International Journal of Environment and Geoinformatics 11(3):106-118 (2024)

Research Article

Classification of Agricultural Crops with Random Forest and Support Vector Machine Algorithms Using Sentinel-2 and Landsat-8 Images

Murat Güven Tuğaç[1]* Fatih Fehmi Şimşek[2] Harun Torunlar[1]

[1] Ministry of Agriculture and Forestry, Soil, Fertilizer and Water Resources Central Research Institute, Ankara/Türkiye.
[2]Ministry of Agriculture and Forestry, General Directorate of Agricultural Reform, Ankara/Türkiye

* Corresponding author: M. G. Tuğaç
E-mail: mgtugac@gmail.com

Received 06.05.2024
Accepted 14.09.2024

How to cite: Tuğaç et al., (2024). Classification of Agricultural Crops with Random Forest and Support Vector Machine Algorithms Using Sentinel-2 and Landsat-8 Images, International Journal of Environment and Geoinformatics (IJEGEO), 11(3): 106-118. doi. 10.30897/ijegeo. 1479116

Abstract

Monitoring crop development and mapping cultivated areas are important for reducing risks to food security due to climate change. Remote sensing techniques contribute significantly to the efficient and effective management of agricultural production. In this study, agricultural fields (sunflower, wheat, maize, oat, chickpea, sugar beet, alfalfa, onion, fallow) and other fields (non-agricultural, pasture, lake) were identified by using Random Forest (RF) and Support Vector Machines (SVM) machine learning algorithms with Sentinel-2 and Landsat-8 images in the area covering Polatlı, Haymana and Gölbaşı districts of Ankara province Multi-temporal images were used to distinguish winter and summer crops, taking into account crop development periods. As a result of classification; the overall accuracy of RF and SVM models with S2 images are 89.5% and 84.6% and kappa coefficients are 0.88 and 0.83, while the overall accuracy of RF and SVM models with L8 images are 79% and 78.1% and kappa coefficients are 0.76 and 0.75. RF model was found to have higher prediction accuracy than SVM. Sentinel-2 imagery has a higher accuracy in all classes compared to Landsat-8, indicating that Sentinel-2 imagery with its high temporal and spatial resolution is more suitable and has a great potential for agricultural crop pattern detection.

Keywords: Crop classification, Random Forest, Support Vector Machine, Landsat-8, Sentinel-2

Introduction

Water is the source of human development (Nadir and Carrivick, 2019) and the initiating place for all the civilisations/ metropolitans which were established and thrived along the water streams (Gimenez, 2015). Today, water resources are used as a source of life, biodiversity, energy generation, irrigation, transportation/ shipping, and primary food sources in the life/ food cycle (Nadir and Ahmed, 2023). The increased use of water resources, wastages, transborder streams inflow/ outflow, regionalisation/ zoning of the earth in cities/ countries, global warming and climatic changes are the main factors causing the drought and availability of a sufficient/ suitable quantity of this precious commodity (Gimenez, 2015; Kumar, 2018; Nadir and Ahmed, 2023).

Climate change and population growth require effective and rational management of agricultural land to ensure food security. Monitoring agricultural lands and identifying potential challenges in crop production in advance is of great importance (Karmakar et al., 2024). Production data for important agricultural products grown in large areas need to be up-to-date and reliable. These data should be monitored spatially and temporally for support policies. In this context, while spatial data is needed, accurate crop mapping is critical for food security (Özdoğan, 2010). In particular, at the level of decision-makers, appropriate plans and policies can be developed to create import and export projections by determining agricultural crop patterns and production quantities (See et al., 2015). In this context, determining the cultivation areas of crops for the accuracy and effectiveness of agricultural production plans and policies is a priority issue in agriculture. Remote sensing has been used effectively and widely in obtaining agricultural production data by determining crop cultivated areas at regional and national scale (Song et al., 2017; Blickensdorfer et al., 2022).

Remote sensing techniques are an effective tool for spatially monitoring the health conditions of agricultural crops throughout the crop growth period (Esetlili et al., 20181; Gorji et al., 2019; Remelgado et al., 2020). Sowing and harvesting periods of crops planted in the same region may vary between years according to seasonal dynamics. The spectral and temporal characteristics of satellite imagery have the potential to improve classification accuracy by identifying crop differences (Zhong et al., 2014; Yi et al., 2020). Crop classification is based on how reflectance features change over the crop growth period (Heupel et al., 2018). To achieve the desired level of accuracy in crop classification, various spectral bands including different temporal images of multispectral time series and vegetation indices derived from these bands are effectively used (Tuvdendorj et al., 2022).

In crop classification, data with different characteristics obtained from many satellite images are used. Among satellite imagery, The high spatial and temporal resolution of Sentinel-2 imagery provides an important opportunity for crop classification (Immitzer et al., 2016; Nasrallah et al., 2018; Gumma et al., 2020; Torunlar et al., 2021; Altun and Türker, 2021; Bantchına and Gündoğdu, 2024). In particular, Sentinel-2 data has a high potential for monitoring crop growth on a plot-by-plot basis and providing real-time spatial data (Chakhar et al., 2020; Segarra et al., 2022). However, Landsat satellite data has also been used extensively in crop classification (Zheng et al., 2015; Anua and Wong, 2022). Landsat imagery has a temporal resolution of 16 days and is likely to provide cloudy images due to this large temporal resolution. Therefore, there are limitations in capturing crop phenology, which is essential for crop discrimination (Cai et al., 2018).

Machine learning (ML) techniques have significantly improved over traditional models and their widespread use has increased. Unlike traditional models that assume a certain statistical distribution of the input data, ML techniques are nonparametric supervised methods that can adapt to the data without imposing any prior assumptions (Liu et al., 2020). ML techniques can learn complex patterns and relationships from data and make accurate predictions for new data points (Liakos et al., 2018). ML models, which are widely used in crop classification, are nonparametric supervised methods that can adapt to the data instead of making statistical assumptions like traditional models (Verma et al., 2020). However, ML models can determine the spatial distribution of crops without very large training data (Debats et al., 2016) and can better classify very different data structures (Basukala et al., 2017). Among ML models, RF and SVM algorithms are widely used in crop classification (Löw et al., 2013; Zheng et al., 2015; Inglada et al., 2015; Savitha and Talari, 2023). ML algorithms define the characteristics of the data to be classified through the learning process of the training data (Fu et al., 2023).

In crop classification, the amount of data available from the field for large areas is limited. On the other hand, simply increasing the number of samples may not be sufficient for accuracy. The quantity as well as the distribution and quality of the reference data have a significant impact on the accuracy of the classification map. (Pott et al., 2021; Alami Machichi et al., 2023). A proportional distribution of the reference data samples taking into account the class areas may lead to better results (Colditz, 2015).

In this study, ML techniques are integrated with multi-temporal Sentinel-2 and Landsat-8 imagery to evaluate the classification accuracy of agricultural crops with different crop growth periods. The accuracy and model performance of the maps obtained using RF and SVM algorithms are compared for large areas.

Materials and Methods

Study Area

The study area is located between longitudes 31^0 49' 50" and 32^0 57' 54" and latitudes 39^0 53' 41" and 38^0 57' 45", covering Polatlı, Haymana and Gölbaşı Districts within the borders of Ankara Province in Central Anatolia Region. The size of its area is 7,089 km^2 and the average elevation is 1035 m. The region has a semi-arid climate regime with dry and hot temperatures in summer and cold temperatures in winter. The average annual temperature is 12^0C and the total annual rainfall is 370 mm. Approximately 75% of the total annual precipitation occurs between November and May. Based on CORINE 2018 (CLC18) land cover classes, the current land cover distribution is mostly dry agriculture (49.9%), while irrigated agriculture (15.9%), pasture (12.8%), mixed agriculture (6.9%) and non-agricultural (14.6%) areas constitute the other classes (Figure 1).

In this region, there are two main cultivation periods. Creals are grown from October to July of the following year, while summer crops are grown from April to November. Fallow agriculture is common in the region and one crop is usually grown during the agricultural season. In the study area where dry farming is predominantly practiced, wheat and barley are the main crops and other crops produced include corn, alfalfa, Sugar beet, sunflower, onion and chickpea.

Satellite Images

Sentinel-2 and Landsat-8 images were used as Remote Sensing data. Sentinel-2 consists of two satellites launched by the European Space Agency (ESA) in 2015 (2A) and 2017 (2B) as part of the European

Union's Copernicus Environmental Monitoring Program, providing global imagery. Sentinel 2 has a temporal resolution of 10 days for a single satellite and 5 days for a dual satellite, and has 13 spectral bands with spatial resolutions of 10, 20 and 60 m (Table 1). Landsat-8, launched by the US Geological Survey (USGS) and NASA in 2013, has a temporal resolution of 16 days, a total of 11 bands and a spatial resolution ranging from 15 m to 100 m (Table 1). In the study, geographic and atmospheric calibration at Level 2 and images with less than 10% cloudiness were preferred for the images of both satellites.

Field Reference Data

Field studies were carried out in 2019 to provide model training and test data for crop classification. Based on CLC18 data, reference sampling areas were determined according to the areal density and distribution of classes. A stratified random sampling approach was used to create the reference dataset. In this context, 498 parcels of different crops were sampled. In addition to field studies, 350 data obtained from the Ministry of Agriculture and Forestry, Farmers Registration System (FRS) were used to obtain crop information for the study area to be used as training data in the modeling phase. These data were pre-processed and missing or incorrect parcel data were corrected or extracted. Winter and summer crops commonly grown in the study area and non-agricultural areas were classified. In this context, twelve classes were created. These classes are Wheat, Sunflower, Maize, Oat, Chickpea, Sugar Beet, Onion, Alfalfa, Fallow, Pasture, Lake and Non-agricultural areas. In the study area, rocky, bare areas, roads and settlement areas were classified as non-agricultural.

Image Classification

In the classification process, some user-defined parameters are optimized to achieve maximum accuracy. After training the models, a crop classification map is created according to the optimum parameter specifications. The accuracy of the models is tested on a reference dataset. RF and SVM supervised machine learning algorithms are used for crop classification. The crop classification methodology includes four main stages: data collection, data processing, classification and accuracy (Figure 2).

In image classification, 18 Sentinel-2 and 14 Landsat-8 images were used as input data between October 2018 and November 2019 (Table 2). RGB (Blue, Green, Red) and near infrared (NIR) bands and NDVI data were used in the classification study. NDVI data were obtained for the images using red and near infrared spectral bands. A multispectral time series was created by combining the RGBNIR and NDVI data obtained for Landsat-8 and Sentinel-2 images.

Table 1. Sentinel 2 and Landsat 8 spectral band characteristics

	Sentinel-2			Landsat-8		
Band	**No**	**Spectral range(µm)**	**Spatial resolution(m)**	**No**	**Spectral range(µm)**	**Spatial resolution(m)**
Blue	2	0.46-0.52	10	2	0.45-0.51	30
Green	3	0.54-0.58	10	3	0.53-0.59	30
Red	4	0.65-0.68	10	4	0.64-0.67	30
Red-edge1	5	0.70-0.71	20			
Red-edge2	6	0.73-0.75	20			
Red-edge3	7	0.76-0.78	20			
NIR	8	0.78-0.90	10	5	0.85-0.88	30
NIR	8A	0.85-0.87	20			
SWIR1	11	1.56-1.65	20	6	1.57-1.65	30
SWIR2	12	2.10-2.28	20	7	2.11-2.29	30
Panchromatic				8	0.50-0.68	15

Table 2. Sentinel 2 and Landsat 8 image acquisition dates

Sentinel-2		Landsat-8	
29.10.2018	26.06.2019	27.10.2018	11.08.2019
08.11.2018	01.07.2019	12.11.2018	27.08.2019
21.02.2019	31.07.2019	20.03.2019	12.09.2019
08.03.2019	10.08.2019	05.04.2019	28.09.2019
23.03.2019	25.08.2019	14.05.2019	14.10.2019
27.04.2019	19.09.2019	08.06.2019	
17.05.2019	29.09.2019	24.06.2019	
27.05.2019	14.10.2019	10.07.2019	
06.06.2019	29.10.2019	27.08.2019	

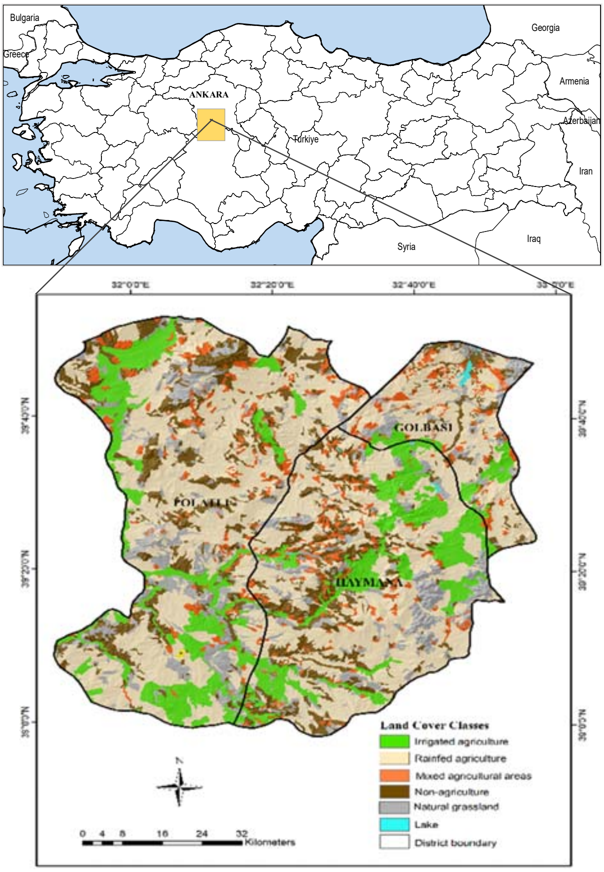

Fig 1. Study area

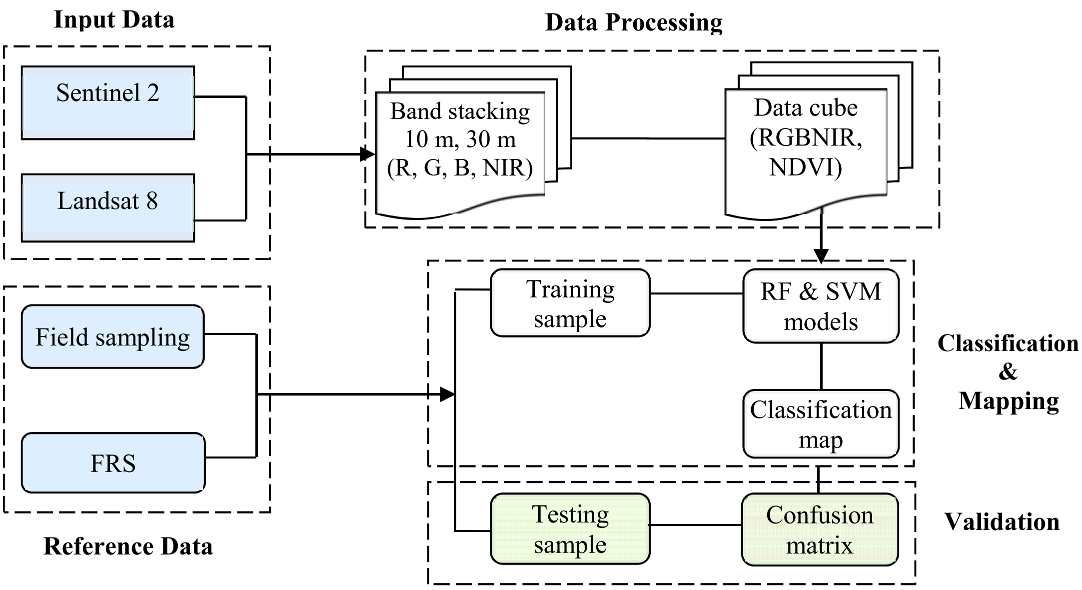

Fig 2. Image classification flowchart

The reference dataset was obtained from crop data collected from the field using a Global Positioning System (GPS) device, and samples of some parcels were also obtained from the FRS data. Parcel boundaries were determined using the FRS and Sentinel 2 data. Then, the number of sampling points was increased by randomly increasing the sampling points within the parcel boundaries. The reference dataset was divided into 75% training data and 25% test data. The datasets were trained in RF and SVM model structure. Implementation of the model algorithms and image classification were performed using the R programming language. Crop classification maps were produced in 10 m and 30 m spatial resolution and in raster data format. ArcGIS 10.5 program was used for visualization and mapping of crop classification

Classification models

The random forest (RF) model creates a forest randomly, so that a direct relationship can be established between the number of trees in the algorithm and the result. The RF classifier is more reliable for large feature size and data noise ranges, and the random process in the algorithm reduces model overfitting (Breiman, 1999). In many crop mapping studies, higher accuracies can be achieved with RF compared to other machine learning algorithms (Zhong et al., 2014; Tatsumi et al., 2015). In the application of the RF model, the number of trees (ntree) and the number of features in each split (mtry) should be determined (Thanh and Kappas, 2018; Çelik et al., 2023). Better results can be obtained by determining the appropriate model parameters (Zhang and Roy, 2017).

The support vector machine (SVM) is a nonparametric learning algorithm often used in remote sensing applications. SVM can partition datasets into different classes where the classes are linearly separable, and a linear decision boundary is determined that leaves the largest margin between the two classes. The margin is defined as the sum of the distances from the closest points of the two classes to the hyperplane (Vapnik, 1995). In the case of nonlinear classification, various types of kernels can be used to define the optimal hyperplane, transforming nonlinear boundaries into linear ones (Huang et al., 2002; Mathur and Foody, 2008). Different functions can be used in SVM algorithms; linear, polynomial, radial basis function (RBF) and sigmoid kernel (Çölkesen and Kavzoğlu, 2008). Among these, the radial basis function (RBF) kernel outperforms the others by playing an important role as the cost (C) and kernel width (γ) parameters (Duro et al., 2012).

The accuracy of classification models depends on user-defined parameters. In this study, hyperparameter optimization is applied on the models. Hyperparameter optimization refers to the tuning of hyperparameters used machine learning models to provide the best performance.

Hyperparameters are user-specified variables that control the structural properties and behavior of the model. Hyperparameter optimization improves the performance of a model by enabling it to generalize better and helping to avoid problems such as overfitting or underfitting, while minimizing computational cost. In the literature, two different methods, grid search and random search, are used for hyperparameter tuning. While the grid search method finds the best parameter value by selecting a certain parameter range and trying all possible combinations, the random search method determines the parameter values by randomly sampling from a certain parameter range (Escabias 2017). In this study, the optimum parameter values for both RF and SVM models were determined. RF model parameters ntree and mtry values are 800 and 10 for Sentinel 2 images and 600 and 8 for Landsat 8 images, respectively. On the other hand, SVM model parameters C and gamma (γ) values are 64 and 0.125 for Sentinel 2 images and 56 and 0.125 for Landsat 8 images, respectively.

Accuracy Assessment

The classification results of remote sensing data are obtained by comparing the images. Classification accuracy is determined by the level of relationship between the class assigned to a pixel and the actual class. In the accuracy assessment, validation criteria obtained from the error matrix are used (Heupel, 2018). Classification accuracy is assessed by constructing the error matrix and calculating the overall accuracy (OA), producer accuracy (PA), user accuracy (UA) and kappa coefficient (Congalton, 1991).

In the evaluation criterion, PA is the ratio of correctly classified data to the sampling data in each class, while UA is the ratio of correctly classified data in each class to the total number of data classified in that class (Tuvdendorj et al., 2022; Savitha and Talari, 2023; Vogiatzis and Eleftheriadis, 2023).

Results and Discussion

In this study, crop types were mapped using Sentinel-2 and Landsat-8 data for the study area in Central Anatolia Region with semi-arid climate and the performance of RF and SVM machine learning algorithms were compared. In determining the seasonal plant growth profiles of the crops, land and FRS data collected from reference sampling plots were evaluated. Parcel data registered in the FRS are used as local reference data in crop pattern extraction studies with image classification methods (Altun and Türker, 2021; Şimşek and Durduran, 2023). In this study, the crop accuracy in the FRS plots was determined by checking whether the NDVI time series of the plots matched the characteristic NDVI curve of the relevant crop in the study area. In this respect, plots that did not match the characteristic NDVI curve of the crops were excluded from the reference dataset. Also, it is possible that a part of a parcel or more than one crop is grown (Yaşar and Yağcı, 2023; Şimşek and Durduran, 2023; Şimşek 2024). In this case, the parcel boundaries were rearranged to obtain the correct reflectance values of the mixed parcels.

NDVI spectral reflectance values obtained from Sentinel-2 time series data of wheat, sunflower, maize, oat, chickpea, beet, dry onion and alfalfa crops produced throughout the study area and fallow areas are shown in Figure 3. The beginning of the rising spectral profile in the crop graphs indicates germination. The germination time for wheat starts from the end of October, for chickpea in late April and early May, for maize, sugar beet and sunflower in late May. The highest biomass period (peak of the season) occurred in mid-May for wheat, early July for chickpea, early August for maize, sugar beet and sunflower. As the harvest period approaches and the ripening period begins, the leaves of the plants begin to turn yellow, and NDVI values decrease from the first week of June for wheat, and from September for maize, sugar beet and sunflower.

Unlike other crops, alfalfa can be harvested three or four times until the end of September depending on the temperature values. The phenological periods of the crops may vary over the years according to climatic differences within the agricultural season. The distinctive spectral characteristics of the crops due to their phenological differences have been utilized to classify the crops. On the other hand, crops with similar developmental periods such as wheat and barley have similar reflectance values, which makes it difficult to distinguish the crops. Furthermore, clear variations in the ndvi value depending on the biomass of crops such as alfalfa, which is cut many times during the season, clarify the differentiation of crops.

In the creation of land use maps of the study area, both Sentinel-2 and Landsat-8 images were classified using RF and SVM models. A total of twelve land use classes, nine for crops and three for other uses, were created for Sentinel-2 (Figure 4) and Landsat-8 (Figure 5) images.

Confusion matrices were created to compare the performance of Sentinel-2 and Landsat-8 images using RF and SVM models for crop classification. In the validation phase; overall accuracy (OA), kappa (K), user (UA) and producer (PA) accuracy values of each class were determined to evaluate the classification accuracies. Accuracy assessments were performed for RF and SVM algorithms of Sentinel-2 and Landsat-8 images using the land reference data (Table 3). The reference data used in the accuracy assessment were obtained from field studies and FRS data for the parcels determined based on land cover/land uses. In this context, a reference dataset containing a total of 40,807-pixel data was created for Sentinel-2 and Landsat-8 images. These points were selected from the whole area representing the land cover classes according to their areal size.

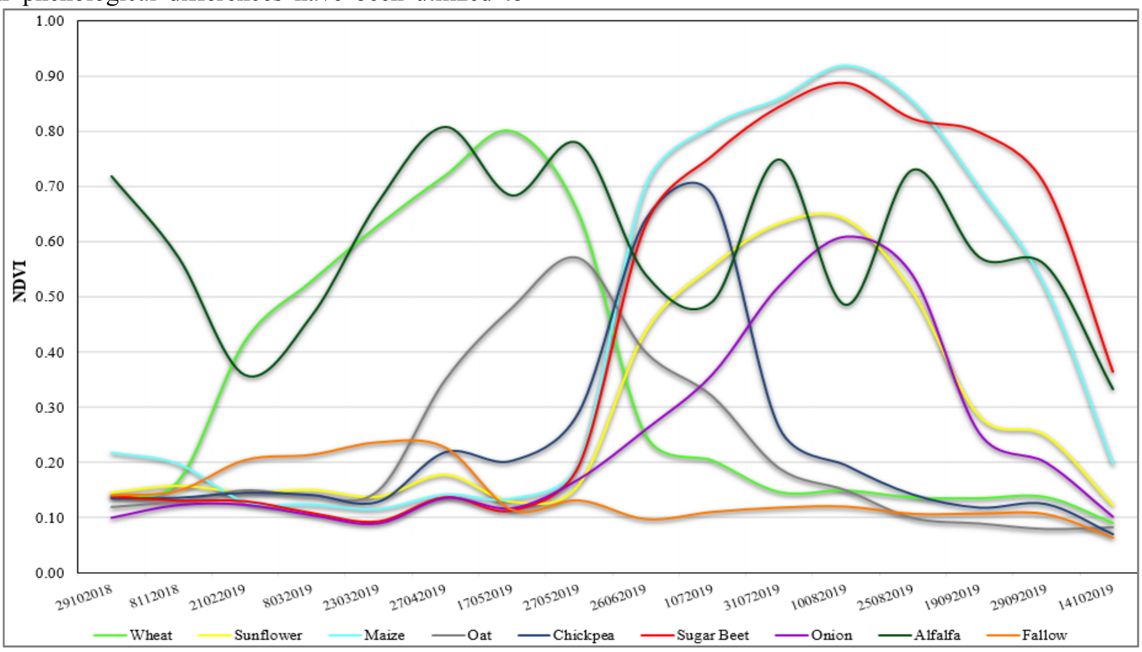

Fig 3. Temporal changes of NDVI values of crops

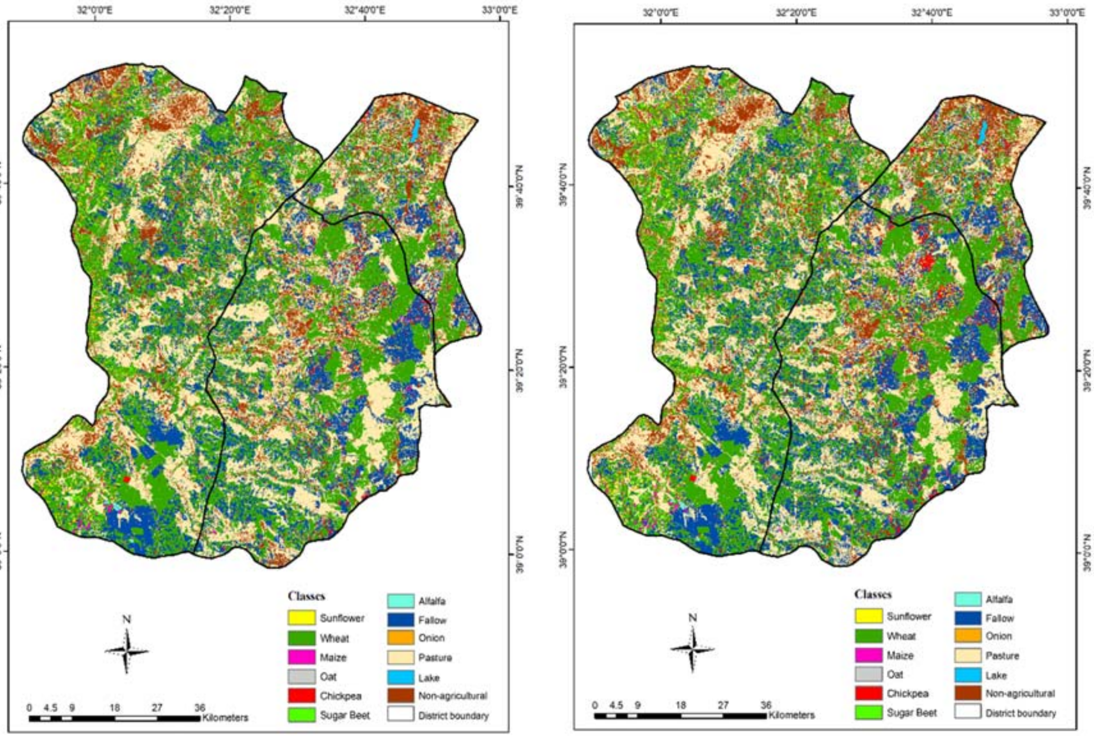

Fig 4. Crop classification of Sentinel-2 images according to RF (left) and SVM (right) models

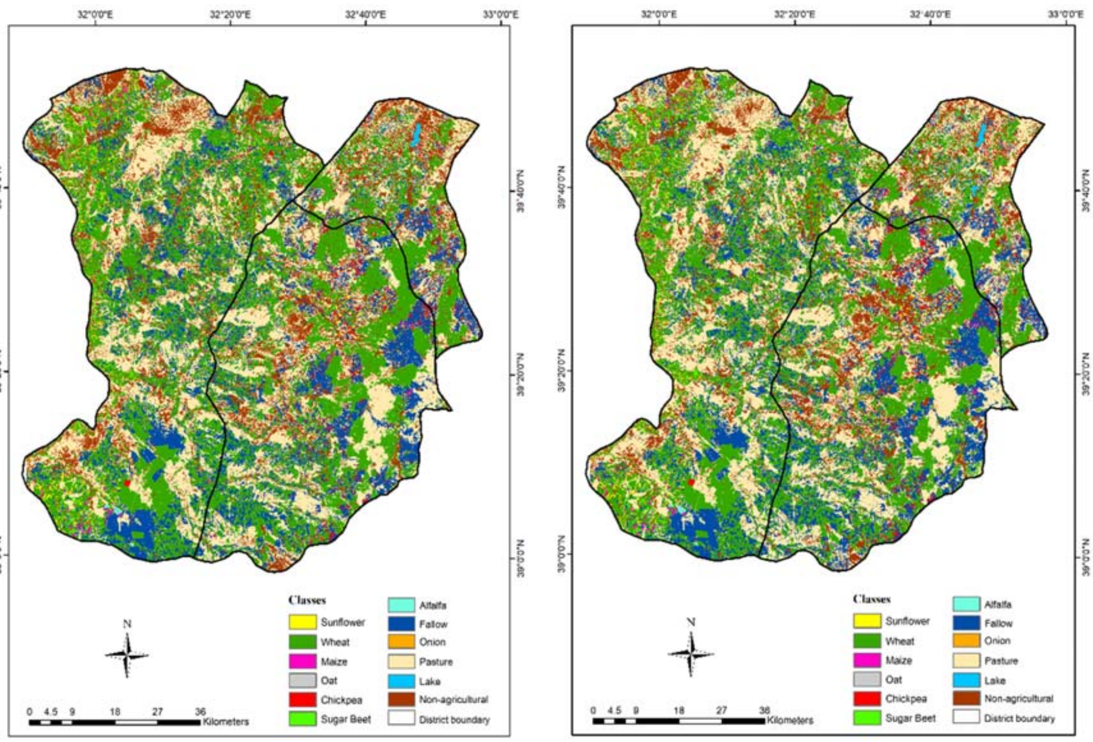

Fig 5. Crop classification of Landsat-8 images according to RF (left) and SVM (right) models

According to the evaluation results, RF model has the highest accuracy with OA (89.5%) and K (0.88) values for Sentinel-2 data. RF model was followed by SVM model with OA (84.6%) and K (0.83) values. On the other hand, the results obtained for RF (OA: 79%, K: 0.76) and SVM (OA: 78.1%, K: 0.75) models with Landsat-8 data were lower than Sentinel-2 (Table 3). For Sentinel-2 data, the user and producer accuracies for each class with the RF model ranged from 77.8% to 99.8%, while for SVM they ranged from 67.2% to 97.5%. For Landsat 8 data, these accuracies ranged between 60.8% and 99.2% for the RF model and between 57% and 97.3% for the SVM model (Table 3). Crop classification accuracies obtained from both RF and SVM models using Sentinel-2 data are higher than Landsat-8 data, which is confirmed by previous studies (Saini and Ghosh, 2018; Htitiou et al., 2019; Bofana et al., 2020; Ahady and Kaplan, 2022; Savitha and Talari, 2023).

Figures 6 and 7 shows the UA and PA accuracies for identifying land use classes using Sentinel-2 and Landsat 8 data for 2019. In the distribution of the accuracy matrix, the user accuracy of all classes is above 80% for Sentinel-2 image and RF model. PA is above 80% except for Oats (77.8%). Wheat is the most widely grown crop as an important crop in the region. In the classification of wheat, RF and SVM models performed well for Sentinel 2 and Landsat 8 images with UA values above 85% and PA values above 95% (Figure 6). For Sentinel-2 and SVM model, user accuracy is above 80% except for Maize (67.2%) and Chickpea (78.3%). PA is above 80% except for Sunflower (74.2%), Chickpea (78.1%) and Non-agricultural (69.5%). In the accuracy matrix evaluation for Landsat-8, the user accuracy of Maize, Oat, Sugar Beet, Alfalfa, Fallow and Pasture classes and PA of Wheat, Sugar Beet, Fallow and Non-Agricultural classes are above 80% for RF model. In Landsat-8 and SVM model, the user accuracy of Wheat, Oat, Sugar Beet, Alfalfa and Fallow classes and the producer accuracy of Wheat and Sugar Beet classes were above 80% (Figure 7).

The highest accuracy values of Sentinel-2 images and RF model were found in Sugar Beet (99.0%) and Alfalfa (99.8%) classes for UA, while for PA, the highest accuracy values were found in Wheat (98.5%) and Sugar Beet (98.5%) classes. According to the results obtained from the SVM model, the highest prediction values were obtained in the Alfalfa (95.9%) class for UA and in the Wheat (97.5%) class for PA (Table 3). A similar evaluation was made for Landsat-8 images. In RF and SVM models, the highest UA and PA values are above 95% for the Alfalfa and Wheat classes. For Sentinel 2 images, the lowest accuracy values were obtained for UA in the Chickpea (80.6%) class in the RF model and in the Maize (67.2%) class in the SVM model, while for PA in the Oat (77.8%) class in the RF model and in the Sunflower (74.2%) and Chickpea (78.1%) classes in the SVM model. Sugar beet and alfalfa crops achieved the highest accuracy values with both UA (99-99.8%) and PA (98.5-96%) in the RF model with Sentinel 2 images (Table 3). However, the lowest accuracy values for Landsat-8 imagery were obtained in the Sunflower (60.8%) class in the RF model for UA and in the Sunflower (66.6%), Chickpea (67.2%) and Onion (67.8%) classes in the SVM model, while for PA, the Oat (61.2%) and Chickpea (63.8%) classes in the RF model and Sunflower (60.4%) in the SVM model were obtained (Table 3).

According to the results, wheat was the most accurately predicted class for PA by RF and SVM models in both Sentinel-2 and Landsat-8 images. The same situation was obtained for UA in the alfalfa class. It was also observed that oats mixed slightly with wheat and chickpeas. Among the summer crops, corn was mixed with sunflower and onion crops, which are in similar phenological periods. Some mixing between sunflower and onion may be due to the canopy structure and surface coverage of onion. In general, the findings of this study showed that Sentinel-2 data can be used successfully in areas with fragmented farmland and heterogeneous crop structure. Satisfactory results were obtained in areas with similar terrain conditions (She vd., 2020).

Since the sample size is limited in large areas, reference data sets have an impact on classification accuracies (Li et al., 2021). In this study, satisfactory results were obtained by selecting plots with controlled sampling over land cover classes with stratified random sampling. Combining sampling points with multi-temporal images linked to phenological periods improves classification accuracy. In similar studies, multi-temporal images are preferred for many crops (Erdanaev et al., 2018; Vuolo et al. 2018; Gumma et al., 2020). RF was used more effectively than SVM in terms of processing satellite data containing crop growth periods spanning the whole season and in terms of memory capacity utilization. Adugna et al. (2022) RF model is more useful than SVM model in terms of processing large input data and memory consumption. Since the use of multi-time high dimensional data will increase the data volume, the creation of optimal data sets can improve model performances.

Conclusion

In this study, agricultural crop pattern performance results were compared by applying different ML algorithms using multi-temporal Sentinel-2 and Landsat-8 images. It was found that satellite images with spatial and temporal differences have significant effects on the crop classification performance of the data obtained. It has been shown that RF and SVM algorithms can be successfully applied from ML techniques using Sentinel-2 and Landsat-8 images. The highest accuracy among the models for classification was obtained from RF results using Sentinel-2 data. When the image and model results were evaluated together, it was determined that the RF and SVM model accuracy results obtained from Sentinel-2 data were higher than Landsat-8 data.

Table 3. Accuracy assessment of Sentinel-2 and Landsat-8 images with RF and SVM models

	Sentinel -2				Landsat-8			
	RF		SVM		RF		SVM	
Classes	**UA (%)**	**PA (%)**	**UA (%)**	**PA (%)**	**UA (%)**	**PA (%)**	**UA (%)**	**PA (%)**
Sunflower	81.8	85.8	82.0	74.2	60.8	76.8	66.6	60.4
Wheat	87.8	98.5	86.5	97.5	79.8	96.5	84.0	97.3
Maize	93.9	82.6	67.2	92.6	84.3	78.5	79.3	78.9
Oat	94.6	77.8	92.9	81.4	88.6	61.2	90.1	74.5
Chickpea	80.6	86.5	78.3	78.1	71.4	63.8	67.2	74.6
Sugar Beet	99.0	98.5	87.0	94.8	96.5	88.8	93.5	86.8
Alfalfa	99.8	96.0	95.9	93.3	99.2	73.5	95.2	75.2
Fallow	91.9	89.1	87.2	80.6	90.7	88.4	82.2	74.4
Onion	82.4	92.0	85.0	80.6	75.1	77.8	67.8	75.7
Pasture	90.3	89.3	85.4	83.3	82.6	73.0	78.2	76.8
Non-agricultural	92.9	84.8	80.8	69.5	65.3	83.6	57.0	64.1
OA (%)	**89.5**		**84.6**		**79.0**		**78.1**	
Kappa	**0.88**		**0.83**		**0.76**		**0.75**	

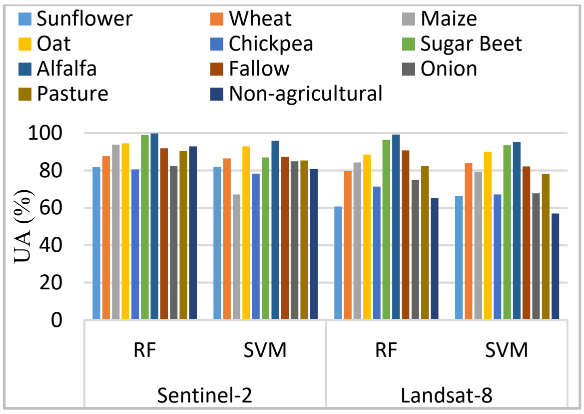

Fig 6. User accuracy of RF and SVM algorithms based on Sentinel-2 and Landsat-8

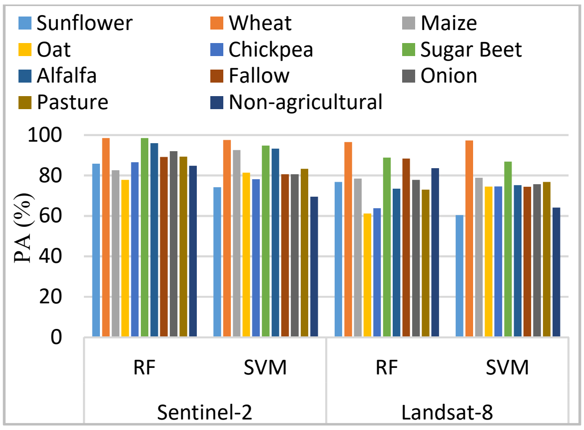

Fig 7. Producer accuracy of RF and SVM algorithms based on Sentinel-2 and Landsat-8

The high temporal and spatial resolution of Sentinel 2 images is considered to be the most important factor in the results. Especially in areas with mixed agriculture and small parcels, this situation is more prominent. The 16-day temporal resolution of Landsat 8 imagery and the unavailability of imagery during critical plant growth periods due to cloudiness caused a decrease in accuracy levels. In addition, the spatial resolution in plots with more than one agricultural crop and plots with low area caused the accuracy performance to decrease for some crop classes. According to the results obtained, it was concluded that Landsat images can be used in the classification of crops with similar phenological periods, but it cannot exceed a certain level due to its spatial and resolution.

Sentinel-2 imagery provided better results in crop classification than Landsat-8 due to its spatial and temporal resolution. In particular, it was concluded that it is more suitable for classification studies in parcels with small areas. Moreover, the RF method performed better in identifying crop types using Sentinel-2 data. Crops in small patches and with different phenological stages were successfully classified using optical remote sensing data such as Sentinel-2. In this study, the use of multi-temporal data provided a good result in distinguishing crops with consecutive phenological periods. The multi-spectral and temporal data of Sentinel-2 images significantly improves the classification accuracy, which is strongly correlated with the band characteristics selected. For crop varieties with early, mid and late growth stages, the combination of temporally different images can provide optimum performance. At this stage, image selection and cloudiness are important.

While high temporal and spatial resolution over large areas such as Sentinel 2 images contributes significantly to classification accuracy, increasing the number and volume of images increases data storage space and processing time. In addition, image selection, downloading and preprocessing can increase the processing time depending on the technical specifications of the computer.

The study results are satisfactory due to the size of the area and the diversity of the cropping pattern. These data are of strategic importance for source data management and agricultural crop planning, especially regional and national studies. Crop classification maps created using satellite images and ML algorithms can play an effective role as basic data for decision makers and producers in creating agricultural planning models such as crop support, yield and agricultural water use.

References

Adugna, T., Xu,W., Fan, J. (2022). Comparison of Random Forest and Support Vector Machine Classifiers for Regional Land Cover Mapping Using Coarse Resolution FY-3C Images. *Remote Sens.*, 14, 574. doi.org/10.3390/rs14030574

Ahady, A. B., Kaplan, G. (2022). Classification comparison of Landsat-8 and Sentinel-2 data in Google Earth Engine, study case of the city of Kabul. *International Journal of Engineering and Geosciences;* 7(1); 24-31.

Alami Machichi, M., Mansouri, L. E., Imani, Y., Bourja, O., Lahlou, O., Zennayi, Y., Hadria, R. (2023). Crop mapping using supervised machine learning and deep learning: a systematic literature review. *International Journal of Remote Sensing,* 44(8), 2717–2753.

Altun M., Türker, M. (2021). Çoklu Zamanlı Sentinel-2 Görüntülerinden Tarımsal Ürün Tespiti: Mardin-Kızıltepe Örneği. *Afyon Kocatepe Üni. Fen ve Müh. Bilimleri Dergisi,* 21(4), 881-899. doi:10.35414/akufemubid.890436

Anua, S. N., Wong, W V C. (2022). Utilizing Landsat 8 OLI for land cover classification in plantations area. IOP Conf. Ser.: *Earth Environ. Sci.* 1053, 012027.

Bantchına, B. B., Gündoğdu, K. H. (2024). Crop Type Classification using Sentinel 2A-Derived Normalized Difference Red Edge Index (NDRE) and Machine Learning Approach. *Bursa Uludağ Üni. Ziraat Fak. Der.,* 38 (1), 89-105.

Basukala, A. K., Oldenburg, C., Schellberg, J., Sultanov, M., Dubovyk, O. (2017). Towards improved land use mapping of irrigated croplands: performance assessment of different image classification algorithms and approaches, *European Journal of Remote Sensing,* 50:1, 187-201, doi.10.1080/22797254.2017.1308235

Blickensdorfer, L., Schwieder, M., Pflugmacher, D., Nendel, C., Erasmi, S., Hostert, P. (2022). Mapping of crop types and crop sequences with combined time series of Sentinel-1, Sentinel-2 and Landsat 8 data for Germany. *Remote Sensing of Environment* 269, 112831.

Bofana, J., Zhang, M., Nabil, M., Wu, B., Tian, F., Liu, W., Zeng, H., Zhang, N., Nangombe, S. S., Cipriano, A. S., Phiri, E., Mushore, D. T., Kaluba, P., Mashonjowa, E., Moyo, C. (2020). Comparison of Different Cropland Classification Methods under Diversified Agroecological Conditions in the Zambezi River Basin. *Remote Sens,* 12, 2096;

Breiman, L (1999*). Random forests-random features. Technical Report 567,* Statistics Department, University of California, Berkeley.

Cai, Y., Guan, K., Peng, J., Wang, S., Seifert, C., Wardlow, B., Li, Z. (2018). A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sens. Environ.,* 210, 35–47.

Çelik, O. İ., Büyüksalih, G., Gazioğlu, C. (2023). Improving the Accuracy of Satellite-Derived Bathymetry Using Multi-Layer Perceptron and Random Forest Regression Methods: A Case Study of Tavşan Island. *Journal of Marine Science and Engineering*, *11*(11), 2090.

Chakhar, A., Ortega-Terol, D., Hernández-López, D., Ballesteros, R., Ortega, F. J., Moreno, A. M. (2020). Assessing the Accuracy of Multiple Classification Algorithms for Crop Classification Using Landsat-8 and Sentinel-2 Data. *Remote Sens.,* 12, 1735.

Colditz, R. (2015). An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms. *Rem. Sens.* 7, 9655. doi.org/10.3390/rs70809655.

Çölkesen, İ., Kavzoğlu, T. (2008). Destek vektör makineleri kullanarak arazi örtüsünün sınıflandırılması: Gebze örneği. 2. Uzaktan

Algılama ve Coğrafi Bilgi Sistemleri Sempozyumu, 13-16 Ekim 2008, 35-45, Kayseri.

Congalton, R. G. (1991). "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." *Remote Sensing of Environment* 37 (1): 35–46. doi:10.1016/0034- 4257(91)90048-B.

Debats, S.R., Luo, D., Estes, L.D., Fuchs, T.J., Caylor, K.K. (2016). A generalized computer vision approach to mapping crop fields in heterogeneous agricultural landscapes. *Remote Sens. Environ.,* 179, 210–221.

Duro D. C., Franklin S. E., Dubé, M. G. (2012). A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sensing of Environment* 118, 259-272.

Erdanaev, E., Kappas, M., Wyss, D. (2018). The Identification of Irrigated Crop Types Using Support Vector Machine, Random Forest and Maximum Likelihood Classification Methods with Sentinel-2 Data in 2018: Tashkent Province, *Uzbekistan Internation Journal of Geoinformatics,* Vol.18, No.2.

Escabias, C. B. (2017). *Tree Boosting Data Competitions with XGBoost* (Master's thesis, Universitat Politècnica de Catalunya).

Esetlili, M. T., Bektas Balcik, F., Balik Sanli, F., Kalkan, K., et al. (2018). Comparison of Object and Pixel-Based Classifications For Mapping Crops Using Rapideye Imagery: A Case Study Of Menemen Plain, Turkey. *International Journal of Environment and Geoinformatics,* 5(2), 231-243. doi.org/10.30897/ijegeo.442002.

Fu, Y., Shen, R., Song, C., Dong, J., Han, W., Ye, T., Yuan. W. (2023). *Exploring the effects of training samples on the accuracy of crop mapping with machine learning algorithm Science of Remote Sensing* 7, 100081.

Gorji, T., Yıldırım, A., Sertel, E., Tanık, A. (2019). Remote sensing approaches and mapping methods for monitoring soil salinity under different climate regimes. *International Journal of Environment and Geoinformatics,* 6(1), 33-49. doi.org/10.30897/ijegeo.500452.

Gumma, M. K., Tummala, K., Dixit, S., Collivignarelli, F., Holecz, F., Kolli, R. N., Whitbread, A. M. (2020). Crop type identification and spatial mapping using Sentinel-2 satellite data with focus on field-level information, *Geocarto International,* doi.10.1080/10106049.2020. 1805029

Heupel, K., Spengler, D., Itzerott, S. A. (2018). Progressive Crop-Type Classification Using Multitemporal Remote Sensing Data and Phenological Information. *PFG*, 86, 53–69, doi:10.1007/s41064-018-0050-7.

Htitiou, A., Boudhar, A., Lebrini, Y., Hadria, R., Lionboui, H., Elmansouri, L., Tychon, B., Benabdelouahab, T. (2019). The Performance of Random Forest Classification Based on Phenological Metrics Derived from Sentinel-2 and Landsat 8 to Map Crop Cover in an Irrigated Semi-Arid Region. *Remote Sensing in Earth Systems Sciences* doi.org/10.1007/s41976-019-00023-9

Huang, C., Davis, L. S., Townshend, J. R. G. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23: 725-749.

Immitzer, M., Vuolo, F., Atzberger, C. (2016). First experience with sentinel-2 data for crop and tree species classifications in central europe. *Remote Sens.* 2016, 8, 166.

Inglada, J., Arias, M., Tardy, B., Hagolle, O., Valero, S., Morin, D., Dedieu, G., Sepulcre, G., Bontemps, S., Defourny, P., Koetz, B. (2015). Assessment of an Operational System for Crop Type Map Production Using High Temporal and Spatial Resolution Satellite Optical Imagery. *Remote Sens.*, *7*,12356-12379. doi.org/10.3390/ rs70912356

Karmakar, P., Teng, S. W., Murshed, M., Shaoning Pang, S., Li, Y., Lin, H. (2024). Crop monitoring by multimodal remote sensing: A review. Remote Sensing Applications: *Society and Environment* 33, 101093

Li, C., Ma, Z., Wang, L., Yu, W., Tan, D., Gao, B., Feng, Q., Guo, H., Zhao, Y (2021). Improving the Accuracy of Land Cover Mapping by Distributing Training Samples. *Remote Sens.,* 13, 4594.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., Bochtis, D. (2018). Machine Learning in Agriculture: A Review. *Sensors,* 18, 2674.

Liu, B., Gao, L., Li, B., Marcos-Martinez, R., Bryan, B. (2020). Nonparametric machine learning for mapping forest cover and exploring influential factors. *Landscape Ecol* 35, 1683–1699 (2020).

Löw, F., Michel, U., Dech, S., Conrad, C. (2013). Impact of feature selection on the accuracy and spatial uncertainty of per-field crop classification using support vector machines. ISPRS J. Photogramm. *Remote Sens.,* 85, 102–119.

Mathur, A., Foody, G. M. (2008). Crop classification by support vector machine with intelligently selected training data for an operational application. *International Journal of Remote Sensing,* 29, 2227-2240. doi.org/10. 1080/01431160701395203.

Nasrallah, A., Baghdadi, N., Mhawej, M., Faour, G., Darwish, T., Belhouchette, H., Darwich, S. A. (2018). A Novel approach for mapping wheat areas using high resolution sentinel-2 images. *Sensors*, 18, 2089.

Ozdogan, M. (2010). The spatial distribution of crop types from modis data: Temporal unmixing using independent component analysis. *Remote Sens. Environ.,* 114, 1190–1204.

Pasternak, M., Pawluszek-Filipiak, K. (2022). The Evaluation of Spectral Vegetation Indexes and Redundancy Reduction on the Accuracy of Crop Type Detection. *Appl. Sci.,* 12, 5067. doi.org/10.3390/app12105067

Pott, L.P., Amado, T.J.C., Schwalbert, R.A., Corassa, G.M., Ciampitti, I.A. (2021). Satellite-based data fusion crop type classification and mapping in Rio Grande do Sul, Brazil. *ISPRS J. Photogrammetry Remote Sens.* 176, 196–210.

Remelgado, R., Zaitov, S., Kenjabaev, S., Stulina, G., Sultanov, M., Ibrakhimov, M., Akhmedov, M., Dukhovny, V., Conrad, C. A. (2020). *Crop Type Dataset for Consistent Land Cover Classification in Central Asia. Sci Data,* 7, 250, doi:10.1038/s41597-020-00591-2.

Saini, R., Ghosh, S. K. (2018). Crop Classification on Single Date Sentinel-2 Imagery Using Random Forest and Support Vector Machine. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-5, *ISPRS TC V Mid-term Symposium "Geospatial Technology – Pixel to People",* 20–23 Nüvemce, Dehradun, India

Savitha, C., Talari, R. (2023). Mapping cropland extent using sentinel-2 datasets and machine learning algorithms for an agriculture watershed. *Smart Agricultural Technology* 4, 100193.

See, L., Fritz, S., You, L., Ramankutty, N., Herrero, M., Justice, C., Becker-Reshef,I., Thornton, P., Erb, K., Gong, P., Tang, H., Van Der Velde, M., Ericksen, P., McCallum, I., Kraxner, F., Obersteiner, M. (2015). Improved global cropland data as an essential ingredient for food security. Glob. *Food Secur.,* 4, 37–45.

Segarra, J., Araus, J. L., Kefauver, S. C. (2022). Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield. *International Journal of Applied Earth Observations and Geoinformation* 107, 102697.

She, B., Yang, Y., Zhao, Z., Huang, L., Liang, D., Zhang, D. (2020). Identification and mapping of soybean and maize crops based on Sentinel-2 data. *Int J Agric & Biol Eng* 13(6) 171.

Şimşek, F. F. (2024). Hızlandırılmış makine öğrenmesi algoritmaları ile tarım parseli tabanlı ürün desen sınıflandırması. *Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi,* 29(1),314-330. doi.org/10.53433/yyufbed.1416820

Şimşek, F. F., Durduran, S. S. (2023). Açık kaynak kodlu Eo-learn kütüphanesi ve çok zamanlı Sentinel-2 görüntüleri ile tarımsal ürün sınıflandırması. *Journal of Geodesy and Geoinformation,* 10(1), 45-62. doi.org/10.9733/ JGG.2023R0004.T

Song, X.-P., Potapov, P.V., Krylov, A., King, L., Di Bella, C. M., Hudson, A., Khan, A., Adusei, B., Stehman, S.V., Hansen, M.C. (2017). National-scale soybean mapping and area estimation in the united states using medium resolution satellite imagery and field survey. *Remote Sens. Environ.,* 190, 383–395.

Tatsumi, K., Yamashiki, Y., Torres, M. A. C., Taipe, C. L. R. (2015). *Crop classification of upland fields using Random forest of time-series Landsat 7 ETM+ data. Computers and Electronics in Agriculture* 115, 171-179.

Thanh, N. P., Kappas, M. (2018). Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors* 18(1) 18. doi.org/10.3390/s18010018

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens Environ* 8:127–150.

Tuvdendorj, B., Zeng, H., Wu, B., Elnashar, A., Zhang, M., Tian, F., Nabil, M., Nanzad, L., Bulkhbai, A., Natsagdorj, N. (2022). Performance and the Optimal Integration of Sentinel-1/2 Time-Series Features for Crop Classification in Northern Mongolia. *Remote Sens.,*14, 1830.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory.* New York: Springer-Verlag, p. 188.

Verma, P., Raghubanshi, A., Srivastava, P. K., Raghubanshi, A. S. (2020). Appraisal of kappa-based metrics and disagreement indices of accuracy assessment for parametric and nonparametric techniques used in LULC classification and change detection. *Model.EarthSyst. Environ.* 6,1045–1059.

Vogiatzis, M., Eleftheriadis, I. (2023). *Comparison of Pixel-Based Classification Algorithms Using Landsat-8 OLI and Sentinel-2 MSI for Land Use/Land Cover Mapping in a Heterogeneous Landscape.* doi:10.20944/preprints202307.1043.v2

Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., Ng, W.-T. (2018). How much does multi-temporal sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.,* 72, 122–130.

Wakulińska, M., Marcinkowska-Ochtyra, A. (2020). Multi-Temporal Sentinel-2 Data in Classification of Mountain Vegetation. *Remote Sensing* 2020, 12, 2696, doi:10.3390/rs12172696.

Yaşar, O., Yağcı, A. L. (2023). Yersel referans verilerinin doğruluğunun çok zamanlı Sentinel-2 uydu görüntüleri ile araştırılması: Arpa ve Buğday örneği. *Geomatik*, 8(3), 277-292

Yi, Z., Jia, L., Chen, Q. (2020). Crop Classification Using Multi-Temporal Sentinel-2 Data in the Shiyang River Basin of China. *Remote Sensing*., 12(24):4052. doi.org/10.3390/rs12244052

Zhang, H. K., Roy, D. P. (2017). Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sensing of Environment* 197, 15-34. doi.org/10.1016/j.rse.2017.05.02

Zheng, B., Myint, S., Thenkabail, P. S., Aggarwal, R. (2015). A support vector machine to identify irrigated crop types using time-series Landsat NDVI data. *International Journal of Applied Earth Observation and Geoinformation* 34, 103–112.

Zhong, L., Gong, P., Biging, G. S. (2014). Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. *Remote Sensing of Environment* 140, 1-13. doi.org/10.1016/j.rse.2013.08.023.

.