EVALUATION OF MACHINE LEARNING BASED ALGORITHMS FOR DETERMINING LAND COVER CHANGE WITH MULTI-TEMPORAL SENTINEL-2 SATELLITE IMAGES

Yaren Ketenci[1], Ugur Alganci[1] and Dursun Zafer Seker[1]
[1]Istanbul Technical University, Geomatics Engineering Department, ITU Ayazaga Campus, Sariyer, Istanbul, Turkey
Email: ketenci15@itu.edu.tr, alganci@itu.edu.tr, seker@itu.edu.tr

KEY WORDS: Land cover, Classification and Regression Trees, Random Forest, Google Earth Engine, Change Detection

ABSTRACT: Land cover maps created with remote sensing images are used for observation in many areas such as urban planning, land resources management, and biodiversity changes. The study of land cover change has a very important place in order to properly analyze the rapidly developing urbanization and to make the right city planning due to the increasing population and the migration wave that has surrounded the whole world. This study investigates the land cover change in the Istanbul province of Turkey, which receives the most irregular and continuous migration. The study was conducted over a five-year period between 2017 and 2021 with the use of Sentinel-2A satellite images. The classification was performed by use of two machine learning algorithms, which are Classification and Regression Trees and Random Forest on Google Earth Engine. The first level land cover class definitions of the CORINE legend were used through analysis. A comparative evaluation of two algorithms reported the higher performance of the random forest algorithm in terms of class-based accuracy, precision, recall, and F1-Score metrics. When the land cover changes are examined, it was determined that the urban land cover class faced a significant increase between 2017 and 2021 at the expense of a reduction in forest class. The biggest reason for this change can be related to the completion of the 3rd Airport project in this period and the construction of new settlements and transportation networks in Istanbul. In addition, water and agricultural land classes showed fluctuations over years, which are related to seasonal variations and cultivation patterns and are considered temporal changes. These results revealed that the use of machine learning algorithms on multi-temporal satellite images provides annual and periodical determination of land cover changes and can be used as a guide for further planning activities.

1. INTRODUCTION

Along with the developing technology, many changes and developments have occurred in the studies on satellite images, as in many areas. Today, processes such as land cover classification, which is an important part of the remote sensing discipline, have gained great speed and accuracy thanks to machine learning.

Remote sensing is the discipline of recording and examining the earth without physically connecting. Various features of the earth can be detected by remote sensing. Machine learning, on the other hand, is the process of learning without direct instructions by a computer using mathematical models, which is considered a subset of artificial intelligence. Many image processing techniques can now be realized with machine learning (Jamali, 2019, Yu et al., 2004). Within the scope of this study, land cover change with machine learning will be examined by using multispectral Sentinel-2 satellite images. Land cover maps created with remote sensing images are used for observation in many areas such as urban planning, land resources management, and biodiversity changes. The study of land cover change has a very important place to properly analyze the rapidly developing urbanization and to make the right city planning due to the increasing population and the migration wave that has surrounded the whole world (Bégué et al., 2018, Rawat et al., 2015).

This study, it is aimed to draw meaningful and interpretable results by using land use classification to examine the land cover change. Land use classification is the process of separating land cover according to its use by areas of economic activity. While monitoring the land cover change in the study, 5 years (2017-2021) Sentinel-2 images belonging to the city of Istanbul, Turkey were used and land use classification was performed on these images by machine learning algorithms. In this context, the used algorithms were Classification and Regression Trees and Random Forest. After all these processes, besides interpreting the land cover change, the accuracy of the algorithms was tested and interpreted by using the accuracy assessment procedure.

1.1 Related Work

In the study conducted by Abdi (2019), 8 land cover and land use were classified with RF, SVM, XgBoost and DL algorithms by using Sentinel-2 satellite images belonging to four seasons of the Uppsala district in Sweden. 70% of training data and 30% of test data were used for each class. When the classification results were compared, it was determined that the highest overall accuracy belonged to the SVM (0.758 ± 0.017) algorithm. The general accuracies of the other algorithms were determined as 0.751 ± 0.017 for XgBoos, 0.739 ± 0.018 for RF and 0.733 ± 0.0023 for DL, respectively.

In the study conducted by Ettehadi Ousgei et al. (2019), the separation of residential areas and bare land areas from each other was emphasized by using Sentinel-2A satellite images. Istanbul was chosen as the study area, and Ankara and Konya were selected as independent test regions. While making LCU (Land Cover Use) classification, a multiple index study was carried out in order to more accurately distinguish between bare area and residential area. Accordingly, it was determined that the combination of multiple indexes consisting of the Normalized Difference Tillage Index (NDTI), The Red-Edge-Based Normalized Vegetation Index (NDVIre), and The Modified Normalized Difference Water Index (MNDWI) gave the most accurate result. These index combinations used produced an excellent result with 93% accuracy and 0.91 kappa value for all LCU classes.

In the study conducted by Alganci (2019), changes in urbanization and land cover use in Istanbul due to population growth and migration were observed. While observing this change, Landsat 8 satellite images were used and a change analysis was made on 5-year images between 2013-2017. While making classification, pixel-based SVM algorithm and object-based GEOBIA algorithms were used and the results were compared. In addition, the images to be studied and algorithms to be applied are primarily built-in index (BUI) and principal component analysis (PCA) methods, and size reduction is performed. As a result of the comparison, the highest accuracy was obtained with the classification result applied with the GEOBIA algorithm to the data set, which was reduced by the BUI method, and overall accuracy of 91.6% and a kappa value of 0.91 were obtained. In addition, as a result of this study, it was concluded that the BUI data set gave more reliable and consistent results for urban classifications compared to the PCA data set.

In the study by McCarty et al. (2020), land use and land cover classification was performed with three machine learning techniques using Sentinel-2 satellite imagery of the Rhine-Ruhr metropolitan area and the results were compared. As a result of the study, it was concluded that the highest overall accuracy rate was the Light Gradient Boosted Machine (0.653). The overall accuracy of the SVM algorithm was determined to be 0.642 and the Random Forest algorithm to be 0.594, and it was concluded that the Light Gradient Boosted Machine algorithm performed 25% faster than these two algorithms.

In the study by Talukdar et al. (2020), a Landsat 8 OLI satellite image of the Ganga river and its surroundings in India was used. Land cover and land use classification was performed using Random Forest (RF), Support Vector Machine (SVM), Artificial Neural Network (ANN), Fuzzy ARTMAP, Spectral Angle Mapper (SAM) and Mahalanobis Distance (MD) algorithms and the results were compared. As a result of the study, it was concluded that the Kappa coefficient results were very close to each other, but the algorithm with the highest accuracy was the RF algorithm with 0.89, and the algorithm with the lowest accuracy was the MD algorithm with an overall accuracy of 0.82.

In the study published by Aghlmand et al. (2021), land use maps were produced with Google Earth Engine. While creating the land use maps, 5 different indexes including NDVI, EVI, NDWI, NDBI, UI indices, and 19 different data combinations were used. The reason for this is to obtain the highest accuracy land use map as a result of the study. Overall accuracy and kappa values were calculated for these 19 combinations classified and the results were compared. In addition, in this study, it was stated that it took less than 60 minutes to classify and calculate the combinations and give results, thus highlighting how attractive the use of GEE is today.

2. STUDY AREA AND DATA

It is thought that Istanbul city of Turkey will be the city where the 5-year land cover change will be observed the most due to the fact that it attracts tens of thousands of people every year due to its education and working opportunities, thus hosting several construction projects including housing and transportation facilities. The study area is illustrated in Figure 1.

In this study, Multispectral Sentinel-2 satellite images will be used. Sentinel-2 includes a twin constellation of satellites Sentinel-2A and Sentinel-2B. They acquire optical images with spatial resolution between 10 and 60 meters over terrestrial and coastal waters. Sentinel-2 satellites are suitable for a variety of applications such as agricultural monitoring, land cover classification, emergency management, natural resources, and water quality monitoring. These satellites, which provide a high frequency revisit feature; provide geographical information production at local, national, and international scales. The study area can be covered with four Sentinel 2 images and acquisition dates for each year are provided in Table 1.

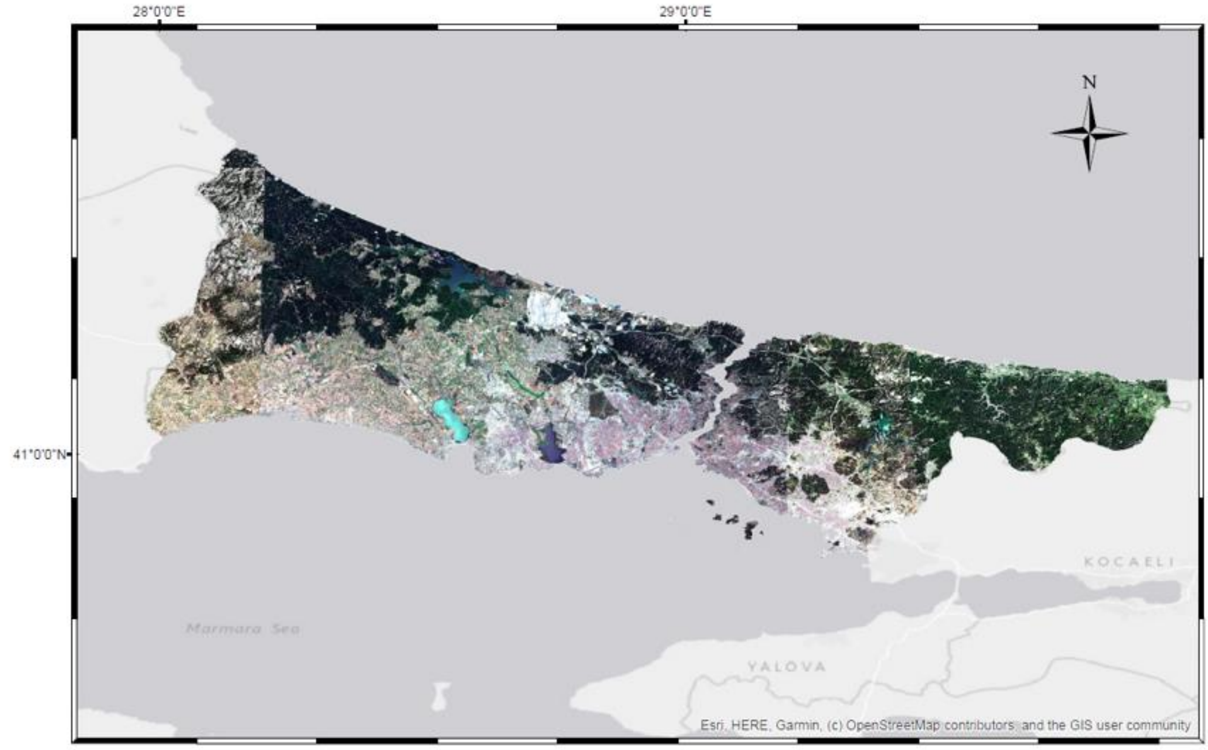

Figure 1. Sentinel 2 satellite image composite of Istanbul dated 2021.

Table 1. Sentinel 2 dataset image acquisition dates.

	2017	**2018**	**2019**	**2020**	**2021**
I	12.07.2017	12.06.2018	2.07.2019	31.07.2020	5.08.2021
II	22.07.2017	17.07.2018	26.08.2019	1.07.2020	5.08.2021
III	12.07.2017	17.07.2018	1.08.2019	1.07.2020	5.08.2021
IV	29.06.2017	4.07.2018	29.07.2019	7.08.2020	2.08.2021

3. METHODOLOGY

3.1 Pre-processing

All preprocessing steps for satellite images were conducted inside Google Earth Engine (GEE) platform. In order to work within the provincial borders of Istanbul, four satellite images covering Istanbul were selected, then these images were mosaiced and cut according to the provincial borders of Istanbul. While selecting these 4 satellite images for each year, care was taken to take the images within the same season and within a 3- months period. Thus, when interpreting the land cover change, the effects of the seasonal differences on the results of the study were avoided.

While selecting the satellite image with GEE, features such as filtering data, choosing a date range, geographical area, and cloud percentage can be selected, so that the desired satellite image of the study area can be easily accessed. In addition, the cloud ratio was considered while selecting the images, and the images with the lowest cloud ratio were preferred.

3.2 Machine Learning Based Classification

This study used two machine learning algorithms that are random forest (RF) and Classification and Regression Trees (CART) for image classification. According to the study of Schonlau et al. (2020), random forest, which is one of the best performing algorithms suitable for large data types, is a community tree-based learning algorithm formed from the trees that form the basis of the algorithm. In other words, RF is an improved version of bagging that creates a broad correlation of uncorrelated trees and then averages them (Breimann,2001). It averages estimates from individual trees using different bootstrap samples instead of the original sample. This tree-based model is a model in which the data set is iteratively divided into two groups under certain criteria until a specified condition is met. Each tree in RF gives a class prediction, and the class with the most votes becomes the model's prediction. (Qin et al., 2019). In the bagging algorithm, multiple preloaded datasets are generated from the original training dataset to train a classifier and a training dataset is assigned to each tree. The trees created are independent of each other and the largest vote is taken as the basis for the prediction (Shetty et al., 2021) Random decision forests that adapt easily even to nonlinear situations tend to predict linearity assumption better than linear regression.

Another decision tree algorithm, the CART algorithm, is a term coined by Leo Breiman in 1984. Its purpose is to first base data properties on simple rules, and then learn these rules, thereby creating a model that predicts the value of a variable. This algorithm can be used for both classification and regression purposes. Its basic principle is to divide the relevant group into two more homogeneous subgroups at each step. In other words, each branch grows by dividing into two sub-branches. The basic idea in the CART algorithm, which provides the opportunity to work with all data types, is to ensure that the units are separated to form homogeneous classes by binary selection while making decisions. This algorithm can use continuous and categorical data. The working method changes according to the variable, and it is called "classification trees" if the variable is in a categorical structure, and "regression trees" if it is continuous. Decision trees are formed with branches that develop according to the yes-no answer to each experiment starting from the root node. These decision trees use only one input variable at each step. If there is more than one input variable with the same value, selection will be performed randomly (Breimann et al., 1984). Trees can explain a single response variable with one or more explanatory variables. These response variables are usually either classification trees or regression trees, and explanatory variables are categorical or numeric. Trees are formed by repeatedly dividing data defined by a simple rule into a single explanatory variable. The data are divided into two groups that are as homogeneous as possible in each compartment. Then, the division process is applied to the two separated groups separately. The purpose of these procedures is to continuously divide the response into the most homogeneous groups possible. (De'ath et al., 2000).

Both classifiers are supervised, thus requiring training samples for class separation. For this purpose, training polygons for 5 main land cover classes, which are agricultural lands, bare lands, vegetation, urban areas, and water bodies, were created in the GIS environment and vector files were transferred to GEE for classification. During this process, care was taken to include training data in every region where the urban and agricultural classes are intertwined, so it is aimed to obtain results with higher accuracy in the distinction between these two classes.

3.3 Accuracy Assessment

In order to analyze the accuracy of the classification results, control points were determined for each year by taking the satellite image of the study area as a reference. Control points were selected for each class on the satellite image, considering the surface area, and then it was determined which land cover class these control points correspond to in the classified images (Figure 2). In this way, an accuracy assessment analysis of the classified images of each year was performed for both algorithms.

Precision, recall, F1 score, and overall accuracy values were calculated over the error matrices created with the help of control points. The precision determines the percentage of points classified as positive (Equation 1), and recall determines the percentage of positives that are true positives (Equation 2). F1 score is the harmonic mean of precision and recall (Equation 3). Overall accuracy measures the percentage of correctly classified points.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \qquad (1)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \qquad (2)$$

$$\text{F1 Score} = 2\text{x}\frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \qquad (3)$$

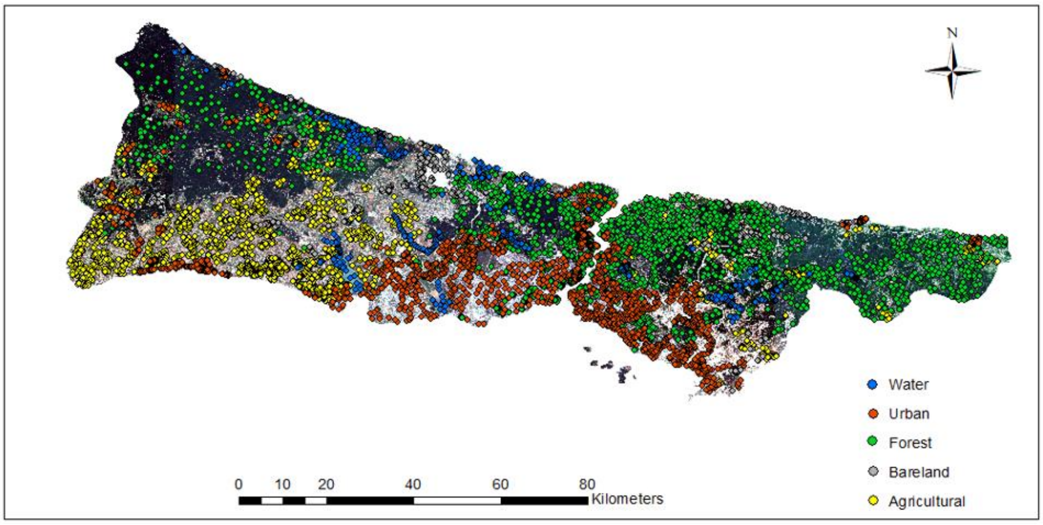

Figure 2. Distribution of the control points over the study area with visualization of class distribution.

4. RESULTS AND DISCUSSION

The visuals of the classification results are provided in Figure 3 and accuracy assessment results are given in Table 2 and Table 3. When these results are interpreted, it can be asserted that the general accuracy value is higher for the RF algorithm. While the overall accuracy is quite high in both algorithms in water bodies and vegetation classes, the results of the RF algorithm are more successful in other land cover classes. In addition, higher accuracy values were obtained with the RF algorithm in determining the land cover separation of areas that are considered bare land, such as dirt roads and small settlements, where agricultural land and urban class are intertwined.

If the algorithms are interpreted in general, it has been observed that RF presents the distinction between classes in a much more observable way and makes the distinction between city-vegetation, barren-cit,y and city-agriculture with high success. Despite the high accuracies obtained with the CART algorithm, it can be said that the RF algorithm is generally more successful in terms of land cover separation.

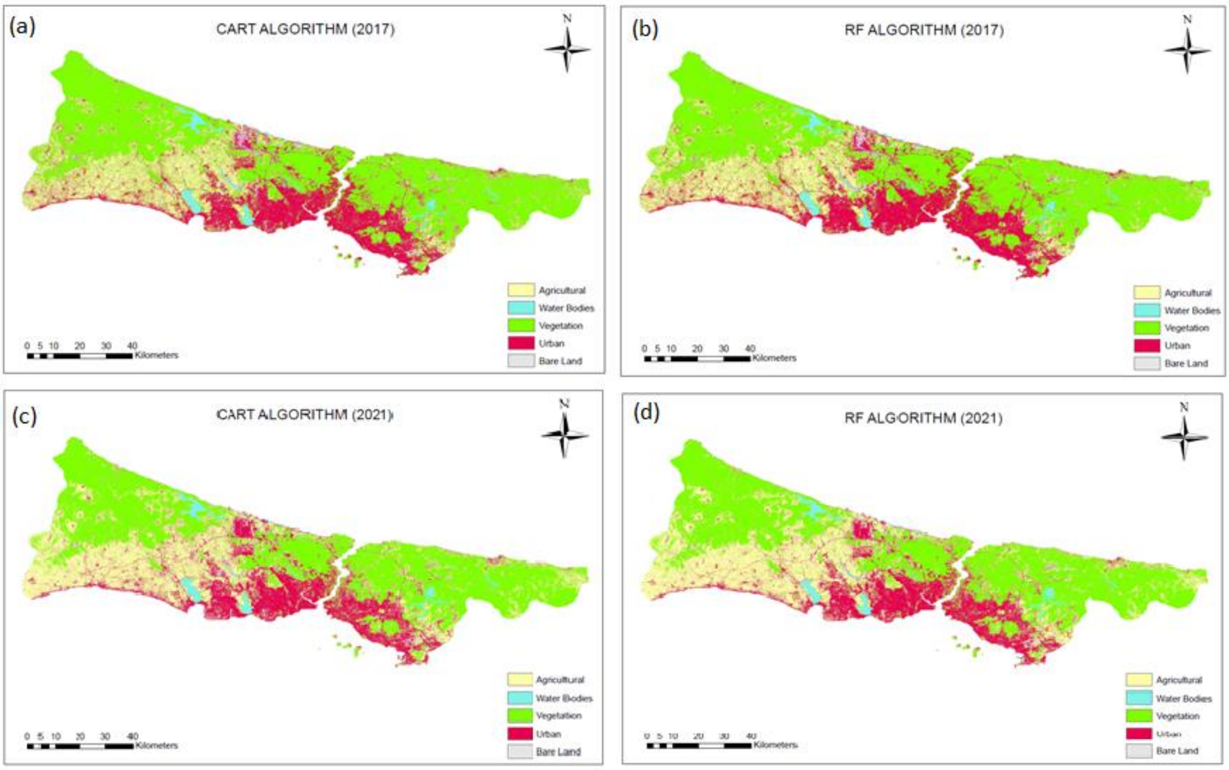

Figure 3. Sentinel 2 classification results of a) 2017 image with CART, b) 2017 image with RF, c) 2021 image with CART, and d) 2021 image with RF.

Table 2. Class based and overall accuracy results for each year

	2017		2018		2019		2020		2021	
	RF	**CART**	**RF**	**CART**	**RF**	**CART**	**RF**	**CART**	**RF**	**CART**
Agricultural	88.57%	80.86%	89.32%	79.93%	95.31%	88.23%	90.43%	81.30%	95.58%	91.32%
Water	97.01%	95.77%	97.48%	97.14%	96.38%	98.02%	95.08%	92.79%	93.78%	92.69%
Vegetation	99.55%	98.60%	97.81%	97.19%	98.12%	92.90%	98.50%	98.24%	96.52%	95.99%
Urban	91.67%	87.23%	88.78%	76.55%	93.90%	86.56%	89.69%	84.43%	91.15%	80.21%
Bare Land	85.76%	82.72%	84.88%	76.70%	89.30%	80.98%	90.50%	83.99%	82.70%	76.18%
Overall Accuracy	93.01%	89.11%	92.02%	85.42%	95.23%	89.18%	92.81%	88.08%	93.53%	88.45%

Table 3. Class based recall, precision, and F1 metrics for each year.

	Random Forest (RF)														
	2017			**2018**			**2019**			**2020**			**2021**		
	Rec	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**
Agricultural	0.89	0.96	0.92	0.89	0.91	0.90	0.95	0.91	0.93	0.90	0.90	0.90	0.96	0.91	0.93
Water	0.97	0.99	0.98	0.98	0.99	0.98	0.96	1.00	0.98	0.95	0.98	0.96	0.94	0.99	0.96
Vegetation	1.00	0.93	0.96	0.98	0.92	0.95	0.98	0.98	0.98	0.99	0.94	0.96	0.97	0.96	0.96
Urban	0.92	0.92	0.92	0.89	0.95	0.92	0.94	0.96	0.95	0.90	0.96	0.93	0.91	0.95	0.93
Bare Land	0.86	0.86	0.86	0.85	0.84	0.85	0.89	0.93	0.91	0.91	0.83	0.87	0.83	0.83	0.83
	CART														
	2017			**2018**			**2019**			**2020**			**2021**		
	Rec	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**	**Rec**	**Pre**	**F1**
Agricultural	0.81	0.93	0.87	0.80	0.89	0.84	0.88	0.81	0.85	0.81	0.90	0.85	0.91	0.82	0.86
Water	0.96	0.98	0.97	0.97	0.94	0.96	0.98	0.99	0.98	0.93	0.98	0.95	0.93	0.98	0.95
Vegetation	0.99	0.88	0.93	0.97	0.87	0.92	0.93	0.95	0.94	0.98	0.88	0.93	0.96	0.96	0.96
Urban	0.87	0.89	0.88	0.77	0.89	0.82	0.87	0.89	0.88	0.84	0.91	0.87	0.80	0.92	0.86
Bare Land	0.83	0.78	0.80	0.77	0.62	0.68	0.81	0.85	0.83	0.84	0.69	0.76	0.76	0.68	0.72

When the 5-year land cover change in the province of Istanbul is examined; it is determined that the land cover class in which the most change occurred in the area between 2017-2021 belonged to the urban land cover (22.92% RF, 36.77% CART) (Table 4). The biggest reason for this change can be estimated as the completion of the 3rd Airport project that turned this region into an urban land cover class, and the construction of new settlements in Istanbul, which receives irregular migration (Figure 4). During this process, the changes in the surface area of agricultural land areas are due to environmental regulations, especially along the Marmara Highway. Despite the decrease and increase in area from year to year, as of 2021, with the completion of the construction of this highway to a large extent, there has not been much change in the agricultural land cover (4.26% RF, -8.81% CART). In addition, it was determined that the greatest decrease occurred in forest areas (-14.06%RF, -9.54% CART) during the construction of new settlements. In addition, there is a significant increase in bareland class (19.16% RF, 7.74% CART) was observed. These barelands have the potential to be turned into urban areas in the near future. It has been concluded that this is one of the important reasons for the increase in agricultural land vegetation, since the regions where the stream beds feeding the lakes are classified as agricultural land due to the change in water bodies. As a result, it was determined that the vegetation and water body land areas within the five land cover classes determined within the scope of the study decreased over time and these areas turned into new settlements or became construction sites in the process.

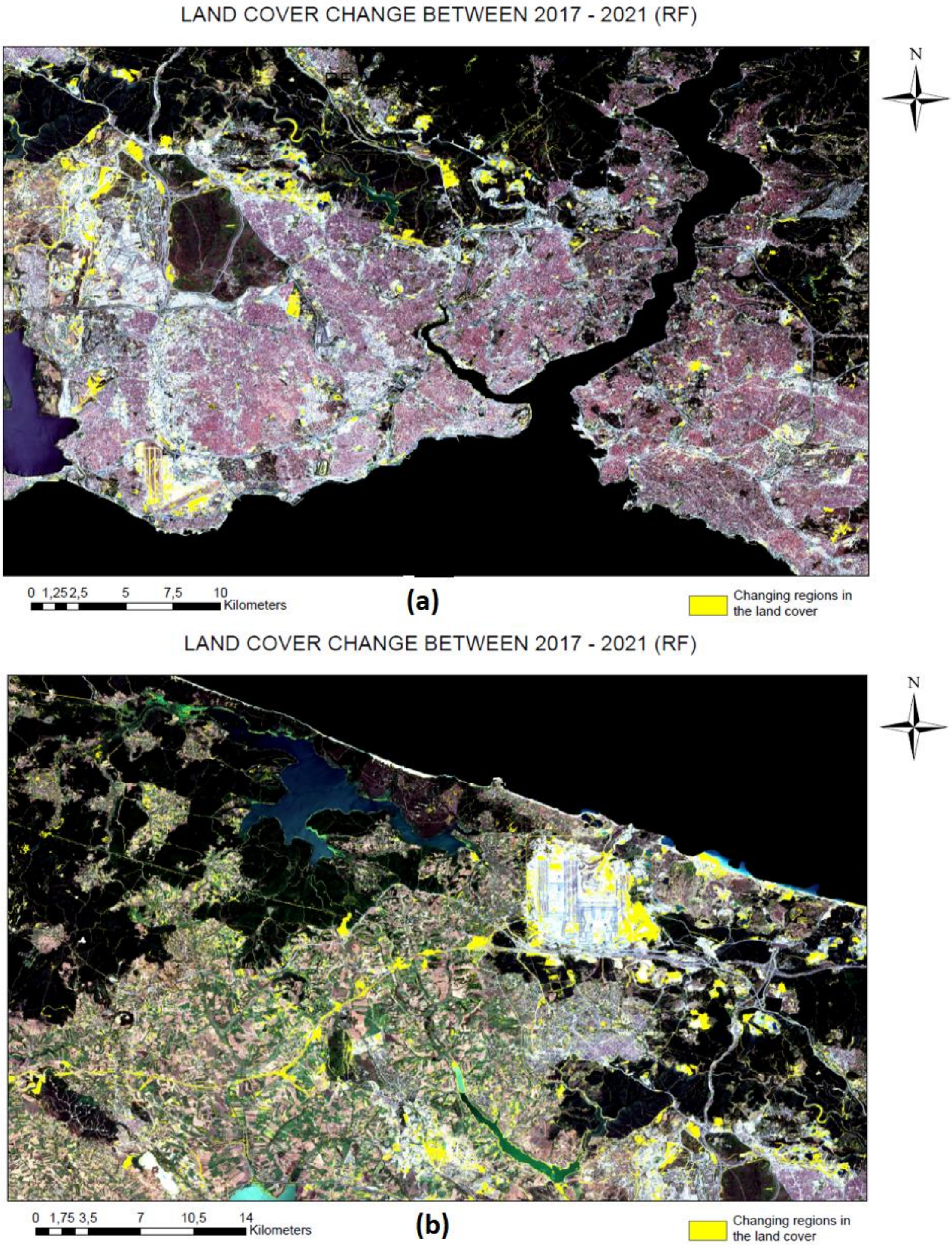

Figure 4. Change detection visualization for a) urban regions, and b) new transportation facilities.

Table 4. Percentage change in the areas of land cover classes compared to 2017.

	2018		2019		2020		2021	
	RF	**CART**	**RF**	**CART**	**RF**	**CART**	**RF**	**CART**
Agricultural	-2.52%	-6.68%	-4.81%	-9.43%	-2.51%	-10.22%	4.26%	-8.81%
Water	10.20%	5.98%	-2.38%	-3.72%	0.69%	3.00%	-4.92%	-6.25%
Forest	-3.32%	-0.02%	-3.24%	1.85%	-10.19%	-6.06%	-14.06%	-9.54%
Urban	4.73%	6.40%	13.21%	11.58%	25.32%	26.01%	22.92%	36.77%
Bareland	37.87%	25.90%	9.53%	7.08%	6.87%	30.72%	19.16%	7.74%

5. CONCLUSIONS

The purpose of this study is to provide a comparative evaluation of different machine learning based algorithms in land cover classification and change detection. Within this context, random forest and Classification and Regression Trees algorithms were utilized on Sentinel 2 images to create annual land cover maps of Istanbul city of Turkey. After classification, annual land cover changes were determined for five main land cover classes. Results proved that there is a drastic decrease in forest and agricultural lands, which increased urban lands. The increase of barren lands also indicates a possible change for further urbanization. Accuracy assessment results provided higher accuracies for both algorithms, however, visual inspections of the results showed that the random forest algorithm is slightly better in separating classes that have high mix potentials such as bare lands and settlements. Moreover, the random forest was found to better discriminate the change regions.

References

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.

Aghlmand, M., Kalkan, K., Onur, M. İ., Öztürk, G. & Ulutak, E. 2021. Google Earth Engine ile arazi kullanımı haritalarının üretimi . Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi , 10 (1) , 38-47.

Alganci, U. 2019. Dynamic Land Cover Mapping of Urbanized Cities with Landsat 8 Multi-temporal Images: Comparative Evaluation of Classification Algorithms and Dimension Reduction Methods. ISPRS Int. J. Geo-Inf. 8, 139.

Bégué A, Arvor D, Bellon B, Betbeder J, De Abelleyra D P D. Ferraz R, Lebourgeois V, Lelong C, Simões M R & Verón S. 2018. Remote Sensing and Cropping Practices: A Review. Remote Sensing, 10(1), 99.

Breiman, L. 1984. Classification And Regression Trees (1st ed.). Routledge.

Breimann, L. 2001. Random Forests. Machine learning, 45(1), 5-32.

De'ath, G. and Fabricius, K.E. 2000. Classification and Regression Trees: A Powerful yet Simple Technique for Ecological Data Analysis. Ecology, 81, 3178-3192.

Ettehadi Osgouei, P.; Kaya, S.; Sertel, E.; Alganci, U. 2019. Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. Remote Sens. 11, 345.

Jamali, A. 2019. A Fit-for-Purpose Algorithm for environmental monitoring based on maximum likelihood, support vector machine and random forest. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-3/W7, 25–32.

McCarty, D.A.; Kim, H.W.; Lee, H.K. 2020. Evaluation of Light Gradient Boosted Machine Learning Technique in Large Scale Land Use and Land Cover Classification. Environments, 7, 84.

Qin, W., Guo, W., Liu, X., & Zhao, H. 2019. A Novel Scheme for Recruitment Text Categorization Based on KNN Algorithm. In International Conference on Smart Computing and Communication, 376-386.

Rawat, J. S., & Kumar, M. 2015. Monitoring land use/cover change using remote sensing and GIS techniques: A case study of Hawalbagh block, district Almora, Uttarakhand, India, The Egyptian Journal of Remote Sensing and Space Science, 18(1), 77-84.

Schonlau M, Zou RY. 2020. The random forest algorithm for statistical learning. The Stata Journal. 20(1), 3-29.

Shetty, S., Gupta, P. K., Belgiu, M., & Srivastava, S. K. 2021. Assessing the Effect of Training Sampling Design on the Performance of Machine Learning Classifiers for Land Cover Mapping Using Multi-Temporal Remote Sensing Data and Google Earth Engine. Remote Sensing, 13(8), 1433.

Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad; Pal, S.; Liou, Y.-A.; Rahman, A. 2020. Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review. Remote Sens. 12, 1135.

Yu, L., Liang, L., Wang, J., Zhao, Y., Cheng, Q., Hu, L., Liu, S., Yu, L., Wang, X., Zhu, P., Li, X., Xu, Y., Li, C., Fu, W., Li, X., Li, W., Liu, C., Cong, N., Zhang, H. … Gong, P. 2014. Meta-Discoveries from a Synthesis of Satellite-Based Land Cover Mapping Research. International Journal of Remote Sensing 35 (13), 4573–4588.