., & Matese, A. (2019). A precision agriculture approach for durum winter wheat yield assessment using remote sensing data and yield mapping. *Agronomy*, 9(8). https://doi.org/10.3390/agronomy9080437

UNESCO. (2024). Sustainable development goals 2030 specialized commite. Retrieved May 1, 2025, from https://www.unesco.org.tr/Pages/108/219/S%C3%BCrd%C3%BCr%C3%BClebilir-Kalk%C4%B1nma-2030-Hedefleri-%C4%B0htisas-Komitesi

Vidican, R., Malinas, A., Ranta, O., Moldovan, C., Marian, O., Gheţe, A., ... & Cătunescu, G. M. (2023). Using remote sensing vegetation indices for the discrimination and monitoring of agricultural crops: a critical review. *Agronomy*, 13(12), 3040. https://doi.org/10.3390/agronomy13123040

Wang, L., Liu, J., Yang, F., Fu, C., Teng, F., & Gao, J. (2015). Early recognition of winter winter wheat area based on GF-1 satellite. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering,* 31(11), 194–201. https://doi.org/10.11975/j.issn.1002-6819.2015.11.028

Waters, R., Allen, R., Bastiaanssen, W., Tasumi, M., & Trezza, R. (2002). S E B A L Surface energy balance algorithms for land idaho implementation advanced training and user's manual list of symbols. EOSDIS, 1.

Weiss, M., Jacob, F., & Duveiller, G. (2020). Remote sensing for agricultural applications: A meta-review. *Remote sensing of environment*, 236, 111402. https://doi.org/10.1016/j.rse.2019.111402

Xiao, G., Zhang, X., Niu, Q., Li, X., Li, X., Zhong, L., & Huang, J. (2024). Winter winter wheat yield estimation at the field scale using Sentinel-2 data and deep learning. *Computers and Electronics in Agriculture*, 216. https://doi.org/10.1016/j.compag.2023.108555

Xing, N., Huang, W., Xie, Q., Shi, Y., Ye, H., Dong, Y., Wu, M., Sun, G., & Jiao, Q. (2020). A transformed triangular vegetation index for estimating winter winter wheat leaf area index. *Remote Sensing,* 12(1). https://doi.org/10.3390/RS12010016

Xu, J., Song, Y., Rui, Z., Zhang, Z., Hu, C., Wang, L., ... & Wang, X. (2025). Trend analysis of the application of multispectral technology in plant yield prediction: a bibliometric visualization analysis (2003–2024). *Frontiers in Sustainable Food Systems*, 9, 1513690. https://doi.org/10.3389/fsufs.2025.1513690

Zhang, X. H. (2011). Identification of Mangrove using decision tree method. Proceeding Book of 4th International Conference on Information and Computing, ICIC 2011, 130–132. https://doi.org/10.1109/ICIC.2011.70

Zhang, Z., Jin, W., Dou, R., Cai, Z., Wei, H., Wu, T., Yang, S., Tan, M., Li, Z., Wang, C., Yin, G., & Xu, B. (2023). Improved estimation of leaf area index by reducing leaf chlorophyll content and saturation effects based on red-edge bands. *IEEE Transactions on Geoscience and Remote Sensing,* 61. https://doi.org/10.1109/TGRS.2023.3270712

Zhao, Y., Han, S., Zheng, J., Xue, H., Li, Z., Meng, Y., Li, X., Yang, X., Li, Z., Cai, S., & Yang, G. (2023). China Winter wheat Yield 30m: a 30m annual winter winter wheat yield dataset from 2016 to 2021 in China. *Earth System Science Data*, 15(9), 4047–4063. https://doi.org/10.5194/essd-15-4047-2023

Zhen, Z., Yunsheng, L., Moses, O. A., Rui, L., Li, M., & Jun, L. (2020). Hyperspectral vegetation indexes to monitor winter wheat plant height under different sowing conditions. *Spectroscopy Letters*, 53(3), 194–206. https://doi.org/10.1080/00387010.2020.1726401

Zheng, Z., Schmid, B., Zeng, Y., Schuman, M. C., Zhao, D., Schaepman, M. E., & Morsdorf, F. (2023). Remotely sensed functional diversity and its association with productivity in a subtropical forest. *Remote Sensing of Environment*, 290, 113530. https://doi.org/10.1016/j.rse.2023.113530

Zhou, X. Y., Lu, G., Xu, Z., Yan, X., Khu, S. T., Yang, J., & Zhao, J. (2023). Influence of Russia-Ukraine war on the global energy and food Security. *Resources, Conservation and Recycling*, 188. https://doi.org/10.1016/j.resconrec.2022.106657

Zhu, X., Guo, R., Liu, T., & Xu, K. (2021). Crop yield prediction based on agrometeorological indexes and remote sensing data. *Remote Sensing*, 13(10), 2016. https://doi.org/10.3390/rs13102016