International Journal of Engineering and Geosciences

https://dergipark.org.tr/en/pub/ijeg

e-ISSN 2548-0960

Comparison of Agricultural Crop Type Classifications with Different Machine Learning Algorithms by Generating Ground Truth Data from Farmer Declaration Parcels

Fatih Fehmi ŞİMŞEK[*1]

[1] TÜBİTAK Space Technologies Research Institute, Ankara , Türkiye fehmi.simsek@tubitak.gov.tr

Cite this study: Şimşek F.F. (2025). Comparison of Agricultural Crop Type Classifications with Different Machine Learning Algorithms by Generating Ground Truth Data from Farmer Declaration Parcels. International Journal of Engineering and Geosciences, 10 (2), 207-220.

https://doi.org/10.26833/ijeg.1538708

Keywords
ANN
FDP
RF
SVM
XGBoost

Research Article

Received: 18.09.2024
Revised: 21.10.2024
Accepted: 30.10.2024
Published: 01.07.2025

Abstract

In large-scale agricultural crop classification studies (Turkey, Adana, Çukurova Plain, 2500 km^2), collecting sufficient and accurate ground truth data is costly, time-consuming, and unsustainable. This study utilized parcels registered in the Farmer Registration System (FRS) as ground truth data. By analyzing time series EVI curves, discrepancies were identified between declared and actual crops. Erroneous parcels were eliminated, and the corrected data were used in the classification process.Using multi-temporal Sentinel-2 images from 2021, this study compared the performance of Random Forests (RF), Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Extreme Gradient Boosting (XGBoost) algorithms for classifying crops like citrus, cotton, maize, peanut, sunflower, watermelon, wheat, and double-crop combinations (e.g., wheat-cotton, wheat-maize). The classification utilized 121 features (11 images × 10 Sentinel-2 bands + EVI). XGBoost achieved the highest overall accuracy (92.14%), followed by RF (89.15%), SVM (86.14%), and ANN (85.48%).The EVI index proved critical, particularly in separating spectral curves of double crops. While single crops like cotton, maize, and wheat yielded high classification accuracy, double crops with overlapping phenological stages had lower accuracy. The study highlighted that crops at distinct phenological stages performed well across algorithms, whereas crops with similar stages struggled to achieve high accuracy.This method of using corrected farmer-declared parcels (FDP) as ground truth data demonstrated high classification performance across all algorithms, proving its reliability. The findings emphasize that FDP can effectively replace traditional field data collection, reducing costs and improving efficiency. This classification approach supports agricultural production monitoring, yield estimation, water resource analysis, and sustainable policy-making, serving as a robust tool for agricultural evaluation.

1. Introduction

Remote sensing technology is increasingly being used to survey and monitor agricultural land, increase productivity and make more efficient use of resources. With the increase in the number of freely available satellite images (Sentinel, Landsat, Modis), spatial, temporal and spectral resolution of satellites, agricultural applications such as detection and determination of agricultural crop types, crop water consumption, yield estimation have become one of the most widely used areas of remote sensing technology [1].

Therefore, professional management and monitoring of agriculture and water consumption are crucial. To effectively manage and monitor agricultural production, it is necessary to have information about the types of crops and cultivated lands. The increasing role of agriculture in the management of sustainable natural resources calls for the development of operational crop type mapping [2]. Traditional methods of obtaining and updating the crop type and planting area information are mainly based on sampling surveys and statistical reports [3] which have problems such as strong subjectivity, time consuming, labor-intensive, delayed updating, and the lack of spatial distribution information [4]. Agricultural applications are one of the most widely used areas of remote sensing technology and can be used to determine crop type classes [5].

Satellite imagery enables observing, identifying, mapping, and evaluating dynamic agricultural areas with different temporal and spatial resolutions [6]. Image classification is the most common method used in agricultural crop type detection with satellite images. In recent decades, satellite images with global coverage and different spectral, spatial temporal characteristics have become freely available, such as Landsat and MODIS [7]. MODIS imagery has been used extensively for over a decade, but the spatial resolution of these products needs to be revised to characterize individual parcels of farmland [8]. MODIS images are only appropriate for parcels larger than 32 ha [9], which limits their usefulness in assessing small cultivated plots. Although the Landsat satellite has alleviated this problem, In particular, the temporal resolution (16 days) and spatial resolution (30 m PAN - 15 m RGB) have been insufficient again. Furthermore, the 16-day revisit cycle of Landsat is insufficient to capture the different phenological information of crops. High-resolution and very-high-resolution (<1 m) imagery contain rich crop texture and structural information and can be used for crop type mapping under complex terrain and planting conditions, but their high cost, large data storage and computational requirements limit their application [10]. Sentinel-2A and Sentinel-2B satellites, with their 5-day acquisition period, 13 spectral bands and high resolution (10m-20m-60m), provide a major advantage in crop type classification studies. Especially, revisit cycle of Sentinel-2 is only five days, which is helpful for capturing phenological information of crops during, the sowing, growing and harvesting period [11]. Crop type classification studies can be carried out using either a single image or multiple images. It has been found that using multi-temporal images for crop type classification is more accurate than using single date images [12]. Phenological differences can be effectively used to distinguish one type of crop from another, as some crop types have similar phenological stages during certain periods. For this reason, phenological information of crops is also very important for the classification study. For the past few decades, satellite remote sensing images have been used for crop classification, employing various image processing and classification techniques. Many parametric (linear regression, naive bayes, neural network and deep learning) and non-parametric (SVM, RF, KNN, DT, bagging and boosting algorithms) algorithms are used in classification studies. Many object or pixel-based machine learning algorithms have been used in the literature for crop classification with different bands, indexes, product types, and single-image or multi-temporal images. Zhang et al. [13] in Heilongjiang, China, crop classification was performed using 12 bands of Sentinel-2 imagery and four plant indices (NDVI, NDSVI, NDRI, NDTI). They employed SVM, DT, and RF algorithms for classification. The RF algorithm achieved the highest accuracy at 97.85%, followed by the SVM algorithm at 97.22%, and the DT algorithm at 95.92%. Li et al. [14] using Sentinel-2 imagery and different vegetation indices in Minnesota, USA, corn, soya, wheat and sugar beet crops were classified. As a result of this study using deep learning (CNN) algorithm, 97.48% overall accuracy was achieved. They said that experimental results show limited accuracy for crop classification based on single temporal features (OA: ~81) whereas multiple spectral or spatial information of temporal features significantly improves (OA: ~97.43) the classification accuracy. Vuolo et al. [15] in the Marchfeld region of Austria, Sentinel-2 imagery and RF algorithm were used for crop type classification using single-time and multi-temporal imagery. Classification using the RF algorithm showed that reached low accuracies (OA: ~0.50) during the growing season (March-April - single date) and an increase in OA (~70) between the highest (May-June single date), with the highest accuracy achieved (OA: ~95) by using images from nine different dates (March-October). Arvor et al. [16] in Brazil, 250 m MODIS imagery and Enhanced Vegetation Index (EVI) time series were used to classify soya beans, maize and cotton (OA: ~74). Belgiu and Csillik [17] crop type classification was performed using 4-band Sentinel-2 imagery (R-G-B-NIR) and NDVI index in Italy, Romania and USA. In this study, object and pixel based classification was performed and RF algorithm and Time Weighted Dynamic Time Warping (TWDTW) algorithm were used. Object-based TWDTW outperformed pixel-based TWDTW in all three-study areas and the overall accuracy ranged between 78.05% and 96.19%.

2. Materials

2.1. Study area

The study area (2500km^2), Çukurova Plain, is located in southern Adana, a highly productive agricultural province in Turkey. The region has a continental climate that is mild in winter and hot with an average annual precipitation of 650 mm. The highest temperatures occur in August and the lowest in January, with an average of 255 sunny days per year. The climatic and temperature conditions make two different seasons of crop cultivation the main cropping. The major crops grown in this region are wheat, sunflower, peanuts, watermelon, maize, cotton, soybeans, and citrus trees, which are planted in about 95% of the study area.

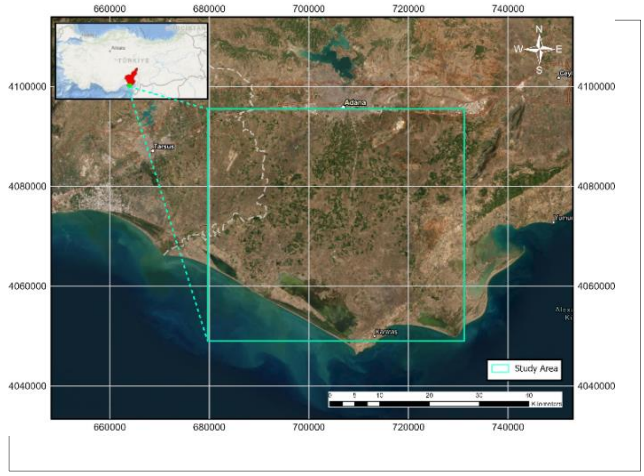


Figure.1 Study area Çukurova Plain- Adana.

2.2. Data

2.2.1. Sentinel-2

Sentinel-2 images were downloaded from the Copernicus (https://scihub.copernicus.eu/.com). Cloud-free images from March 2021 to October 2021 were selected to encompass the entire crop-growing season. A total of 10 Sentinel-2 images (Tables 1 and 2) were selected as the primary input data of the experiment. Data preparation included stacking and resampling the 20 m spectral bands to 10 m and removing the coastal band, water vapor, and the cirrus band, accomplished through the Sentinel Application Platform (SNAP).

Table 1. Spectral bands of Sentinel-2 images.

Band Names	Spectral Band	Spatial Resolution (m)
Blue	B2	10
Green	B3	10
Red	B4	10
Red-Edge	B5	20
Red-Edge	B6	20
Red-Edge	B7	20
NIR	B8	10
NIR	B8a	10
SWIR	B11	20
SWIR	B12	20

Table 2. Acquisition time of Sentinel-2

Day of Year (DOY)	Acquisition Time
88	29 March 2021
105	5 April 2021
125	5 May 2021
140	20 May 2021
165	6 June 2021
195	14 July 2021
223	11 August 2021
250	7 September 2021
275	2 October 2021
290	17 October 2021
300	27 October 2021

2.2.2. Farmer declaration parcels (train-test data)

The Farmer Declaration Parcels (FDP) in Turkey, also known as the Farmer Registration System (FRS), is a government initiative aimed at registering and tracking agricultural activities and farmers in the country. This system was implemented to improve the efficiency and transparency of agricultural practices and to provide various benefits to registered farmers. Under the FRS, farmers are required to register themselves and their agricultural activities. This registration process involves providing personal information and details about the land they are cultivating. Parcels that include agricultural activities registered in the system by farmers are called farmer declaration parcels (FDP). Parcels registered in the Farmer Registration System were used as ground truth data in the study. In addition to the geometric information of the parcels, the system also includes information on the province, district, parcel number, agricultural parcel number, cultivated product information, area, surface area, cadastral area, cultivation date and harvest date of each parcel. There were 87.692 parcels registered in 2021 in the study area. When these declaration-based parcels were examined, it was found that there had been systematic and non-systematic differences between the geometry and attribute information of the parcels comparing to the field reality. Due to these differences, the declared parcels underwent editing and deletion processes. At the end of these processes, ground truth data were produced from the declared parcels and used as training and test data in the classification process. After farmer declaration parcels (FDP) were used as reference data in the classification study.

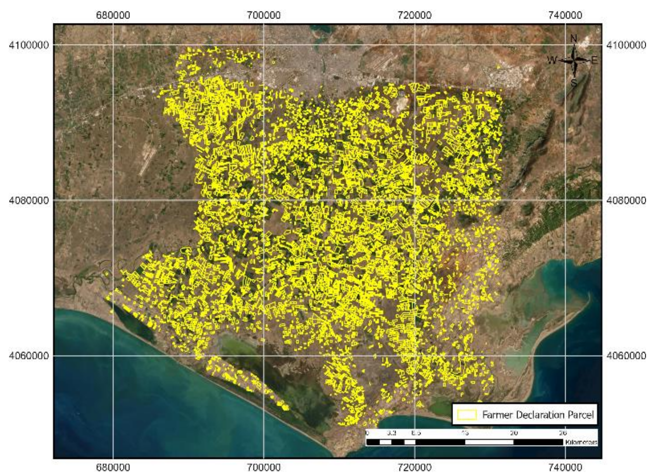


Figure.2 Map of the study area in Adana, with the reference

3. Methodology

3.1. Methodology overview

The overall methodology of the study is described in Figure 3. The method consists of four basic steps: satellite image pre-processing, reference data preparation, machine learning classification, and accuracy analysis. Firstly, the pre-processing steps of Sentinel-2 images were applied and the bands and indices used in the classification study were determined. Afterward, editing and deletion operations were performed on the parcels that have declarations. At the end of this process, the spectral separation curves of each product were determined by using the remaining parcels and the images created with the time series EVI index. Thirdly, pixel-based classification was performed using RF, SVM, ANN and XGB algorithms. Finally, the accuracy analyses of the obtained results were calculated and the algorithms' results were compared.

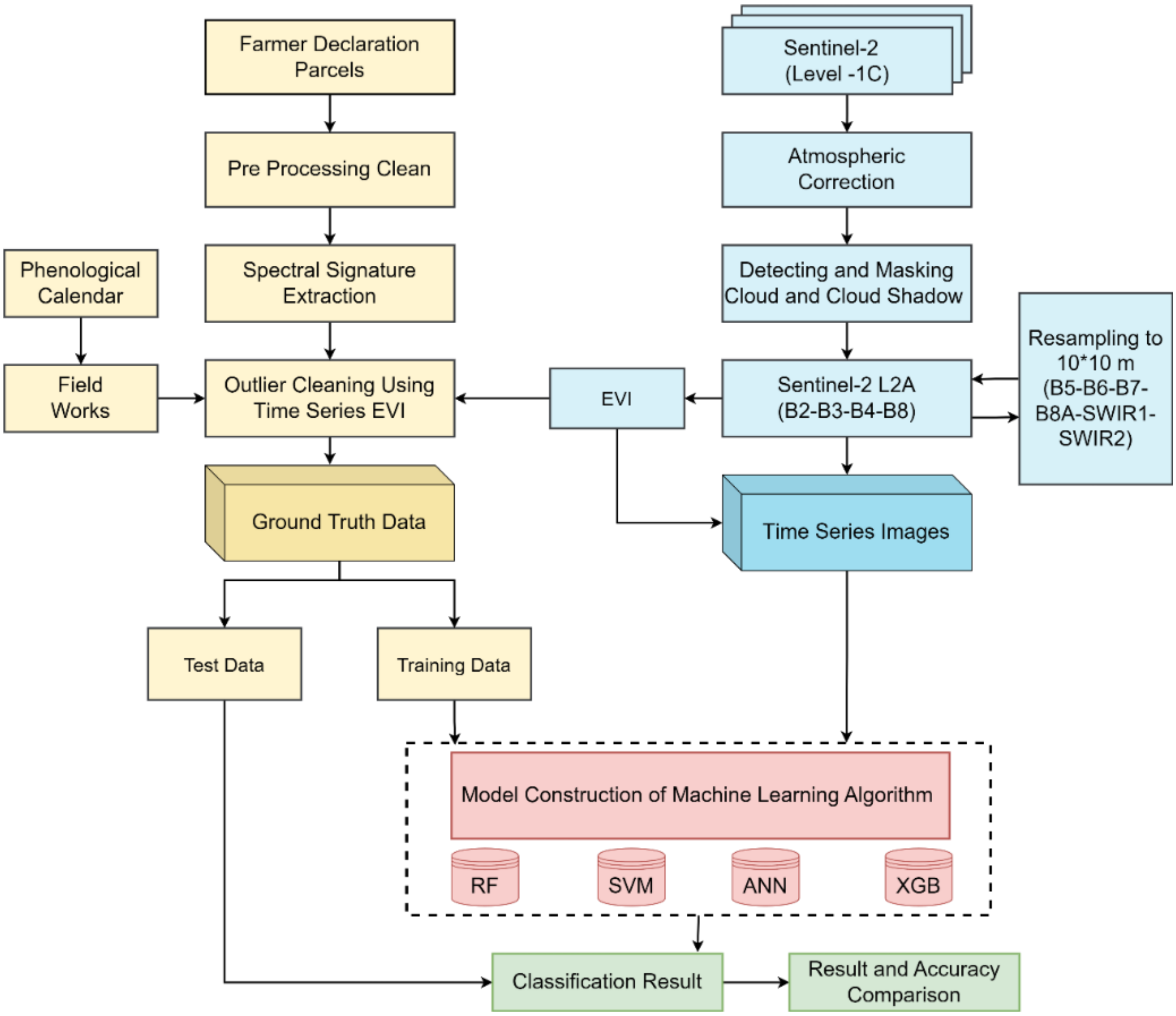


Figure.3 General workflow of this study

3.2. Satellite image pre-processing

In data processing, there is no need for further geometric correction of L1C products; only atmospheric correction and spatial resolution resampling are required [18]. Sentinel-2 images are provided in Level 1C format and contain above-atmosphere reflectance values. To calculate the actual reflectance values of plants in a classification study, top-of-atmosphere (TOA) values should be converted to bottom-of atmosphere reflectance values should be converted to bottom-of-atmosphere reflectance values (BOA) [19].Atmospheric effects were eliminated by converting TOA values to BOA values using the Sen2cor plugin. After Sen2Cor processing, the L1C TOA reflectance values were transformed into Level-2A (L2A) Bottom-of-Atmosphere (BOA) reflectance values [20]. Clouds resulting from atmospheric effects and shadows cast by clouds in satellite images are the primary sources of noise that pose challenges in challenges in image analysis. The brightness caused by clouds and shadows negatively impacts data analysis. These effects can lead to changes in values of NDVI and other indices, causing errors in various analyses and classification processes [21]. Cloud-covered areas cause anomalies in the bands in the images as well as the pixel values in the indexes created from these bands, which adversely affect the classification result [22]. In order to eliminate these anomalies in pixel values and increase the classification accuracy, clouds and shadow areas caused by clouds were detected and masked by the Sen2cor software.

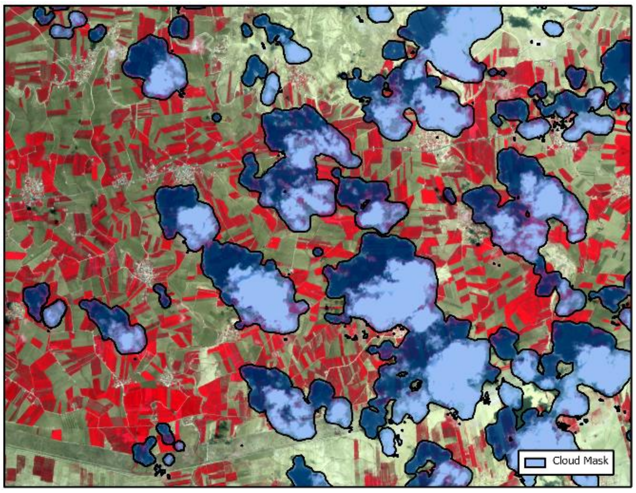


Figure.4 Cloud detection and masking.

3.3. Reference data preparation

The study area is located within the borders of Çukurova Plain and has a rich diversity of agricultural crops. Cotton, maize, wheat, barley, sunflower, sunflower, watermelon and peanuts are the region's most widely grown crops. Cotton, maize and soybeans are also planted as second crops after wheat harvest. The region is also rich with fruit trees of high economic value such as oranges, tangerines, grapefruit and lemons. Among the categorized crops, wheat and barley crops are classified under the wheat class, while orange, mandarin, grapefruit and lemon trees are classified under the citrus class. In addition to these crops, maize, cotton, watermelon, peanuts, sunflower and second crop maize, cotton and soybeans are also included in the classification study.

There are 87.692 farmer declaration parcels in the study area. It was also found that the operator caused errors during data entry in the system (crop planted, area planted, declaration change). Another problem is related to the declared agricultural parcels and it has been determined that some of these parcels have topological errors, overlaps or agricultural parcels within non-agricultural areas [23]. Before farmer declaration parcels could be used as ground truth data, pre-processing was carried out and the parcels with erroneous declaration were eliminated. As a result of these processes, the number of parcels decreased from 45.692 to 16.333. The biggest problem with the FDP is that the declaration and the actual crops are often different. Parcels with agricultural product declarations despite the non-existence of any agricultural activity constitute another problem. These wrong and incorrectly represented declaration parcels need to be removed to be used as ground truth data in the classification process.

An approximate calendar for all crop types in this geographical region is available in Figure.5. The calendar presents information about the sowing, growing and harvesting periods of the crops grown in the region. The exact dates of sowing and harvest vary depending on plots and farmers. There might be differences of up to 1–3 months in some cases. For certain crops, the harvesting of the first planted fields in the season overlaps with the sowing of the last fields. Therefore, no intermediate growth period is shown in this calendar.

Figure.5 Approximate crop calendar in the region

Different vegetation indexes are used to determine agricultural crop classification, NDVI, EVI, SAVI etc. index, which provide information on plant health. In this study, each parcel overlapped with multi-temporal NDVI images and NDVI median values of each plot were calculated in time series. After this process, each parcel's vegetative development and change were determined and the characteristics of NDVI curves showing variability over time were revealed (Figure 9). The characteristics NDVI curves for each crop were checked regarding the phenological calendar and the spectral reflectance values of each crop collected in the field study [24]. For each agricultural class outlier NDVI curves have been deleted and these parcels represent incorrect ground truth The characteristics of separation curves generated with NDVI differed for each agricultural crop (wheat, sunflower, peanut, watermelon, citrus, maize, cotton), which is a single harvest crop, while the characteristics of NDVI separation curves were found to be close to each other for agricultural crops with double harvests (wheat-maize, wheat-cotton, wheat-soybean) similar, this similarity decreased when EVI values were used (Figure8a-Figure8b). Therefore, the characteristic curves were reconstructed for all crops using EVI values instead of NDVI values (Figure 9). As a result of eliminating outlier parcels, 8549 FDP were used in the classification study as ground truth data. After atmospheric correction, cloud dedection and masking, 10 bands (B2, B3 B4, B5, B6, B7 B8, B8A, B11, and B12) were composited with 10 m resolution. The images produced as a result of the EVI index of each image were also included in the classification process. A total of 11*11 (121) bands composite images were used in the classification process.

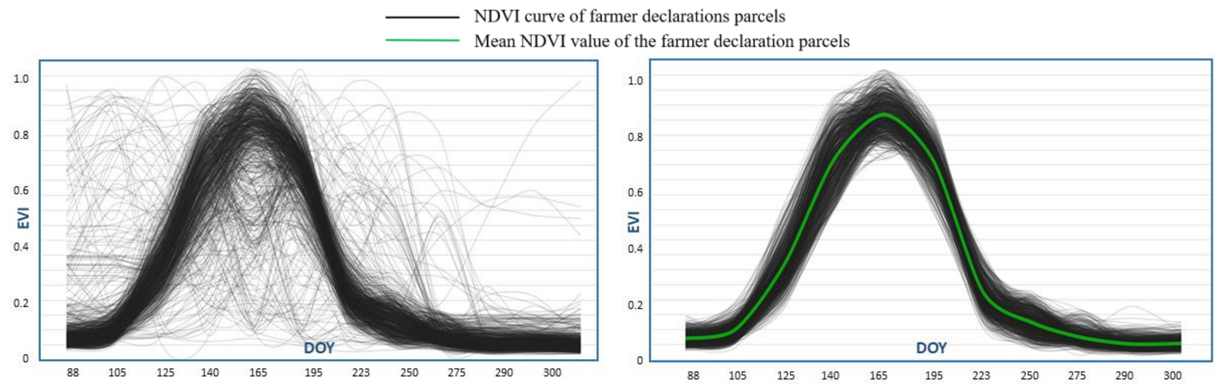


Figure.6a Non-clean farmer declaration parcels (Sunflower).

Figure.6b Cleaned farmer declaration parcels (Sunflower).

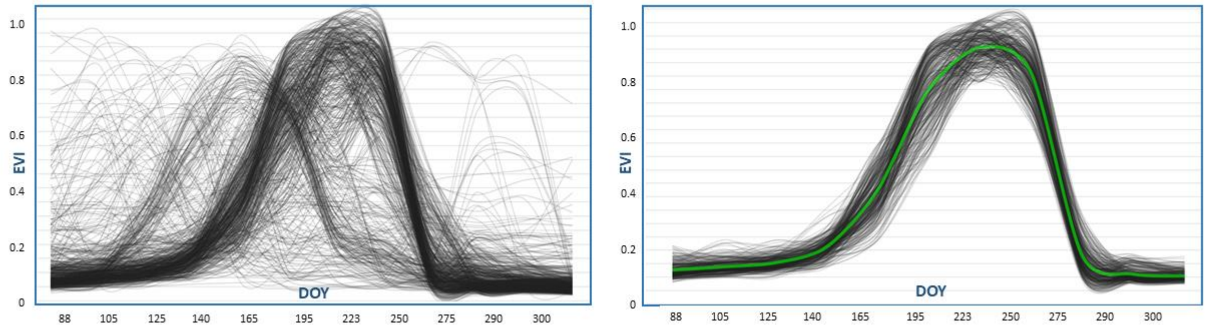


Figure.7a Non- clean farmer declaration parcels (Cotton).

Figure.7b Cleaned farmer declaration parcels (Cotton).

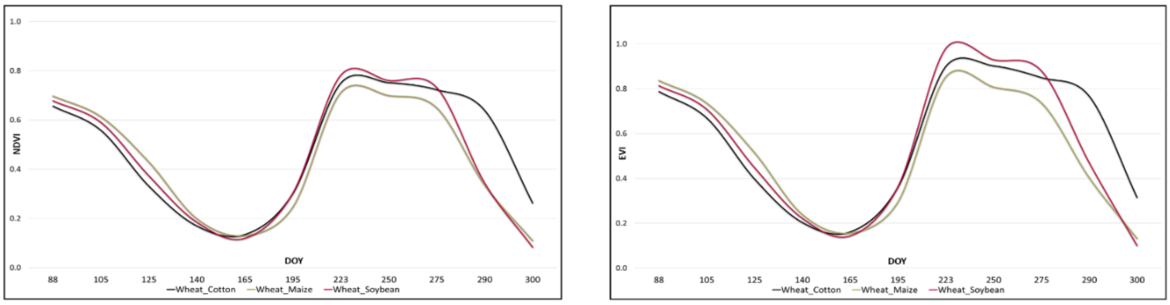


Figure.8a NDVI curve characteristics of double crops (Wheat-Cotton, Wheat-Maize, Wheat-Soybean).

Figure.8b EVI curve characteristics of double crops (Wheat-Cotton, Wheat-Maize, Wheat-Soybean).

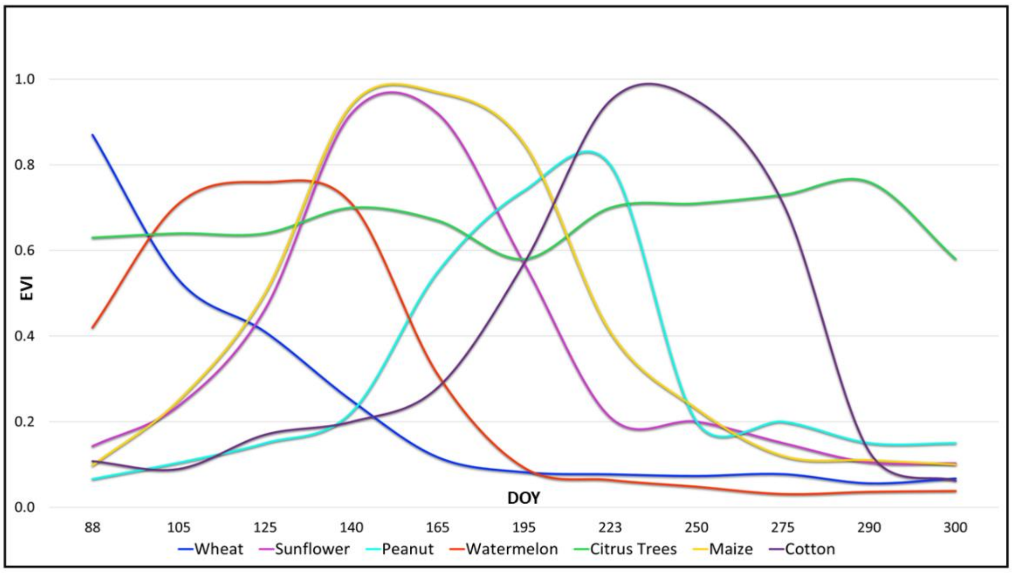


Figure.9 NDVI curves characteristics of single crops

4. Classification Algorithms and Tuning Parametres

4.1. Artificial neural network (ANN)

An artificial neural network (ANN) is a machine-learning algorithm that was originally developed to model the ability of a human brain to solve pattern recognition problems, and it has been increasingly used in remote sensing for image classification in recent years . The basic ANN framework consists of dense networks that are made up of interconnected neurons that are organized in layers and weights are assigned to these connections. These weights are first randomly determined and then iteratively adjusted for training; then, the effect on the output nodes is observed until the separation of the inputs and the predefined classes incur an error [25]. In this study, three parameters need to be set up: learning rate 0.001 to 0.1, number of layers: 1 to 5, batch size 16 to 512 [26].

4.2. Support vector machine (SVM)

A support vector machine (SVM) is a non-parametric learning algorithm that has frequently been used in remote sensing applications. In case of nonlinear classification, SVMs can perform the classification by using various types of kernels, which turns nonlinear boundaries to linear ones in the high-dimensional space to define optimal hyperplane [27];[28]. There are four kernel functions, which are linear, polynomial, radial basis function (RBF) and sigmoid kernels, which are the most frequently used types in SVM algorithms. Among them, the radial basis function (RBF) kernel outperforms the others with two parameters of the costs (C) and the kernel width parameter (γ) [29] [30].

4.3. Random forest (RF)

RF is a bagging-type ensemble algorithm that synthesizes predictions utilizing multiple decision trees. The RF classifier generally has a higher classification accuracy than a single decision tree. The random forest (RF) is an ensemble classifier that has been widely used in remote sensing studies currently due to its classification accuracy [31];[32]. Higher accuracies have been achieved with RF as compared to other machine learning algorithms in many crop mapping studies [33];[34]. In order to implement the RF, two parameters need to be set up: the number of trees (ntree) and the number of features in each split (mtry) [35]. Several studies have stated that satisfactory results could be achieved with the default parameters [36];[37]. In this study, to find the optimal RF model for classification, a range of values for both parameters were tested and evaluated: ntree = 100, 200, 400, 600 and 1000; mtry = 1:12 with a step size of 1.

4.4. Extreme gradient tuning machine (XGBoost)

The extreme gradient boosting (XGBoost) algorithm is a tree-based machine-learning algorithm and is a scalable implementation of the gradient boosting machines (GBM) algorithm, especially in data science studies where self-learning models achieve high performance [38]. XGBoost is a scalable machine-learning algorithm that minimizes the error rate by learning step by step by increasing the tree Structure [39]. XGBoost generates a series of decisions to predict the variable and each tree is designed to reduce the prediction errors of the previous trees. In this study, for XGBoost parametres max depth, min child weight, gamma, subsample and learning rate were used.

4.5. Tuning parametres

In the field of machine learning, tuning parameters, known as hyperparameters, refer to settings or configurations that are not learned from the training data but are established before training a model. These parameters exert control over various aspects of the learning process and can significantly influence the model's performance. The process of discovering the best parameters for a particular machine learning algorithm and dataset is commonly referred to as hyperparameter tuning or optimization. Tuned parameters are crucial for achieving high-accuracy results when utilizing artificial neural networks (ANN), support vector machines (SVM), random forests (RF), and XGBoost. Each classifier involves distinct tuning steps, tuned parameters, and optimal settings determined based on the highest overall classification accuracy. In this study, the optimum parameters of each classifier are shown in Table 3

Table.3 The optimal parametres for ML algorithm.

Algorithm	Hyperparameter	Optimum value
RF	ntree	800
	mtry	11
SVM	C	64
	gamma (γ)	0.125
ANN	learning rate	0.01
	number of layers	3
	batch size	32
XGB	max depth	10
	min child weight	3
	gamma	0.2
	subsample	1
	learning rate	0.05

In the classification study, 4274 of 8549 FDP parcels were selected as training data and 4275 of them were selected as test data to be used in the accuracy analysis Table 4.

Table.4 Training and testing data for classification

Class Name	Training Data	Test Data
Citrus tree	1516	1516
Cotton	844	844
Maize	980	980
Peanut	98	98
Sunflower	131	131
Watermelon	58	58
Wheat	230	231
Wheat-Cotton	139	139
Wheat-Maize	16	16
Wheat-Soybean	262	262
Total	**4274**	**4275**

A pixel-based classification study was performed with R software using RF, SVM, ANN and XGB machine learning algorithms.

5. Results and Discussion

Within the scope of this study, the performance of machine learning algorithms including RF, SVM, ANN, and XGB for agricultural crop classification has been compared. Parcels whose land cover has been classified by using pixel-based classification have been overlapped with the physical block created within the scope of the Integrated Administration and Control System (IACS) project. Non-agricultural areas (roads, artificial surfaces, forests, wetlands, water, bare land) were masked and eliminated using polygons within the physical blocks. As a result of this process, only the classification result representing the agricultural crop classification has been obtained Figure10-11-12.

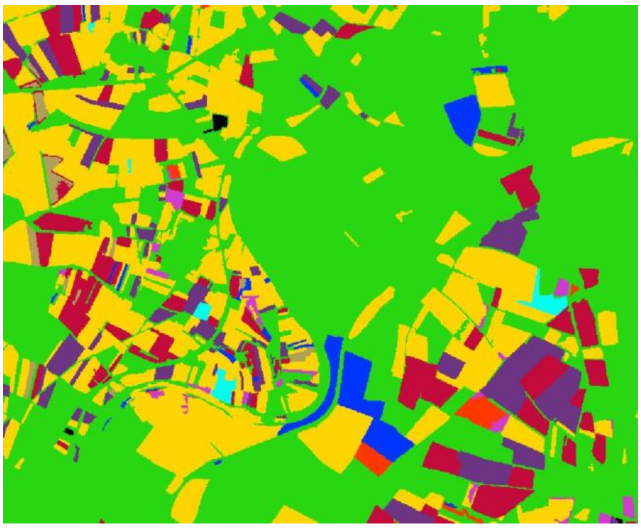


Figure.10 Crop classification result

Figure.11 Crop classification with overlap physical blocks

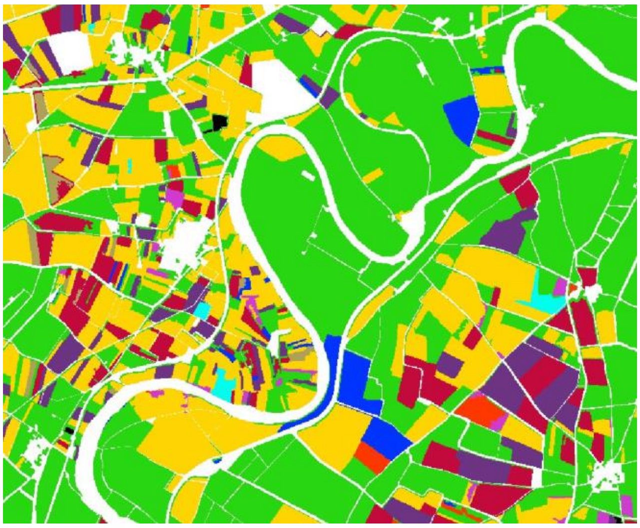
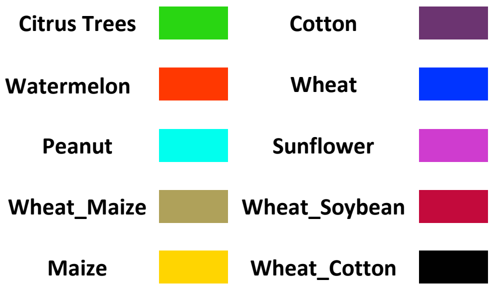


Figure.12 Masked crop classification result with physical blocks



In order to assess the accuracy of classification performance, there are many metrics available in the literature. In this work, Recall, Precision and F1-score have been used to assessment class accuracy. Precision is a metric that quantifies the number of the true positives that actually belong to true positives. In other words, it is the measure of accuracy. On the other hand, recall is a metric that quantifies the true positives made out of all true positive samples. The lower limit for both precision and recall scores is 0 and the upper limit is TP, FP, TN, and FN denote true positive, false positive, true negative, and false negative where each of the terms represent a special case depending on the relation between prediction and the ground truth. By using the Recall and Precision of each class, and the F1 score values, which are the harmonic average of these two values, the overall accuracy (Table 5-6-7-8) and each algorithm's confusion matrix have also been calculated (Figure.13-Figure14).

Table.5 F1 score, recall and precision of RF classification results.

	Class Name	F1	Recall	Precision
RF	Citrus Trees	89.86	91.70	88.10
	Cotton	92.36	94.20	90.60
	Maize	91.79	94.20	89.50
	Peanut	89.79	90.50	89.10
	Sunflower	87.63	85.20	90.20
	Watermelon	86.29	87.40	85.20
	Wheat	93.75	93.20	94.30
	Wheat-Cotton	86.69	85.80	87.60
	Wheat-Maize	79.90	78.00	81.90
	Wheat-Soybean	80.15	79.70	80.60
	Overall Accuracy : % 89.56			

Table.6 F1 score, recall and precision of SVM classification results.

	Class Name	F1	Recall	Precision
SVM	Citrus Trees	85.62	87.20	84.10
	Cotton	85.90	86.20	85.60
	Maize	87.08	85.70	88.50
	Peanut	84.18	87.50	81.10
	Sunflower	85.46	81.20	90.20
	Watermelon	85.91	88.80	83.20
	Wheat	87.75	88.20	87.30
	Wheat-Cotton	82.69	83.80	81.60
	Wheat-Maize	76.42	75.00	77.90
	Wheat-Soybean	80.54	77.70	83.60
	Overall Accuracy : % 86.14			

Table.7 F1 score, recall and precision of ANN classification results.

	Class Name	F1	Recall	Precision
ANN	Citrus Trees	84.94	85.80	84.10
	Cotton	86.62	88.20	85.10
	Maize	86.46	83.80	89.30
	Peanut	84.67	86.30	83.10
	Sunflower	85.24	84.20	86.30
	Watermelon	85.18	89.70	81.10
	Wheat	90.15	93.20	87.30
	Wheat-Cotton	81.28	82.50	80.10
	Wheat-Maize	77.00	77.10	76.90
	Wheat-Soybean	79.76	77.10	82.60
	Overall Accuracy : % 86.56			

Table.8 F1 score, recall and precision of XGBoost classification results

	Class Name	F1	Recall	Precision
XGBoost	Citrus Trees	90.95	90.40	91.50
	Cotton	93.92	95.60	92.30
	Maize	92.58	96.10	89.30
	Peanut	89.25	89.20	89.30
	Sunflower	88.60	88.80	88.40
	Watermelon	84.44	83.50	85.40
	Wheat	95.59	94.50	96.70
	Wheat-Cotton	85.66	87.60	83.80
	Wheat-Maize	82.43	79.40	85.70
	Wheat-Soybean	81.87	83.50	80.30
	Overall Accuracy : % 92.14			

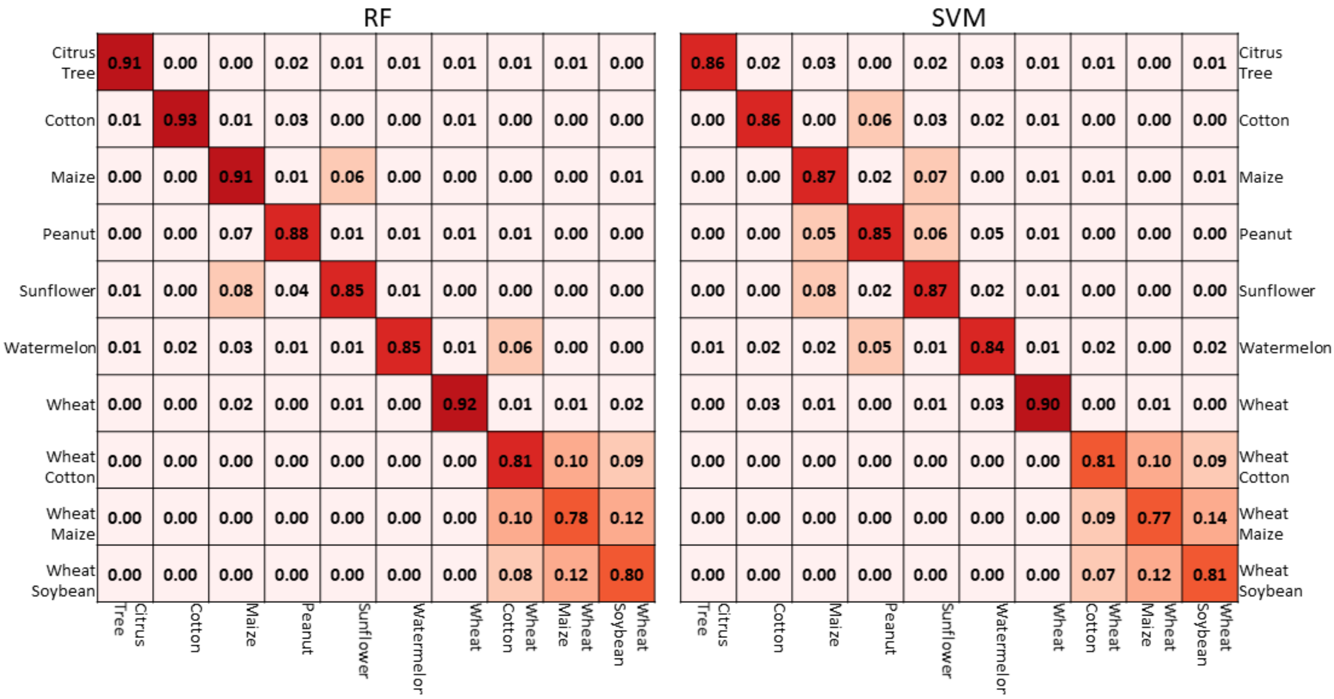


Figure.13 Confusion matrix for RF and SVM classification

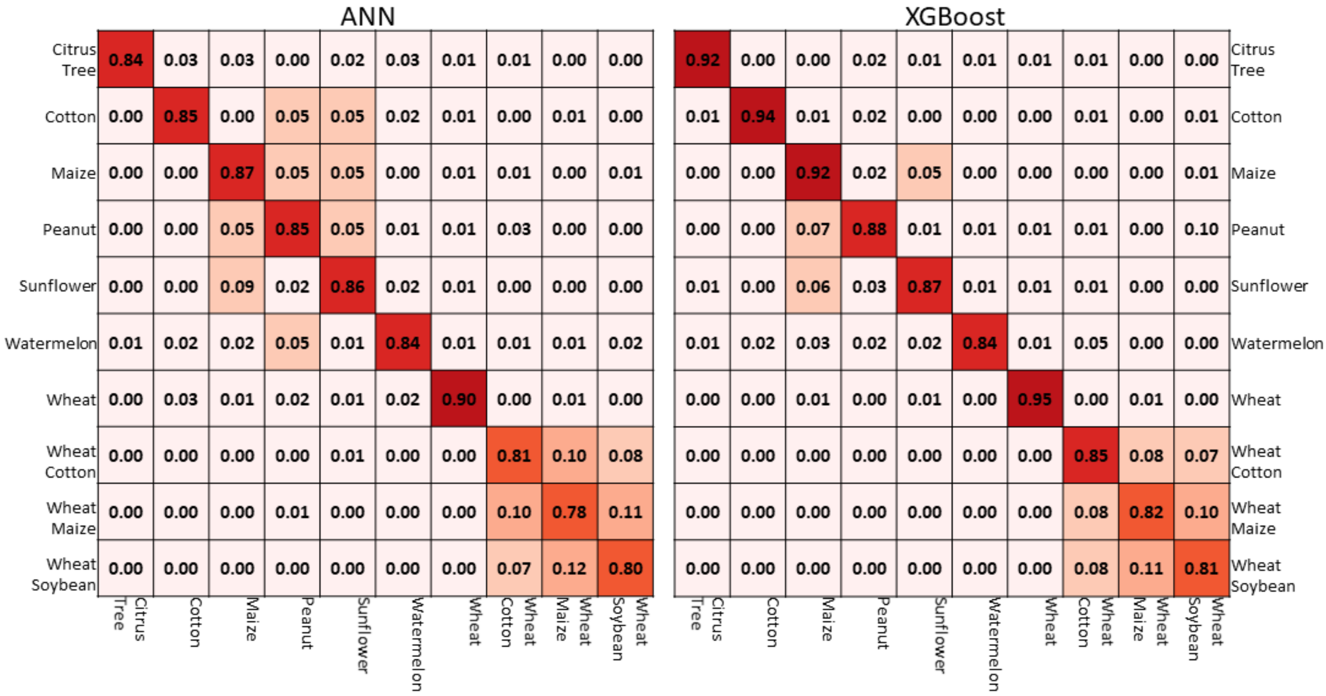


Figure.14 Confusion matrix for ANN and XGBoost classification.

As shown in Table 5-6-7-8, XGBoost showed the most accurate results OA (%92.14), followed by RF OA (%89.56), SVM OA (%86.14) and ANN OA (%85.48). SVM and ANN classifiers both in OA and all classes were only slightly different. The highest accuracy in all classifiers was for wheat (RF-F1-%93.75, SVM-F1-%87.75, ANN-F1% 90.15, XGBoost-F1-%95.59) cotton (RF-F1-%92.36, SVM-F1-%85.90, ANN-F1% 86.62, XGBoost-F1-%93.92) and maize (RF-F1-%91.79, SVM-F1-%87.08, ANN-F1% 86.46, XGBoost-F1-%92.58) classes and the lowest for the double crops, wheat-cotton (RF-F1-%86.59, SVM-F1-%82.69, ANN-F1% 81.28, XGBoost-F1-%85.66) wheat-maize (RF-F1-%79.90, SVM-F1-%76.42, ANN-F1% 77.00, XGBoost-F1-%82.43) wheat-soybean (RF-F1-%80.15, SVM-F1-%80.54, ANN-F1% 79.76, XGBoost-F1-%81.87) All four models exhibited an overall accuracy of approximately %84-89 for watermelon, peanut and sunflower classes. Citrus trees and wheat are distinguished from other crops because they have different phenological information.

As observed in the error matrix for all classifiers, there is some degree of confusion between maize and sunflower. This can be attributed to the fact that maize and sunflower exhibit similar characteristics in their NDVI curves. In the case of SVM and ANN classifiers, cotton is misclassified as peanut, and peanut is misclassified as maize. Additionally, sunflower and watermelon are a confused with peanuts to some extent. In the crop classes especially double crops, wheat-cotton, wheat-maize, and wheat-soybean, there was a notable degree of confusion between these classes across all algorithms. Particularly, the classes wheat-maize and wheat-soybean obtained low accuracy values for all machine learning algorithms. Especially the wheat-maize and wheat-soybean classes could not reach a high accuracy value for all classifiers. The reason for the low accuracy is that double crops have similar or identical sowing time, length of production period, and harvest time, their growth status at each stage are almost identical. In this study shows that RF and XGBoost algorithms perform better than ANN and SVM. It was also shown that crops with differing phenological periods achieved high performance.

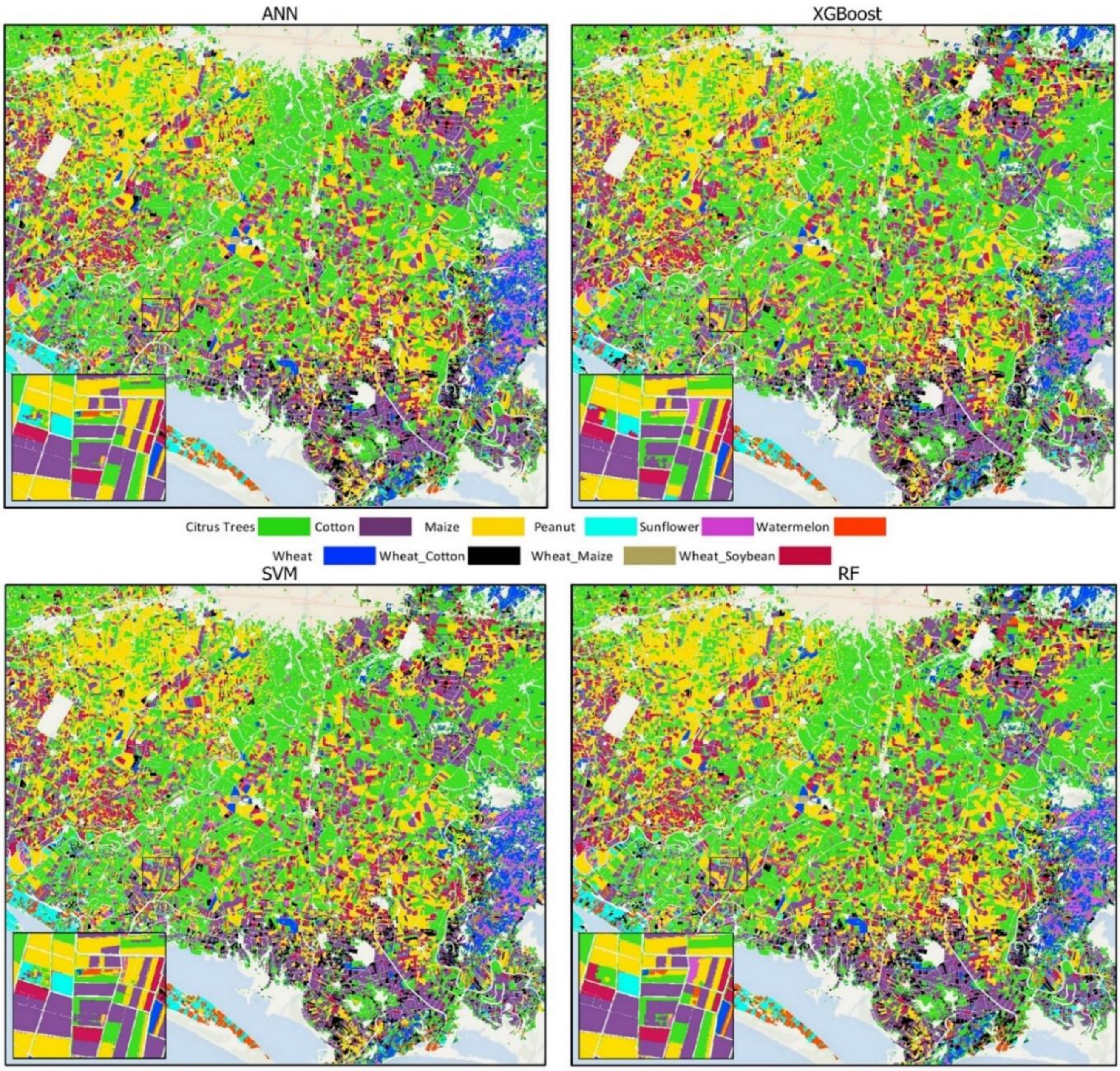


Figure.15 Crop classification results for all algorithm .

6. Conclusion

In this study, the performances of four machine learning algorithms, namely, random forest (RF), support vector machine (SVM), artificial neural network (ANN) extreme gradient boosting machine (XGBoost), and using a feature space that contained eleven temporal phases, with 10 spectral bands and EVI indices per each phase for the crop type classification of the Çukurova Plain in Adana were evaluated and compared (Figure 15).

The comparisons involved analyzing the impact of tuning parameters on classification results and assessing the classification accuracy achieved by each classifier with their respective optimal parameters. When comparing processing unit times, (64 GB RAM, Core i7-9750H CPU hardware) the fastest algorithm was XGBoost, (46min. 48sec.) followed by RF as the second-fastest (1h. 12 min. 24sec.) and ANN as the third-fastest (1h 23min. 12 sec.). SVM (1h 57min 12sec.) on the other hand, was notably slower in comparison to the other algorithms.

In this study using multi-temporal satellite imagery, the selection of image dates should be done carefully due to the dynamic structure of agricultural areas. In multi-temporal image classification, satellite image dates have a significant effect on classification accuracy. Therefore, the phenological characteristics of agricultural crops should be carefully studied, and appropriate dates should be determined by taking phenological characteristics into account.

As a result of this study, agricultural crop type classification using the bands and indices of Sentinel-2 image showed that the classification accuracy of agricultural crops with different phenological periods was high, while agricultural crops with close phenological periods could not reach high accuracy due to confusion with each other. For each of classifier performed high accuracy values for single crops and low accuracy values for double crops. In order to increase the classification accuracy, especially for double crops, Sentinel-1 radar images and indices generated from these images, as well as deep learning methods with different algorithms, will be used in the next study.

The fact that the images are freely available, that the 10-metre resolution bands are sufficient for agricultural applications in regions with certain sizes of agricultural parcels, and that they have a high temporal resolution of 5 days make Sentinel-2 data very valuable for the detection of agricultural crops. Reference data were used to train the models created in the machine learning algorithms. In order to obtain the reference data, it is necessary to go to the field and carry out a study or product information of agricultural parcels that can replace the reference data collected in the field.

To train the machine learning model employed in classification studies, reference (ground truth) data must be utilized. Another objective of this study is to generate ground truth data from farmer declaration parcels (FDP) for the classification study. Prior to using these farmer declaration parcels as ground truth data, preprocessing was conducted. Any parcels with errors or outliers in the FDP parcels were excluded, resulting in the creation of reference parcels to be employed in the classification process.

In study areas with large areas such as this study, collecting sufficient and homogeneous data from the land for each class for which the crop type is to be determined is costly and time-consuming, and is not sustainable. Due to this situation, the alternative of using FDP as ground truth data was used. By performing various processing steps (topological correction, editing, deletion, elimination of outliers about characteristics of EVI curves), ground truth data were generated from these parcels and used in the classification study. By transforming the FDP into ground truth data due to various processes, high accuracy was obtained from the classification result. It has also shown that it can be used as ground truth data in this and similar studies.

Conflicts of interest

The authors declare no conflicts of interest

References

1. Şimşek, F.F. (2023). Optik ve radar görüntüleri ile aşırı gradyan artırma algoritması kullanılarak tarımsal ürün desen tespiti. Geomatik Dergisi, 9(1),54–68 https://doi.org/10.29128/geomatik.1332997
2. Matton, N., Canto, G.S., Waldner, F., Valero, S., Morin, D., Inglada, J., Arias, M., Bontemps, S., Koetz, B., & Defourny, P. (2015). An automated method for annual cropland mapping along the season for various globally-distributed agro systems using high spatial and temporal resolution time series. Remote Sensing, 7 (10), 13208-13232.
3. Qiong, H., Wen-bin, W., Qian, S., Miao, L., Di, C., Qiang-yi, Y., & Hua-jun, T. (2017). How do temporal and spectral features matter in crop classification in Heilongjiang Province, China Journal of Integrative Agriculture, 16(2), 324–336.https://doi:10.1016/S2095 3119(15)61321-1
4. Zhang, C., Zhang, H., Du, J., & Zhang, L. (2018). Automated paddy rice extent extraction with time stacks of sentinel data: a case study in Jianghan plain, Hubei, China. 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), 1-6 https://doi:10.1109/AgroGeoinformatics.2018.8476119
5. Altun, M., & Turker, M. (Year). Integration of Sentinel-1 and Landsat-8 images for crop detection: The case study of Manisa, Turkey. Advanced Remote Sensing, 2(1), 23-33
6. Waldner, F., Canto, G.S., Defourny, P. (2015). Automated annual cropland mapping using

knowledge-based temporal features. ISPRS Journal of Photogrammetry and Remote Sensing,110:1-13. https://doi:10.1016/j.isprsjprs.2015.09.013
7. Csillik, O., Belgiu, M., Asner, G.P., & Kelly, M. (2016). Object-based time-constrained dynamic time warping classification of crops using sentinel-2. Remote Sensing, 11(10),1257. https://doi.org/10.3390/rs11101257
8. King, L.M., Adusei, B., Stehman, S., Potapov, P.V., Song, X., Krylov, A., Bella, C.M., Loveland, T.R., Johnson, D.M., & Hansen, M.C., (2017). A multi-resolution approach to national-scale cultivated area estimation of soybean. Remote Sensing of Environment, 195, 13-29. https://doi.org/10.1016/j.rse.2017.03.047
9. Wardlow, B.D., & Egbert, S.L. (2008). Large-area crop mapping using time-series MODIS 250m NDVI data: An assessment for the U.S. Central Great Plains. Remote Sensing of Environment 112, 1096-1116 https://doi.org/10.1016/j.rse.2007.07.019
10. Wang, S., Azzari, G., & Lobell, D. (2019). Crop type mapping without field-level labels: Random forest transfer and unsupervised clustering techniques. Remote Sensing of Environment, 222,303-317. https://doi.org/10.1016/j.rse.2018.12.026
11. Kang, J., Zhang, H., Yang, H., & Zhang, L. (2018). Support vector machine classification of crop lands using sentinel-2 imagery. 2018 7th International Conference on Agro-geoinformatics (Agro-geoinformatics), Hangzhou, China, pp. 1-6. https://doi: 10.1109/Agro-Geoinformatics.2018.8476101
12. Gomez, G., Shi, Z., Zhu, Y., Yang, X., & Hao, Y. (2020). Land use/cover classification in an arid desert-oasis mosaic landscape of China using remote sensed imagery: Performance assessment of four machine learning algorithms. Global Ecology and Conservation, 22, e00971. https://doi.org/10.1016/j.gecco.2020.e00971
13. Zhang, H., Kang, J., Xu, X., & Zhang, L. (2020). Accessing the temporal and spectral features in crop type mapping using multi-temporal Sentinel-2 imagery: A case study of Yi'an County, Heilongjiang province, China. Computer Electronic Agriculture, 176, 105618. https://doi.org/10.1016/j.compag.2020.105618
14. Li, Q., Tian, J., & Tian, Q. (2023). Deep learning application for crop classification via multi-temporal remote sensing images. Agriculture, 13(4),906.https://doi.org/10.3390/agriculture13040906
15. Vuolo, F., Neuwirth, M., Immitzer, M., Atzberger, C., & Ng, W. (2018). How much does multi-temporal Sentinel-2 data improve crop type classification? Int. J. Appl. Earth Obs. Geoinformation, 72, 122-130 https://doi.org/10.1016/j.jag.2018.06.007
16. Arvor, D., Jonathan, M., Simoes, M., Dubreuil, V., & Durieux, L. (2011). Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil. International Journal of Remote Sensing, 32(22), 7847-7871. https://doi.org/10.1080/01431161.2010.531783
17. Belgiu, M., & Csillik, O. (2018). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. Remote Sensing of Environment, 204, 509-523. https://doi.org/10.1016/j.rse.2017.10.005
18. Zheng, H., Du, P., Chen, J., Xia, J., Li, E., Xu, Z., Li, X., & Yokoya, N. (2017). Performance evaluation of downscaling sentinel-2 imagery for land use and land cover classification by spectral-spatial features. Remote Sensing, 9(12),1274. https://doi.org/10.3390/rs9121274
19. Müller, U. &, Wilm, U. (2017). Sen2Cor configure ration and user manual. Ref. S2-PDGS-MPC-L2A-SUM-V2.4, 1, 9-12
20. Müller, U., Wilm, U., Louis, J., Richter, R., Gascon, F., & Niezette, M. (2013). Sentinel-2 level 2a prototype processor: architecture, algorithms and first results. ESA Living Planet Symposium vol. 722, p. 98
21. Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat images for tropical forest monitoring. Remote Sensing of Enviroment, 118,83-94. https://doi.org/10.1016/j.rse.2011.10.028
22. Unel, F. B., Kusak, L., & Yakar, M. (2023). GeoValueIndex map of public property assets generating via Analytic Hierarchy Process and Geographic Information System for Mass Appraisal: GeoValueIndex. Aestimum, 82, 51-69. https://doi.org/10.3390/rs13153031
23. Morsy, S., & Hadi, M. (2022).Investigation of phenological stages of wheat plant using vegetation index. Mersin Photogrammetry Journal, 2 (1) , 24-28.
24. Kaya, Y., & Polat, N. (2020). Impact of land use/land cover on land surface temperature and its relationship with spectral indices in Dakahlia Governorate, Egypt . International Journal of Engineering and Geosciences, 7 (3) , 272-282. https://doi.org/10.26833/ijeg.978961
25. Maxwell, A.E., Warner, T.A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. International Journal of Remote Sensing, 39(9),2784-2817. https://doi.org/10.1080/01431161.2018.1433343

26. Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: a review. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 247-259 https://dx.doi.org/10.1016/j.isprsjprs.2010.11.001
27. 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. http://dx.doi.org/10.1080/01431160110040323
28. 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. https://doi.org/10.1080/01431160701395203
29. 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. http://dx.doi.org/10.1016/j.rse.2011.11.020.
30. Ustuner, M., & Sanli, F.B. (2021). Crop classification from multi-temporal PolSAR data with regularized greedy forest. Advanced Remote Sensing, 1(1), 10-15
31. Belgiu M., Dragut, L. (2016). Random forest in remote sensing: a review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114, 24-31.https://doi.org/10.1016/j.isprsjprs.2016.01.011
32. Avcı, C., Budak, M., Yağmur, N., Balçık, F. (2023). Comparison between random forest and support vector machine algorithms for LULC classification. International Journal of Engineering and Geosciences, 8(1), 1-10. https://doi.org/10.26833/ijeg.98760533.
33. Tatsumi, M., 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. https://doi.org/10.1016/j.compag.2015.05.01
34. 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. https://doi.org/10.1016/j.rse.2013.08.023
35. Kusak, L., Unel, F. B., Alptekin, A., Celik, M. O., & Yakar, M. (2021). Apriori association rule and K-means clustering algorithms for interpretation of pre-event landslide areas and landslide inventory mapping. Open Geosciences, 13(1), 1226-1244
36. Terzi Türk, S., & Balçık, F. (2023). Rastgele orman algoritması ve Sentinel-2 MSI ile fındık ekili alanların belirlenmesi: Piraziz Örneği. Geomatik, 8(2), 91-98. https://doi.org/10.29128/geomatik.1127925
37. Zhang, H. K., & Roy, D.P. (2014). 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. https://doi.org/10.1016/j.rse.2017.05.02
38. Chen, T.Q., & Guestrin, C. (2016). Xgboost: a scalable tree boosting system. proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, San Francisco, 13-17 August 2016, 785-794. https://doi.org/10.1145/2939672.2939785
39. Farid, D.M., Maruf, G.M., & Rahman, C .M. (2013). A new approach of boosting using decision tree classifier for classifying noisy data. 2013 International Conference on Informatics, Electronics and Vision (ICIEV), Dhaka, Bangladesh, 2013, pp. 1-4, https://dx.doi.org/10.1016/j.isprsjprs.2010.11.001