Mapping of Poplar Tree Growing Fields with Machine Learning Algorithms using Multi-Temporal Sentinel-2A Imagery

Muhammed Yusuf Ozturk (1), Ismail Colkesen (1)

[1] Department of Geomatics Engineering, Gebze Technical University, 41400, Kocaeli, Turkey
Email: m.ozturk2020@gtu.edu.tr; icolkesen@gtu.edu.tr

KEY WORDS: Machine Learning, Multi-Temporal, Supervised Classification, Poplar Trees, Sentinel-2A

ABSTRACT: The poplar trees used in peeling, packaging, furniture, fiber chip, cellulose industry and construction sector are one of the most significant wood supply sources of the countries. Monitoring the development stage of cultivated poplar trees, determination of their boundaries and mapping their fields in cheaper and more accurate ways plays an important role for the poplar tree growing sector. The main goal of this study is to map the cultivated hybrid Poplar (*P. deltoides*) fields in Akyazi district of Sakarya, Turkey using multi-temporal Sentinel-2A satellite imagery and different spectral band combinations. For this purpose, pixel-based supervised classification procedure was selected and three machine learning algorithms, namely, random forest (RF), support vector machines (SVMs) and Adaboost (AdaB) were applied to produce thematic map of the study area. In order to meet desired the goals of the study, three Sentinel-2A imagery from April, July and September, 2019 were used as a multi-temporal dataset consisted of three band combinations. In addition, classification results of multi-temporal datasets compared with single-dated datasets belonged to April for evaluate the effect of using multi-temporal imagery on classification accuracy. Overall accuracy and McNemar's test were used for accuracy assessment. According to classification results, overall accuracies of multi-temporal datasets became superior from single-dated datasets. Furthermore, McNemar's test results also affirmed there is significantly difference datasets formed by 20 m resolution bands and pan-sharpened bands of Sentinel-2A imagery between April datasets and multi-temporal datasets. One of used algorithms, AdaB, showed weaker classification performance with respect to RF an SVMs. Furthermore, according to results of F-score, poplar label class reached up to %99 at Dataset-6.

1. INTRODUCTION

Remote sensing technologies provide valuable information about the Earth's surface for the visualization, processing and managing of the limited natural resources. . Remotely sensed imagery have been commonly used as fundamental dataset for producing variety of thematic maps representing spatial locations and distributions the natural or man-made resources of the Earth's surface. These thematic maps are primarily utilized for forestry, agriculture, geology, natural resource, land use and land cover (LULC) detection and land management plans (Kavzoglu and Colkesen, 2013; Machala and Zejdová, 2014; Colkesen et al., 2016; Tonbul et al., 2020). Detection of cultivated lands and tree species from the multi-temporal and single-dated imageries using supervised or unsupervised image classification algorithms have been widely-used in many studies. The cropland mapping using multi-temporal imageries have been found to be effective method in terms of maps accuracy compared to single-dated thematic mapping methods (Gómez et al., 2016; Long et al., 2013; Belgiu and Csillik, 2018).

Poplar trees have important features compared to other species due to their fast-growing structures, their ability to be used as a basic raw material in many industrial areas and their high economic returns. There are more than 100 species of poplar trees, subspecies variability, and countless hybrids and clones located in the world. Above 100 species of poplar, subspecies variability and numerous hybrids exist within the worldwide (Tonbul et al., 2020). Determination of poplar cultivated fields and creating LULC maps with effectively, low-costly, rapidly and high accurately ways is significantly important for poplar tree industry. Thus, producing a LULC map using remote sensed imageries has been one of the most applied procedure. Several studies have been conducted to classify cultivated crops and croplands using imageries captured by Landsat, MODIS, SPOT or other satellites. The Sentinel-2, relatively new Earth observation satellite, have been recently used in many studies such as LULC classification and cropland monitoring due to its suitable spatial (i.e. 10 m and 20 m), spectral (i.e. 10 spectral bands) and high temporal (i.e. 5 days) resolutions.

The main purpose this study is to determine poplar trees cultivated lands by means of pixel-based classification of multi-temporal Sentinel-2A imagery using robust machine learning algorithms (i.e. RF, SVMs and AdaB). Classification results derived from multi-temporal datasets were compared with single dated dataset using overall accuracy, F-score and McNemar's test.

2. STUDY AREA AND DATASET

The study area covers approximately 250 km^2 agricultural lands located in Akyazi district of Sakarya province, located in northwest of Turkey (Fig. 1). Agriculture takes place one of the essential parts of region's economy, especially with the production of sugar beet, corn and hazelnut. Furthermore, weather conditions and soil characteristics of the study region provides convenient field for cultivating poplar trees. Study region compose of five LULC class such as poplar, urban, cultivated, non-cultivated, forest.

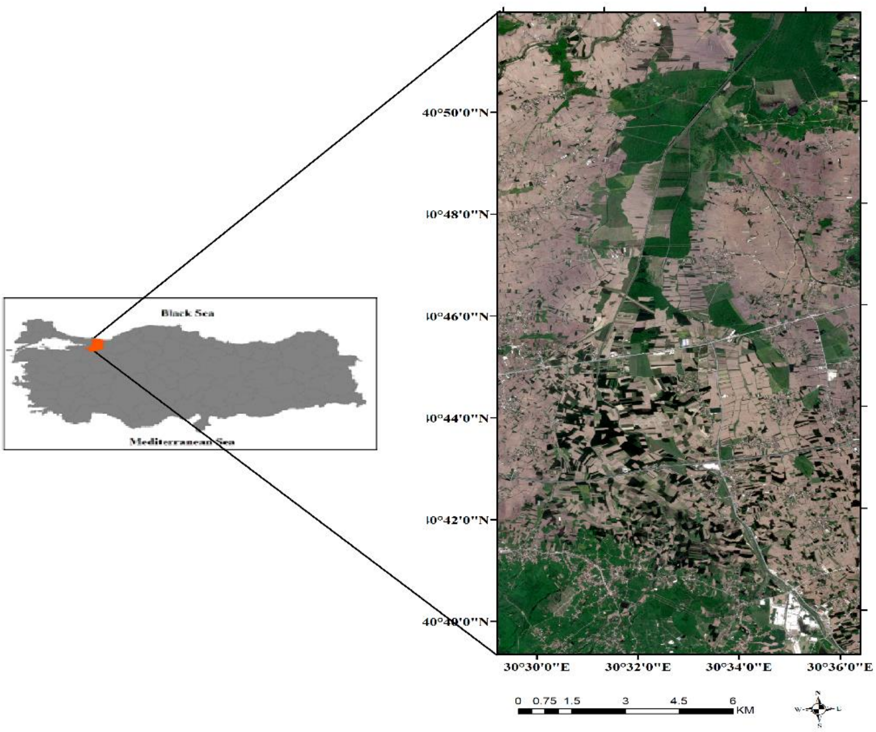

Figure 1. Location of study area

In this study multi-temporal satellite imagery captured by Sentinel-2A satellite on 30 April 2019, 29 July 2019 and 17 September 2019 were used. The Sentinel-2 imagery are Level-2A products that provide the bottom of atmosphere reflectance. To achieve the aim of study six different datasets were formed as follow: Dataset-1 includes four bands of April image at 10m spatial resolution, Dataset-2 includes six bands of April image at 20m spatial resolutions, Dataset-3 includes pan-sharpened 10 bands of April image. Dataset-4 comprises 12 bands (i.e. four bands of April, July and September images at 10m spatial resolution), Dataset-5 comprises 18 bands (i.e. six bands of April, July and September images at 20m spatial resolution and Dataset-6 comprises 30 bands (i.e. pan-sharpened 10 bands of April, July and September images). all datasets and their ingredients are given in Table 1.

Table 1. Sentinel-2 imagery properties and datasets created in this study

Date	Band	Name	Res(m)	Dataset-1	Dataset-2	Dataset-3	Dataset-4	Dataset-5	Dataset-6
April	2	Blue	10	✓	×	✓	✓	×	✓
	3	Green	10	✓	×	✓	✓	×	✓
	4	Red	10	✓	×	✓	✓	×	✓
	5	Red-Edge	20	×	✓	✓	×	✓	✓
	6	Red-Edge	20	×	✓	✓	×	✓	✓
	7	Red-Edge	20	×	✓	✓	×	✓	✓
	8	NIR	10	✓	×	✓	✓	×	✓
	8A	Narrow-NIR	20	×	✓	✓	×	✓	✓
	11	SWIR	20	×	✓	✓	×	✓	✓
	12	SWIR	20	×	×	×	×	✓	✓
July	2	Blue	10	×	×	×	✓	×	✓
	3	Green	10	×	×	×	✓	×	✓
	4	Red	10	×	×	×	✓	×	✓
	5	Red-Edge	20	×	×	×	×	✓	✓
	6	Red-Edge	20	×	×	×	×	✓	✓
	7	Red-Edge	20	×	×	×	×	✓	✓
	8	NIR	10	×	×	×	✓	×	✓
	8A	Narrow-NIR	20	×	×	×	×	✓	✓
	11	SWIR	20	×	×	×	×	✓	✓
	12	SWIR	20	×	×	×	×	✓	✓
September	2	Blue	10	×	×	×	✓	×	✓
	3	Green	10	×	×	×	✓	×	✓
	4	Red	10	×	×	×	✓	×	✓
	5	Red-Edge	20	×	×	×	×	✓	✓
	6	Red-Edge	20	×	×	×	×	✓	✓
	7	Red-Edge	20	×	×	×	×	✓	✓
	8	NIR	10	×	×	×	✓	×	✓
	8A	Narrow-NIR	20	×	×	×	×	✓	✓
	11	SWIR	20	×	×	×	×	✓	✓
	12	SWIR	20	×	×	×	×	✓	✓

3. METHODOLOGY

In this study, aforementioned six datasets including different combinations of spectral bands of multi-temporal Sentinel-2 imagery were classified using RF, SVMs and AdaB algorithms. Therefore, supervised pixel-based image classification steps including selection of training and test samples, classification of each datasets using machine learning algorithms, accuracy assessment and creation of thematic map were applied respectively.

3.1. Machine Learning Algorithms

In recent years, researchers have shown an increased interest in the use of machine learning algorithms for producing thematic maps in remote sensing (Petropoulos et al., 2012; Keshtkar et al., 2017; Feyisa et al., 2020; Abdi, 2020; Ge et al., 2020). In this study, three machine learning algorithms frequently preferred in supervised classification process due to their robustness namely, random forest, support vector machines and AdaBoost were utilized to produce LULC map of the study area.

RF, SVMs and AdaB that are frequently preferred due to their high performance in machine learning algorithms were evaluated in this study.

3.1.1. Random forest

Random Forest (RF), tree-based ensemble learning algorithm proposed by Breiman (2001), has been commonly used in the classification of remotely sensed imagery due to its ability to accurately determine class boundaries between the spectrally similar LULC classes (Kavzoglu et al., 2019). The algorithm uses bootstrapped samples as training dataset to build decision trees in the forest. Each tree in the ensemble model is trained using nearly 2/3 bootstrapped samples (i.e. in bag) and the rest (i.e. out of bag) are used to accuracy assessment of ensemble model. Majority vote of decision trees is utilized to prediction of class label of samples (Kavzoglu et al., 2018; Sahin et al., 2020).

3.1.2. Support vector machines

The support vector machines (SVMs), supervised non-parametric machine algorithm proposed by Cortes and Vapnik, 1995, have been commonly preferred to classification of remotely sensed data due to its classification performance (Colkesen, 2015; Mirończuk and Hościło, 2017; Saini and Ghosh, 2018). Initially, SVMs algorithm was developed for binary classification purposes but it was adapted to multiclass classification problem (Ergun et al., 2010). The main idea behind of binary classification with SVMs is that define of hyperplane which ideally splits different two classes using train data. The test data is used for confirm generalization capability of hyperplane (Kavzoglu and Colkesen, 2009). The multiclass classification using SVMs is conducted with Radial Basis Function kernel having two fundamental parameters namely penalty parameter C and kernel width γ. In this study, a mesh grid search procedure has been applied to determine optimum C and γ parameters (Saini and Ghosh, 2018).

3.1.3. Adaboost

Adaboost (AdaB) proposed by Freund and Schapire (1997) has been widely-used for the classification of remotely sensed data for several decade (Ramzi et al., 2014). The algorithm is mainly based on the iterative training procedure. At the start of each iteration, each pixel in the training data is associated with a new weight and a classifier is added to model end of iteration. The training data chosen for a classifier are determined depending on the performance of its previous classifier (Kavzoglu and Colkesen, 2013). In other word, misclassified pixels in previous iteration are assigned with higher weight value and accurately classified pixels are taken lower. In the next iteration, misclassified pixels are utilized more than accurately classified ones. So, the AdaB builds a next classifier that aim to true errors in the last iteration (Chen et al., 2017; Tonbul et al., 2020).

3.2. Accuracy Assessment

In the literature, many accuracy assessment techniques have been suggested to evaluate thematic map accuracy produced by the image classification operation. (Kavzoglu and Colkesen, 2013). In this study, to analysis of classification performance of the produced thematic maps, overall accuracy estimated from error matrix were used. Additionally, F-score measures, calculated from harmonic mean of precision Producer’s accuracy and User’s accuracy was also conducted to evaluate LULC class-based performance. Furthermore, the McNemar’s test was utilized to assess whether there was a statistically significant difference between the accuracies of the datasets belonged to April and the multi-temporal datasets. McNemar's test, the non-parametric test based on χ^2 distribution and constructed as a 2 by 2-dimensional matrix, is utilized to determine whether there is a statistically significant difference between the accuracies of two different classifiers. If computed test value is higher than χ^2 tabular value such as 3.84 at 95% confidence interval, it means that there is significantly difference between two classifiers (Kavzoglu and Colkesen, 2009)(Feyisa et al., 2020).

4. RESULTS

In order to determinate boundaries of cultivated poplar trees, pixel-based image classification was conducted using machine learning algorithms (i.e. RF, SVMs and AdaB). Six different datasets were created and used for LULC classification process. All image classification steps and accuracy assessment analysis were carried out using *R* software. In order to apply supervised classification procedure, 700 pixels were chosen for each LULC classes as training samples to build machine learning models and 300 pixels for per LULC classes were selected as validation (i.e. test) dataset to conduct accuracy assessment procedure. Overall accuracies (OA) and F-score measures of each dataset and pairwise McNemar's test results between April datasets and multi-temporal datasets were given in Table 2.

Table 2. Classification performance of algorithms for each dataset

LULC label-class	Random Forest						Support Vector Machines						AdaBoost					
	F-score (%)						F-score (%)						F-score (%)					
	Dat-1	Dat-2	Dat-3	Dat-4	Dat-5	Dat-6	Dat-1	Dat-2	Dat-3	Dat-4	Dat-5	Dat-6	Dat-1	Dat-2	Dat-3	Dat-4	Dat-5	Dat-6
Cultivated	0.86	0.88	0.80	0.89	0.87	0.90	0.89	0.91	0.81	0.86	0.87	0.87	0.86	0.89	0.78	0.89	0.86	0.88
Forest	0.85	0.88	0.84	0.92	0.86	0.91	0.91	0.91	0.87	0.91	0.86	0.89	0.84	0.89	0.84	0.91	0.88	0.90
Non-cultivated	0.91	0.90	0.90	0.94	0.95	0.98	0.92	0.92	0.92	0.96	0.97	0.96	0.91	0.89	0.89	0.92	0.93	0.95
Poplar	0.92	0.93	0.94	0.98	0.93	**0.99**	0.94	0.91	0.94	0.95	0.97	**0.98**	0.92	0.90	0.93	0.96	0.94	**0.97**
Urban	0.91	0.94	0.90	0.93	0.97	0.98	0.91	0.98	0.92	0.95	0.97	0.99	0.92	0.94	0.89	0.93	0.94	0.97
OA (%)	89.1	90.7	87.7	93.3	91.6	94.9	91.7	92.4	89.5	92.9	93.1	93.9	89.1	90.1	86.5	92.3	90.4	93.7
McNemar's	1.21		20.10		14.02		0.64		8.01		5.02		0.57		18.86		8.02	

Results clearly indicated that the highest OA values were estimated with use of multi-temporal datasets (i.e. Dataset-4, -5 and -6) compared to single-date (i.e. April) dataset in all cases considered in this study.

When 20 m resolution bands were utilized with different classifiers, difference of OAs increased up to about %6. On the other hand, when 10m resolution imageries were classified using different algorithms, the difference between the OA values are limited. When the results of the single-date datasets were analyzed, the lowest accuracy values were estimated with the use of 20m bands of April dataset (Dataset-2), whereas the highest accuracy values were estimated with the use of pan-sharpened bands of April dataset (Dataset-3). On the other hand, the results of the multi-temporal datasets showed that the lowest accuracy values were for Dataset-4, whereas the highest accuracy values were estimated for Dataset-6 for all classification algorithms used in this study. This could be result of the use of having increased spectral and spatial information of pan-sharpened datasets for both April and multi-temporal imagery. In addition, the classification performances of the algorithms were also assessed for the datasets used in this study. The classification results indicated that the SVMs algorithms showed superior classification performances of poplar label-class compared to that of the RF and AdaB for April datasets. On the other hand, for all multitemporal datasets, the classification of poplar label-class performances with RF were found to be superior compared the others. These findings clearly indicated that while, SVMs algorithm successfully determined class boundaries in the case of limited spectral information (i.e. only use of 10 bands of April dataset), the RF algorithms showed superior performance in the case of more spectral information is available.

McNemar's test was conducted to find out whether there is a statistically significant difference between the April datasets with the same band combination and the multi-temporal datasets. McNemar's test results showed that there is no significantly difference between dataset-1 and dataset-4 instead of use different algorithms. However, using of multi-temporal bands such as 20 m resolution dataset-5 and pan-sharpened dataset-6 were increased the classification accuracy.

Thematic maps given in Fig. 2 were created using all algorithms for visually analysis the classification results of each dataset. It is clearly shown that poplar trees are clustered within under-middle and northeast of study region. It is obviously notable that although 20 m resolution bands belonged to April (Dataset-2) was classified with various algorithms, urban label class still tend to misclassification. On the other hand, using the multi-temporal data set (Dataset-5) with a resolution of 20 m, it can be seen that the problem of misclassification of the urban label class had been overcome. In addition, when April datasets were used, all algorithms misclassified the growing poplar trees as associating with forest label class. Moreover, boundaries of cultivated poplar trees could not exactly determine with all April dataset due to limitation of spectral information. However, the classification of poplar trees using multi-temporal datasets, especially when having more spectral information such as pan-sharpened imageries (Dataset-6), gave better results.

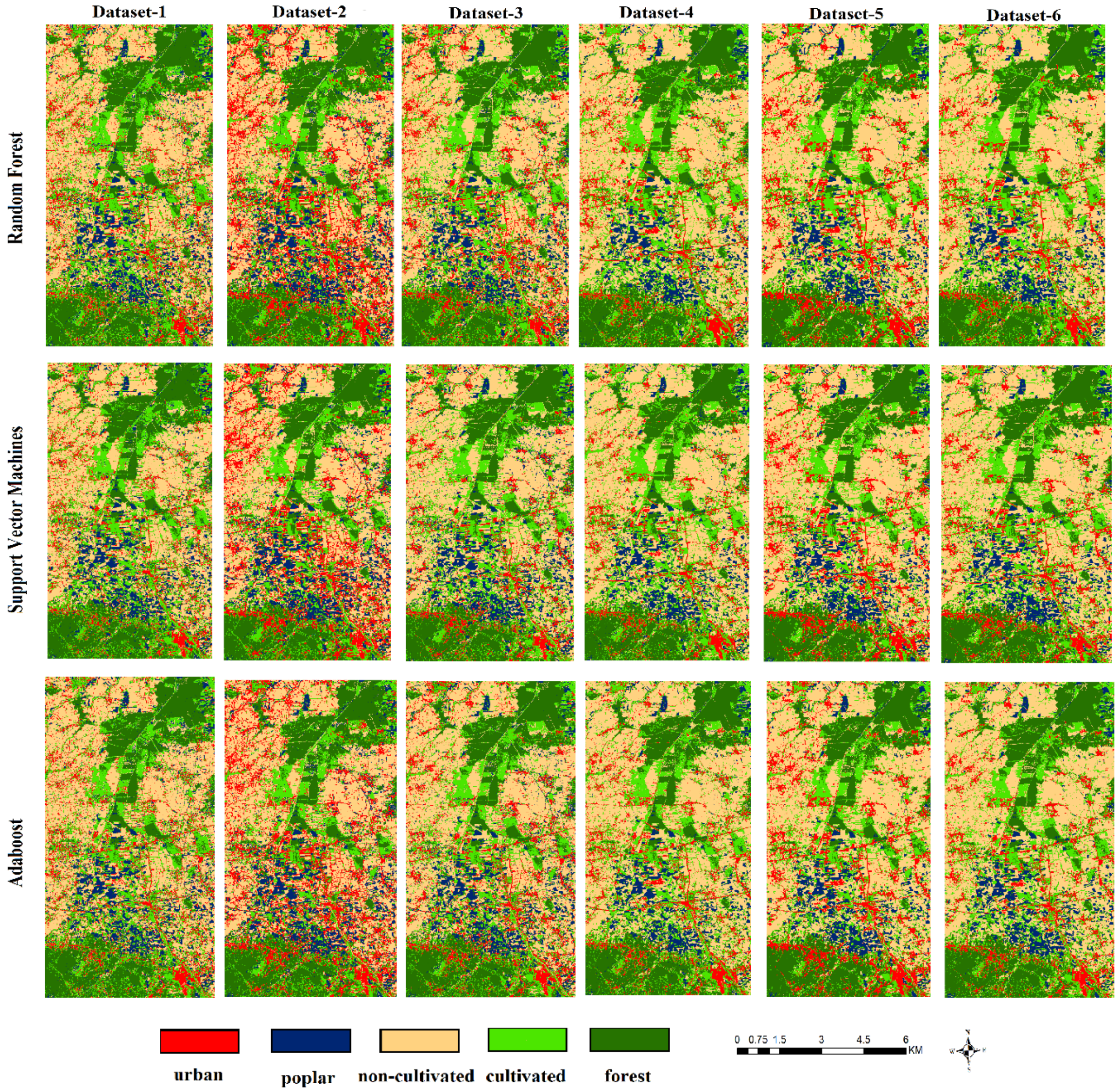

Figure 2. Visualization results created by all machine learning algorithms of each dataset

5. CONCLUSIONS

The goal of this study, determinate the boundaries of poplar trees within chosen study region with supervised pixel-based image classification using some machine machine learning algorithms such as RF, SVMs and AdaB. To achieve this purpose, six different datasets with three different band combination were created using single-dated and multi-temporal Sentinel-2A imageries for evaluate the effeteness of using multitemporal bands on classification performance instead of using only single-dated bands. The accuracy assessments were conducted using overall accuracy and McNemar’s test. LULC classification results confirmed that higher overall accuracy was achieved using multi-temporal imageries, resulting in more visually satisfying and accurate. McNemar's test confirmed that there was no significant change in the results of the classification of the 10 m resolution bands (Dataset-1 and Dataset-4) even if the multi-temporal image was used and more statistically significant results were obtained using all algorithms with the classification of single-dated and multi-temporal 20 m resolution and pan-sharpened imageries. Therefore, we can conclude that red edge bands and SWIR bands have a very important place in LULC classification. According to the comparison of learning algorithms, RF and SVMs had always outperformed from AdaB. Therefore, RF and SVMs, which are among the most preferred algorithms for image classification, will continue to be used due to their robust and accurate performance. According to F-scores measures, the highest classification accuracy of the poplar class label was achieved by using Dataset-6 and RF algorithm. Thus, Dataset-6 was used for calculate of cultivated poplar tree fields and was approximately 17 km^2 cultivated poplar tree field was determined.

6. REFERENCES

Breiman L., (2001), Random Forests. Mach. Learn., 45(1), 5–32, doi:10.1023/A:1010933404324

Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, *57*(1), 1–20. https://doi.org/10.1080/15481603.2019.1650447

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*(September 2017), 509–523. https://doi.org/10.1016/j.rse.2017.10.005

Chen, Y., Dou, P., Yang, X. (2017). Improving Land Use/Cover Classification with a Multiple Classifier System Using AdaBoost Integration Technique. *Remote Sensing*, *9*(10), 1055. https://doi.org/10.3390/rs9101055

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Ergun, B., Kavzoglu, T., Colkesen, I., Sahin, C. (2010). Data filtering with support vector machines in geometric camera calibration. *Optics Express*, *18*(3), 1927. https://doi.org/10.1364/OE.18.001927

Feyisa, G. L., Palao, L. K., Nelson, A., Gumma, M. K., Paliwal, A., Win, K. T., Nge, K. H., Johnson, D. E. (2020). Characterizing and mapping cropping patterns in a complex agro-ecosystem: An iterative participatory mapping procedure using machine learning algorithms and MODIS vegetation indices. *Computers and Electronics in Agriculture*, *175*, 105595. https://doi.org/10.1016/j.compag.2020.105595

Freund, Y., Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/jcss.1997.1504

Ge, 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

Gómez, C., White, J. C., Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 116*, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Kavzoglu, T., Colkesen, I. (2009). A kernel functions analysis for support vector machines for land cover classification. *International Journal of Applied Earth Observation and Geoinformation, 11*(5), 352–359. https://doi.org/10.1016/j.jag.2009.06.002

Kavzoglu, T., Colkesen, I. (2013). An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping. *International Journal of Remote Sensing, 34*(12), 4224–4241. https://doi.org/10.1080/01431161.2013.774099

Kavzoglu, T, Colkesen, I., Tonbul, H. (2019). *Agricultural Crop Type Mapping Using Object-Based Image*. *Acrs*, 1–8.

Kavzoglu, Taskin, Tonbul, H., Colkesen, I. (2018). Dimensionality reduction for hyperspectral images to improve object-based image classification using feature selection and principal components analysis. *Proceedings - 39th Asian Conference on Remote Sensing: Remote Sensing Enabling Prosperity, ACRS 2018, 4*(October), 2314–2318.

Keshtkar, H., Voigt, W., Alizadeh, E. (2017). Land-cover classification and analysis of change using machine-learning classifiers and multi-temporal remote sensing imagery. *Arabian Journal of Geosciences, 10*(6), 154. https://doi.org/10.1007/s12517-017-2899-y

Long, J. A., Lawrence, R. L., Greenwood, M. C., Marshall, L., Miller, P. R. (2013). Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest. *GIScience and Remote Sensing*, *50*(4), 418–436. https://doi.org/10.1080/15481603.2013.817150

Machala, M., Zejdová, L. (2014). Forest Mapping Through Object-based Image Analysis of Multispectral and LiDAR Aerial Data. *European Journal of Remote Sensing*, *47*(1), 117–131. https://doi.org/10.5721/EuJRS20144708

Mirończuk, A., Hościło, A. (2017). Mapping tree cover with Sentinel-2 data using the Support Vector Machine (SVM). *Geoinformation Issues*, *9*(1), 27–38. https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Petropoulos, G. P., Arvanitis, K., Sigrimis, N. (2012). Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping. *Expert Systems with Applications*, *39*(3), 3800–3809. https://doi.org/10.1016/j.eswa.2011.09.083

Ramzi, P., Samadzadegan, F., Reinartz, P. (2014). Classification of hyperspectral data using an AdaBoostSVM technique applied on band clusters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *7*(6), 2066–2079. https://doi.org/10.1109/JSTARS.2013.2292901

Sahin, E. K., Colkesen, I., Kavzoglu, T. (2020). A comparative assessment of canonical correlation forest, random forest, rotation forest and logistic regression methods for landslide susceptibility mapping. *Geocarto International*, *35*(4), 341–363. https://doi.org/10.1080/10106049.2018.1516248

Saini, R., Ghosh, S. K. (2018). Crop classification on single date sentinel-2 imagery using random forest and suppor vector machine. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–5*(November), 683–688. https://doi.org/10.5194/isprs-archives-XLII-5-683-2018

Tonbul, H., Colkesen, I., Kavzoglu, T. (2020). Classification of poplar trees with object-based ensemble learning algorithms using Sentinel-2A imagery. *Journal of Geodetic Science, 10*(1), 14–22. https://doi.org/10.1515/jogs-2020-0003