Intercontinental Geoinformation Days

igd.mersin.edu.tr

Sentinel-1 and -2 time-series data-fusion for olive tree identification in heterogeneous land surfaces using Google Earth Engine

Haydar Akcay*, **Sinasi Kaya**, **Elif Sertel**, **Ugur Alganci**, **Samet Aksoy**

Istanbul Technical University, Civil Engineering Faculty, Geomatics Engineering Department, Istanbul, Turkey

Keywords
Time-series data fusion
Sentinel-1 Sentinel-2
DVI red index (DVIR)
Random forest classification
Google earth engine

ABSTRACT

Olive, a crucial crop for the economies of Mediterranean countries, is expanded to Aegean, Mediterranean, Marmara, South-East and Black Sea regions of Turkey. Identification of olive trees in heterogeneous land surfaces, particularly in mountainous regions is essential for exploitation of un-grafted olive trees. In this study, several samples of olive tree, agriculture, bare-land, urban, forest and sparse vegetation fields located between Bayındır and Tire districts of Izmir province in Turkey, are randomly selected. Independent two sample sets are generated to train the classifier (70%) and for the validation (30%). Several data fusion combinations of time series of Sentinel-1 and Sentinel-2 satellite data with various spectral indices are performed with random forest classifier on Google Earth Engine environment. A new spectral index, named as "DVI Red index (DVIR)" is generated and experimented in the study, as well. Results demonstrated that "Sentinel-1, Sentinel-2 and 10 indices" data fusion performed best overall accuracy (95.5%) as "Sentinel-1 and new ratio index (DVIR)" data fusion performed highest user's accuracy (97.2%) for olive class. Of 10 spectral indices standalone classifications, DVIR ranked the first for overall accuracy (94.8%) and the third for olive class user's accuracy (84.4%).

1. INTRODUCTION

Olive, one of the most widely distributed crops around the world, is cultivated in each continent (Khan et al., 2018). It is crucial for the economies of Mediterranean countries and can be considered as a strategic crop (Akcay et al., 2019). Turkey is ranked fourth in 2017 with 2,100,000 tons of olive production in the world (FAO, 2019) and olive is the fourth most produced fruit following grape, apple and orange within the country (TUIK, 2019). Olive production is primarily a family business of which 320.000 families live off (Directorate General of Cooperatives of Turkey, 2018). Determination and continuous monitoring of olive trees particularly un-grafted ones are vital when evaluating olive as a strategic crop. It is time and labor-intensive work with traditional methods as olive is an indigenous crop of Mediterranean as well as exists in mountainous regions that are difficult to access. Hence, remote sensing is an efficient option and a very popular technique in order to overcome these challenges.

Sentinel-1 and -2 supply free data in the microwave and optical range of the electromagnetic spectrum in high spatial and temporal resolution. Recent studies exhibit the potential of Sentinel-2 satellite data for land cover determination with high accuracy. Besides, different studies experimented the performance of Sentinel-1 and Sentinel-2 data fusion of which were reported improvements in classification accuracy (Heckel et al., 2020).

Vegetation indices generally provide more sensitive results than individual spectral bands of satellite images (Avola et al., 2019). Numerous spectral indices are experimented to find the best index for separating different land covers or a specific one. For instance, Pena-Barragan et al. (2002) calculated various indices to identify cover crops, bare soil and olive trees and used (Blue+Green+Red)/3 ratio index to discriminate bare soil from vegetation cover and olive trees. Ettehadi et al. (2019) who focused on separating built-up areas from bare-lands exploited NDVI, SAVI and NDVIre for determination of vegetation cover. As such, several spectral indices are employed in different studies to find

* **Corresponding Author**

*(akcayhay@itu.edu.tr) ORCID ID 0000-0002-2404-2390
(kayasina@itu.edu.tr) ORCID ID 0000-0002-4962-0492
(sertele@itu.edu.tr) ORCID ID 0000-0003-4854-494X
(alganci@itu.edu.tr) ORCID ID 0000-0002-5693-3614
(aksoysa@itu.edu.tr) ORCID ID 0000-0001-6420-7316

Cite this study

Akcay H, Kaya S, Sertel E, Alganci U & Aksoy S (2021). Sentinel-1 and -2 time-series data-fusion for olive tree identification in heterogeneous land surfaces using Google Earth Engine. 2nd Intercontinental Geoinformation Days (IGD), 159-162, Mersin, Turkey

the best index to separate different land covers or a specific one.

In this study, multi-sensor time-series satellite data are utilized for identification of olive trees. In this regard, a total of 42 Sentinel-1 and Sentinel-2 satellite image pairs were accessed for 21 dates during two-year period between 2019 and 2021 in Google Earth Engine (GEE) platform. 10 different spectral indices, including a new experimental index, which we named as “DVI Red index (DVIR)” are calculated for each date and different combinations of bands and indices are stacked. Random forest classifier is performed to each stack. The results were assessed with accuracy assessment to compare the performance of different data fusions.

2. MATERIALS AND METHODS

2.1. Study Area

The study area is located between Bayındır and Tire districts of Izmir province in Turkey with a size of 225 km^2 (15 km x 15 km) (Fig. 1). North and south sides of the landscape are mountainous with heterogeneous plantations of which olive trees, sparse vegetation and bare-lands largely cover the area. Agricultural fields are mainly concentrated in the middle of the study site while olive trees also exist.

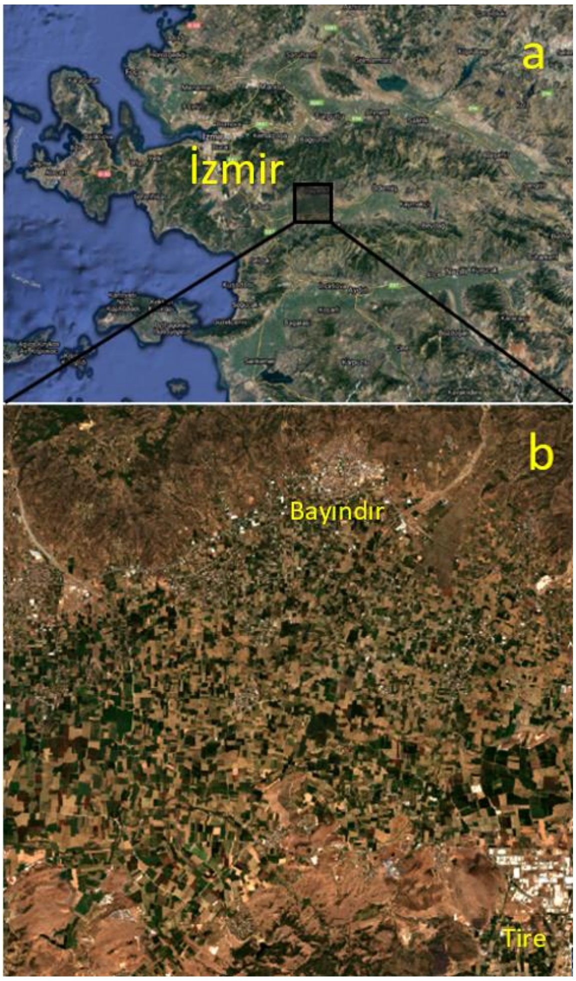

Figure 1. a) Study area from Google Earth© 2021 and b) Sentinel-2 (RGB: B4, B3, B2)

90 olive orchards, 70 agricultural parcels, 60 urban, 30 forest, 20 bare-land and 20 sparse vegetation fields are randomly selected from the study site such that samples for each class are scattered within the site. 70% and 30% of the sample sets are separated and used independently to train the random forest classifier and the validation of the classification, respectively.

2.2. Data Preparation

GEE (https://earthengine.google.com/), a cloud-based planetary scale geospatial analysis platform, is accessed and controlled through Internet-accessible application programming interface (API) (Gorelick et al., 2017). It enables exploiting several data collections of various remote sensing satellites (e.g. Landsat-5, -7, -8, Sentinel-1, -2, MODIS).

Sentinel-1 satellites are operating in C-band Synthetic Aperture Radar (SAR) and provide data at dual polarization (VH, VV) in Interferometric Wide Swath mode (IW) with 10 x 10 m pixel spacing (Torres et al., 2012). Sentinel-1 image collection on GEE platform includes calibrated and ortho-corrected Ground Range Detected (GRD) scenes of which images are processed using the Sentinel-1 Toolbox. Sentinel-2 satellites were launched in 2015 (S2A) and 2017 (S2B) and have 13 spectral bands with 10-, 20- and 60-m spatial resolutions (Laurin et al., 2018). Note that band 10 (SWIR-cirrus) is not available for level 2A products which are atmospherically corrected and provide bottom-of-atmosphere reflectance values.

In this study, a total of 42 Sentinel-1 and -2 satellite scenes (21 for each satellite sensor) between 01.01.2019 and 01.01.2021 were accessed in GEE environment. Closest image dates of each satellite were matched. VH/VV ratio band, radar vegetation index (RVI), normalized difference vegetation index (NDVI), Red-Edge NDVI (NDVIre), (Red+Green+Blue)/3 ratio index, difference vegetation index (DVI), normalized difference tillage index (NDTI), soil-adjusted vegetation index (SAVI), normalized difference built-up index (NDBI), bare soil index (BSI) and a new index named as “DVI red index (DVIR)” were calculated and included to Sentinel-1 and -2 data sets. Note that DVIR is generated using DVI and Red band which were noticed as distinctive when separating olive orchards. Table 1 shows the formulas of utilized spectral indices (B: Blue G: Green R: Red, NIR: Near-infared, SWIR: Shortwave-infrared).

Table 1. Spectral Indices

Index Name	Formula
NDVI	(NIR-R) / (NIR+R)
NDVIre	(RE1-R)/ (RE1+R)
(B+G+R)/3	(B+G+R)/3
DVI	NIR – R
NDTI	(SWIR1-SWIR2) / (SWIR1+SWIR2)
SAVI	(NIR-R) x (1+L) / (NIR+R+L) L=1.5
NDBI	(SWIR-NIR) / (SWIR+NIR)
BSI	((SWIR+R) - (NIR + B)) / ((SWIR+R) + (NIR + B))
RVI	4 x VH / (VV+VH)
DVIR	(R-DVI) / (R+DVI) = (2R-NIR) / NIR

2.3. Classification

Random forest (RF) classifier is a modern approach proved to perform fast and accurate for large and variegated data sets (Bargiel et al., 2017), and uses Classification and Regression Trees (CARTs) as working principle. A CART attempts to segment a predictor space into a number of homogenous regions which can be predicted by a generated rule set based on the input data. RF generates large number of CARTs based on various selections of the input data sets that are later summed up to one final result (Braun et al., 2015). Therefore, RF was selected as the classifier for the study.

Sentinel-2 images selected for the study are mostly cloud free. However, cloud mask was applied to the image collection as few parts of the scene covered by clouds. These regions, hence, were unclassified and displayed in white color in the classification image (Fig. 2). In the next step, Sentinel-1 and Sentinel-2 bands and calculated indices were combined. (21 dates, 4 Sentinel-1 bands/indices and 21 Sentinel-2 bands/indices). Therefore, an image with 21 x 25 = 525 layers were acquired. Different stack combinations were generated (e.g. S1, S2, S1+S2, S1+DVIR, S2+9 indices, S1+S2+10 indices) and random forest classification was implemented to each stack. 70% of samples were used for training the classifier and 30% of samples were used for the validation. It should be noted that the sample sets for training and the validation were divided and used independently. Table 2 demonstrates the accuracy metrics of each data fusion. Table 3 shows the confusion matrix of "S1 + S2 + 10 Indices" data fusion classification which performed the highest overall accuracy. Note that numbers refer to pixels of which an average olive orchard has approximately 50 pixels (with 10 m spatial resolution).

Table 2. Accuracy Metrics

Data Stack	User's Accuracy of Olive %	Overall Accuracy %	Kappa
S1+DVIR	97.2	94.2	0.92
S2+DVIR	93.2	94.9	0.93
S2	93.1	93.0	0.90
S2+9 indices	92.4	95.1	0.93
S1+S2+10 indices	91.4	95.5	0.94
S1+S2	91.2	94.7	0.93
(R+G+B)/3	86.1	91.8	0.89
BSI	85.3	88.2	0.84
DVIR	84.4	94.8	0.93
NDBI	83.5	87.1	0.82
NDVI	82.5	94.0	0.92
S1+NDVI	81.4	93.5	0.91
SAVI	79.2	92.0	0.89
DVI	76.1	90.7	0.87
NDVIre	65.2	85.2	0.79
NDTI	62.2	83.1	0.76
S1	55.5	66.4	0.52
S1+RVI	55.2	66.0	0.52
(VH/VV)	42.7	53.6	0.31
RVI	42.6	53.6	0.31

3. RESULTS

DVIR standalone classification succeeded 84.4% user's accuracy for olive class, following the (R+G+B)/3 ratio index (86.1%), and BSI (85.3%). Sentinel-1 standalone bands and RVI gave the lowest user's accuracies as expected (Table 2). S1+DVIR stack increased the accuracy remarkably and performed the highest user's accuracy (97.2%) as S1+NDVI, S1+(R+G+B)/3 and S1+BSI stacks remained at 81.4%, 82.3% and 82.7% respectively for olive class. S2+DVIR performed the second highest user's accuracy (93.2%) while S2 standalone classification performed 93.1% user's accuracy for olive class.

Regarding to overall classification accuracies, S1+S2+10 indices fusion succeeded the highest (95.5%) as S2+DVIR, S1+S2 and S1+DVIR stacks performed slightly lower accuracies of 94.9%, 94.7% and 94.2% respectively. DVIR standalone classification performed the highest (94.8%) compared to rest of the standalone index classifications.

Overall accuracy of DVIR standalone classification (94.8%) performed slightly higher than S1+DVIR (94.2%) but slightly lower than S2+DVIR stack (94.9%). This result can demonstrate the effectiveness of DVIR in land cover land use (LCLU) mapping. Fig. 2 describes the LCLU classification image of the study site.

Table 3. Confusion Matrix of "S1 + S2 + 10 Indices" Data Fusion Classification

Class	Bare-land	Agr.	Olive	Urban	Forest	Sparse Veg.	Total	User's Acc.
Bare-land	1208	0	1	21	0	0	1230	98.2
Agr.	0	2992	43	0	61	0	3096	96.6
Olive	0	3	1714	3	156	0	1876	91.4
Urban	1	5	1	220	0	0	227	96.9
Forest	0	0	0	0	477	0	477	100
Sparse Veg.	18	0	6	0	1	103	128	80.5
Total	1227	3000	1765	244	695	103		kappa
Producer's Acc.	98.5	99.7	97.1	90.2	68.6	100	kappa	0.94

4. DISCUSSION

As the purpose of the study was identifying olive orchards in heterogenous land surfaces with multi-sensor time-series data fusion, Sentinel-1 and -2 data were accessed and several spectral indices were calculated, combined and processed in Google Earth Engine environment. Results indicated that multi-sensor data fusion significantly increased the classification accuracy of SAR data. Furthermore, results show that just one optical index can improve the accuracy of SAR data more than %50 percent. For instance, S1+DVIR data fusion performed the highest user's accuracy for olive class (97.2%) whereas S1 and DVIR standalone classifications remained at 55.5% and 84.4% respectively. This explains the efficiency of SAR data when fused with optical indices. Hence, it is obvious that DVIR has significant role in identifying olive trees. It should be noted that the positive impact of time-series data is also clear in this study. According to Table 3, olive mostly mismatched with forest class as expected. Because different tree groups including olive can be located in the same field (Fig. 2). Bare-land mismatched mostly with urban and sparse vegetation as expected. Large discrepancies between user's and producer's accuracies of forest and sparse vegetation classes are thought to be due to the lack of samples of these classes. Lastly, it was observed that joint use of (VH/VV) and RVI has negative effect on classification accuracy which may be explained with the negative effect of data redundancy.

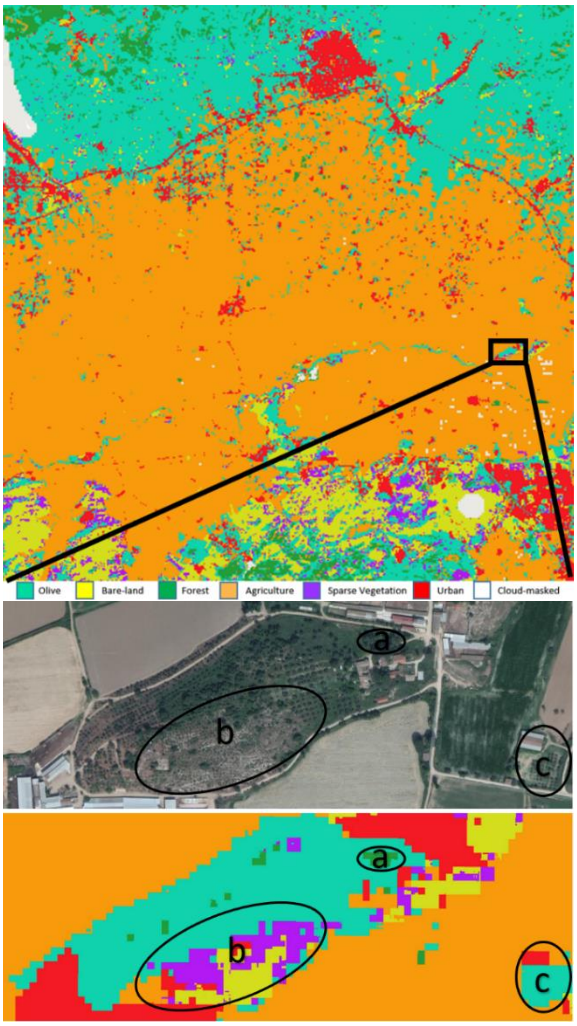

Figure 2. Classified image of "S1+S2+10 indices" data fusion. "Circle a" shows a group of trees, correctly classified as forest. In contrast, "circle b" demonstrates misclassification of olive class with sparse vegetation, agriculture, and bare-land. "Circle c" shows correct classification of urban and olive fields.

5. CONCLUSION

In this study, classification performances of data fusion of time series Sentinel-1, Sentinel-2 satellite data and several spectral indices experimented. It is observed that, DVIR has significant positive impact for not only in identifying olive trees but also land cover land use (LCLU) mapping. However, more comprehensive future works focused on DVIR are necessary to verify the efficiency of this index for both olive tree determination and LCLU mapping.

REFERENCES

Akcay, Haydar, et al. "Determination of Olive Trees with Multi-sensor Data Fusion." 2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics). IEEE, 2019.

Avola, G. et al., (2019). Remotely Sensed Vegetation Indices to Discriminate Field-Grown Olive Cultivars. Remote Sensing, 11(10), 1242.

Braun, A., and V. Hochschild. "Combined use of SAR and optical data for environmental assessments around refugee camps in semiarid landscapes." International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences (2015).Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Belmont, Wadsworth.

Directorate General of Cooperatives of Turkey (2018). 2017 Yılı Zeytin ve Zeytin Yağı Raporu. http://koop.gtb.gov.tr/data/5ad06f17ddee7dd8b42 3eb2e/ 2017%20Zeytinya%C4%9F%C4%B1%20Raporu.pd f (Accessed 25 June 2019).

Earth Engine Data Catalog, https://developers.google.com/earth-engine/datasets/catalog (accessed on 28 March 2021).

Ettehadi P. et al., (2019). Separating Built-Up Areas from Bare Land in Mediterranean Cities Using Sentinel-2A Imagery. Remote Sensing, 11(3), 345.

FAO, 2019. Agricultural Statistics of the Food and Agriculture Organization of the United Nations, http://www.fao.org/faostat/en/#data/QC. (Accessed 27 June 2019).

Google Earth Engine, https://earthengine.google.com/ Accessed (30 March 2021).

Gorelick, Noel, et al. "Google Earth Engine: Planetary-scale geospatial analysis for everyone." Remote sensing of Environment 202 (2017): 18-27.

Heckel, Kai, et al. "Predicting forest cover in distinct ecosystems: The potential of multi-source Sentinel-1 and-2 data fusion." Remote Sensing 12.2 (2020): 302.

Khan, Aftab, et al. "Remote Sensing: An Automated Methodology for Olive Tree Detection and Counting in Satellite Images." IEEE Access 6 (2018): 77816-77828

Laurin, G. V. et al., (2018). Above-ground biomass prediction by Sentinel-1 multitemporal data in central Italy with integration of ALOS2 and Sentinel-2 data. Journal of Applied Remote Sensing, 12(1), 016008.

Peña-Barragán, J. M. et al., (2004). Assessing land-use in olive groves from aerial photographs. Agriculture, ecosystems & environment, 103(1), 117-122.

Torres, Ramon, et al. "GMES Sentinel-1 mission." Remote Sensing of Environment 120 (2012): 9-24.

TUIK, (2019). Turkish Statistical Institute, http://tuik.gov.tr/UstMenu.do?metod=temelist (Accessed 25 June 2019).