Classification of Seasonal Sentinel-2 Imagery for Mapping Vegetation in Mediterranean Ecosystems

Konstantinos Antoniadis[a], Nikos Georgopoulos[a], Thomas Katagis[b], Dimitris Stavrakoudis[a], Ioannis Z. Gitas[a]
[a] Laboratory of Forest Management and Remote Sensing, Department of Forestry and Natural Environment, Aristotle University of Thessaloniki, P.O. Box 248, 54124 Thessaloniki, Greece;
[b]Department of Forestry and Management of the Environment and Natural Resources, Democritus University of Thrace, 68200 Orestiada, Greece

ABSTRACT

Earth Observation satellite systems are considered the main source of information used for delivering up-to-date land cover/use maps. Medium to high spatial resolution images, such as the ones provided by Sentinel-2 sensors, can improve significantly mapping and monitoring of vegetation communities and are utilized in a wide range of applications such as the management of natural resources and forest inventories. The aim of this work was to employ Sentinel-2 images for accurately classifying vegetation cover in selected areas of Greece that present diverse vegetation characteristics. Cloud-free Sentinel 2 (L2A) images were acquired for each area during 2021 for the months of February, June, and September, in order to capture the reflectance changes due to seasonal variations. Two machine-learning techniques, namely Random Forest (RF) and Support Vector Machines (SVM), were applied and assessed for their performance in mapping vegetation cover and species in the study areas. The training patterns, used as input in both classifiers, were acquired through photo-interpretation of stratified random points, distributed across forested areas. Consequently, validation of the classification results was performed, in order to estimate accuracy metrics for each model per site. More specifically, the kappa coefficient, overall (OA), user's and producers' accuracy were calculated. The accuracy results demonstrated higher scores for RF (OA over 90% for all areas) than SVM (OA ranging from 81 to 89%, respectively). Overall, our study demonstrates the capability of seasonal Sentinel-2 data to accurately discriminate vegetation communities over diverse biomes, when combined with advanced classification methods.

Keywords: Remote Sensing, Sentinel 2, Land cover, vegetation mapping, machine learning

1. INTRODUCTION

Land cover/use (LCLU) mapping provides essential information in climate change research[1]. Classifying vegetation characteristics is a crucial process for terrestrial carbon estimation, which facilitates climate change mitigation[2]. Vegetation maps are also used, in a plethora of applications in the field of natural resources management such as forest inventory, wildfire mapping, and water resources management[3].

Traditionally, LCLU mapping is based on field measurements, providing accurate results, although time-consuming and costly, especially in large and remote areas. Contrarily, remote sensing (RS) offers the opportunity for accurate and cost-effective LCLU mapping on different scales, utilizing airborne and spaceborne sensors [4]. With recent developments in RS systems, satellites can provide data at various spatial and temporal resolutions[5]. In the last decades, many studies on LULC mapping and monitoring have been carried out employing multispectral imagery from satellites, such as Landsat, Satellite for observation of Earth (SPOT), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Moderate Resolution Imaging Spectroradiometer (MODIS), and more[3], [4], [6], [7].

However, several authors have reported that medium to low-resolution observations, have a negative impact on the accuracy of the final product[8]–[11]. To overcome these limitations, machine learning (ML) approaches have been applied, instead of more traditional approaches (e.g. Maximum Likelihood and Minimum Distance) to classify remotely sensed images. Over a wide range of ML methods (e.g. Decision Tree, K-Nearest Neighbor, Artificial Neural Network, and XGBoost) Random Forest (RF) and Support Vector Machines (SVM), have gained a lot of attention in recent

Ninth International Conf. on Remote Sensing and Geoinformation of the Environment (RSCy2023), edited by K. Themistocleous, et al., Proc. of SPIE Vol. 12786, 1278609 · · doi: 10.1117/12.2680752

studies, due to their ability to provide reliable classification results, in different biomes and vegetation characteristics [12], [13].

With full constellation employed, Sentinel- 2A and Sentinel-2B satellites (launched on 23 June 2015 and 7 March 2017 respectively) provide high spatial (10m, 20m and 60m) and temporal resolution (approximately 5-day revisit cycle). In this study, we introduce a classification approach, based on machine learning techniques, for mapping mainly forest vegetation in three study areas in Greece. More specifically, we implement RF and SVM classification algorithms and compare their performance in mapping accurately vegetation species and cover, by employing seasonal Sentinel-2 multispectral imagery. The areas, selected for classification present different vegetation characteristics, including Greece's most dominant vegetation species. Therefore, one of the goals of this research is to also provide a reliable mapping methodology that could be used at national level for supporting forest management and forest inventory planning.

2. MATERIALS AND METHODS

2.1 Study area

The study was conducted in 2021 in the municipalities of Arta, Pella, and Korinthos (Fig.1) with approximate areas of 401km^2, 2506km^2, and 2297 km^2 respectively. The landscape of the selected sites exhibits a complex vegetation structure and consists of various species, with coniferous and broadleaves trees, shrubs, and grasslands covering the main bulk of the area. Moreover, the selected study sites include some of the most widespread forest species in Greece, such as Quercus coccifera, Pinus halepensis, Pinus brutia, and Abies cephalonica. The climate in all areas can be characterized as typical Mediterranean, with hot, dry summers and mild, rainy winters.

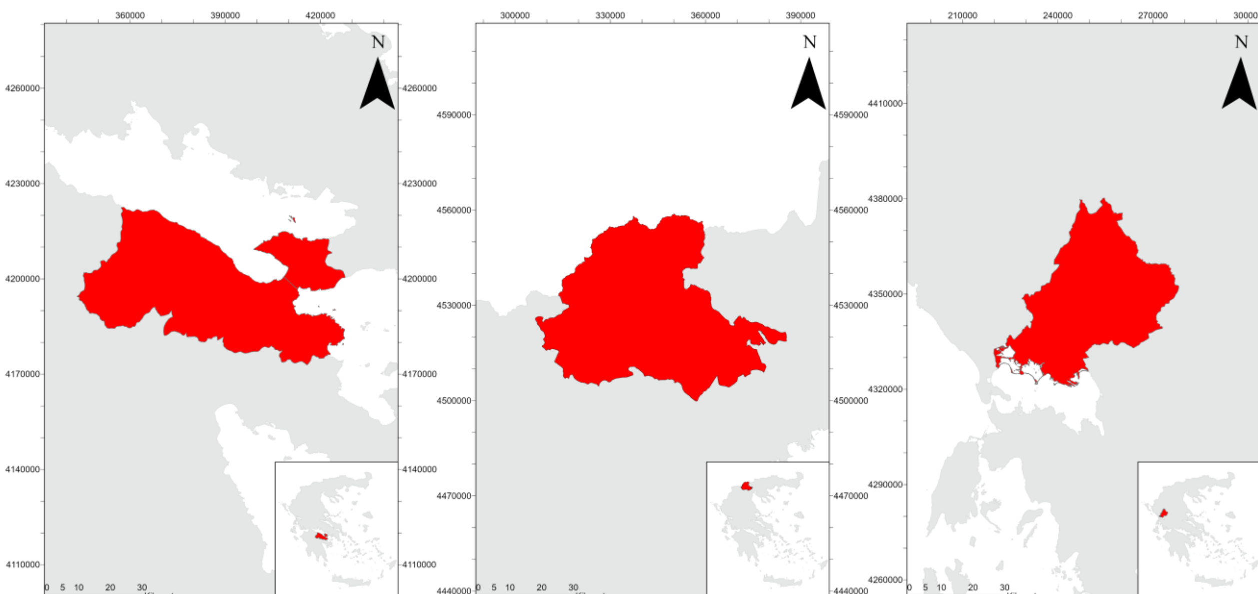

Figure 1. Study areas in Greece; Municipalities of Korinthos (left), Pella (center), Arta (right)

2.2 Satellite imagery

Three Sentinel-2 (Level-2A) images were acquired for each study area. More specifically, cloud-free images selected over three different seasons, namely from February, June and September of 2021, were utilized in the study. The scope here was to derive the seasonal spectral variations captured by Sentinel-2A images, in order to assist the mapping process and species discrimination. For the classification, eight spectral bands were utilized (Table 1). Basic pre-processing operations involving downloading, subsetting, cloud masking and stacking, were carried out in the Google Earth Engine (GEE) platform.

Table 1. Sentinel-2 bands used in this study.

Band	Description	Resolution
B02	Blue	10m
B03	Green	10m
B04	Red	10m
B05	Red-edge	20m
B06	Red-edge	20m
B8A	Narrow-NIR	20m
B11	SWIR1	20m
B12	SWIR2	20m

2.3 Vegetation mapping

Overall, the classification process included the implementation of two ML models (RF and SVM) for producing vegetation maps in principally forested areas. In order to ensure that only forested areas will be classified, a masking process was performed based on official land cover maps acquired from the Cadastral Agency of Greece[14]. Polygons that are characterized as “non-forested” or “other type of cover” were excluded from further analysis. The rest of the polygons representing forested areas were used to clip the Sentinel-2 images.

To prepare the classification set, a number of stratified random points were distributed across the forested areas and labeled with photo-interpretation, using Google Earth imagery. For Korinthos, Arta, and Pella 954, 1386, and 1415 random points were used respectively. The dataset was labeled according to eight vegetation types namely *Oaks, Fir, Pines, Conifers, Beech, Broadleaves, Evergreen Broadleaves, and Grasslands*. Prior to classification, the dataset was separated into two sets, namely training (70%) and testing (30%).After the training phase, the models were applied to Sentinel-2 images to derive the thematic maps with the aforementioned classes. Finally, the validation of the results was performed, in order to obtain the classification accuracy of each model. More specifically, the kappa coefficient, overall (OA), user’s (UA), and producers’ (PA) accuracy were calculated. For training, testing and classification ArcGIS Pro was used.

3. RESULTS AND DISCUSSION

Overall, the results of this study showcased that RF provided stronger classification capability (OA= 0.96, 0.90, and 0.95 with kappa=0.94, 0.85, and 0.93) than SVM (OA=0.89, 0.85, and 0.81 with kappa=0.84, 0.77, and 0.85) in all tested regions (table 2). The highest classification performance for both classifiers was observed in the Korinthos region. This can be attributed to the fact that Korinthos consists of a more homogenous landscape than the rest of the study areas. Also, the spectral discrimination of the vegetation types is more explicit in Korinthos than in Arta and Pella.

Table 2. Classification results in the three study areas using RF and SVM.

Study area	Model	Overall accuracy	Kappa
Korinthos	RF	0.96	0.94
	SVM	0.89	0.84
Arta	RF	0.90	0.85

	SVM	0.85	0.77
Pella	RF	0.95	0.93
	SVM	0.81	0.85

To access the performance of the best model (RF) for each vegetation type, Producer's accuracy (PA) and User's accuracy (UA) metrics were investigated (table 3). More specifically, the classes that were better distinguished were *Evergreen Broadleaves, Grasslands and Beech* achieving PA and UA above 90%. However, it should be noted that *Beech* was not observed in two of the three study areas. In the case of *Oaks,* RF achieved good results in all tested sites (above 80%), while the lower PA value was observed in Arta (0.8). This is due to the misclassification of pixels between *Oaks* and *Evergreen Broadleaves,* which can be explained by the fact that both vegetation types are phenologically similar. *Fir* also provided good results (above 85% in both metrics) in Korinthos and Arta. A lower PA value (0.83) in Pella, indicates a minor underestimation of *Fir*, as some pixels were omitted to *Beech* and *Pines* classes. Regarding *Pines*, over 0.95 values were observed in Pella and Korinthos. On the other hand, *Pines* cannot be distinguished clearly from *Fir* and *Evergreen Broadleaves,* in Arta (PA=0.65 and UA=0.75). For *Broadleaves,* although, high accuracy metrics were obtained in Korinthos, lower PA (0.74) was reported in Pella. After a closer examination of the results, it is observed that *Broadleaves* were misclassified as *Oaks*. The reason behind the confusion between these classes is the same described for the misclassification between *Oaks and Evergreen Broadleaves* in Arta. Overall, all classes provided good results (PA and UA above 70%). Only one exception can be observed, for *Conifers* in Korinthos, where significant omission errors were reported (PA=0.55). This can be attributed to the fact that *Conifers* and *Evergreen Broadleaves* could not be discriminated properly, as some species (mainly shrubs) tend to have similar phenological characteristics. The vegetation maps resulted from RF implementation are illustrated in figure 2.

Table 3. User's and producer's accuracy for vegetation types in the three study areas using RF.

	Korinthos		**Arta**		**Pella**	
Vegetation Type	**Producer's**	**User's**	**Producer's**	**User's**	**Producer's**	**User's**
Oaks	0.91	1	0.8	0.9	0.98	0.93
Fir	1	0.91	0.88	0.93	0.83	0.9
Pines	0.96	0.99	0.65	0.75	0.98	0.97
Beech	-	-	-	-	0.97	0.97
Broadleaves	1	0.83	-	-	0.74	0.98
Conifers	0.55	1	-	-	-	-
Evergreen Broadleaves	0.98	0.95	0.95	0.9	0.92	0.95
Grasslands	0.94	0.94	0.94	0.94	0.89	0.95

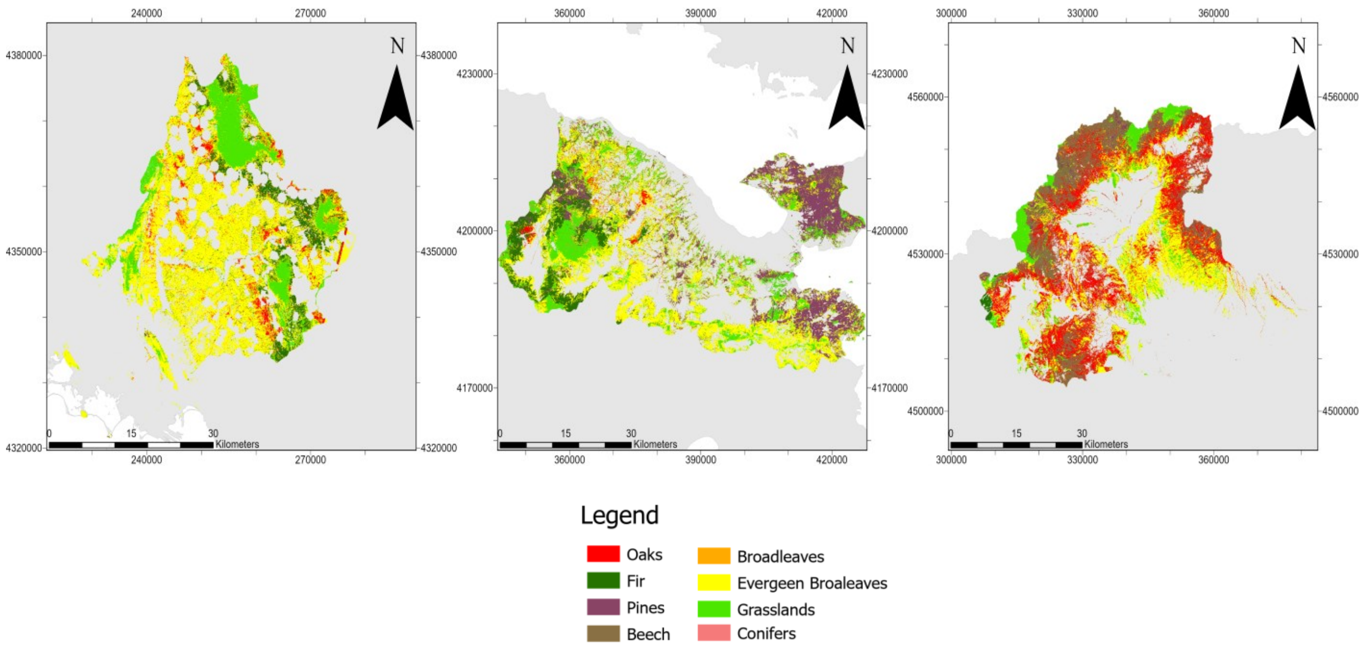

Figure 2. Vegetation maps (2021) for Arta (left), Korinthos (center) and Pella (right).

4. CONCLUSIONS

In this study we investigated the potential of Sentinel-2 imagery in vegetation mapping, using two popular ML methods, namely RF and SVM. The proposed method relies upon multi-temporal imagery, in order to capture the spectral variations, occurring due to different phenological phases. To test the proposed methodology, we selected 3 study areas in Greece, based on their different vegetation characteristics. The results of the study demonstrated that the combination of Sentinel-2 imagery with both methods can provide reliable vegetation maps in Mediterranean ecosystems. Both classifiers achieved high classification accuracies, with Random Forest outperforming SVM in all of the study areas. Despite the fact that this research reached its goals, some limitations should be further investigated in future work. More specifically, the rather small sample size could be enhanced, in order to avoid uncertainty in the results. Furthermore, the extraction of additional phenological traits, potentially utilizing time-series imagery, could contribute to better discrimination of *Oaks* and *Broadleaves*. Finally, a more holistic approach will also encompass a more detailed validation of the results, using very high-resolution imagery and/or field data. Overall, our study provides a reliable framework that can be used in the future as the basis for accurately mapping vegetation patterns in Mediterranean ecosystems, at regional and national scale.

REFERENCES

[1] Y. Xie, Z. Sha, and M. Yu, “Remote sensing imagery in vegetation mapping: a review,” Journal of Plant Ecology, vol. 1, no. 1, pp. 9–23, Mar. 2008, doi: 10.1093/jpe/rtm005.

[2] X. Xiao, “Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data,” Remote Sensing of Environment, vol. 91, no. 2, pp. 256–270, May 2004, doi: 10.1016/j.rse.2004.03.010.

[3] G. A. Carpenter, S. Gopal, S. Macomber, S. Martens, C. E. Woodcock, and J. Franklin, “A Neural Network Method for Efficient Vegetation Mapping,” Remote Sensing of Environment, vol. 70, no. 3, pp. 326–338, Dec. 1999, doi: 10.1016/S0034-4257(99)00051-6.

[4] S. K. Langley, H. M. Cheshire, and K. S. Humes, “A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland,” Journal of Arid Environments, vol. 49, no. 2, pp. 401–411, Oct. 2001, doi: 10.1006/jare.2000.0771.

[5] S. Talukdar et al., “Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review,” Remote Sensing, vol. 12, no. 7, p. 1135, Apr. 2020, doi: 10.3390/rs12071135.

[6] W. L. Stefanov and M. Netzband, “Assessment of ASTER land cover and MODIS NDVI data at multiple scales for ecological characterization of an arid urban center,” Remote Sensing of Environment, vol. 99, no. 1–2, pp. 31–43, Nov. 2005, doi: 10.1016/j.rse.2005.04.024.

[7] S. I. Toure, D. A. Stow, H. Shih, J. Weeks, and D. Lopez-Carr, “Land cover and land use change analysis using multi-spatial resolution data and object-based image analysis,” Remote Sensing of Environment, vol. 210, pp. 259–268, Jun. 2018, doi: 10.1016/j.rse.2018.03.023.

[8] R. Manandhar, I. Odeh, and T. Ancev, “Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement,” Remote Sensing, vol. 1, no. 3, pp. 330–344, Jul. 2009, doi: 10.3390/rs1030330.

[9] C. Yang, G. Wu, K. Ding, T. Shi, Q. Li, and J. Wang, “Improving Land Use/Land Cover Classification by Integrating Pixel Unmixing and Decision Tree Methods,” Remote Sensing, vol. 9, no. 12, p. 1222, Nov. 2017, doi: 10.3390/rs9121222.

[10] S. Pal and S. Talukdar, “Assessing the role of hydrological modifications on land use/land cover dynamics in Punarbhaba river basin of Indo-Bangladesh,” Environ Dev Sustain, vol. 22, no. 1, pp. 363–382, Jan. 2020, doi: 10.1007/s10668-018-0205-0.

[11] R. Latifovic and I. Olthof, “Accuracy assessment using sub-pixel fractional error matrices of global land cover products derived from satellite data,” Remote Sensing of Environment, vol. 90, no. 2, pp. 153–165, Mar. 2004, doi: 10.1016/j.rse.2003.11.016.

[12] M. Carranza-García, J. García-Gutiérrez, and J. Riquelme, “A Framework for Evaluating Land Use and Land Cover Classification Using Convolutional Neural Networks,” Remote Sensing, vol. 11, no. 3, p. 274, Jan. 2019, doi: 10.3390/rs11030274.

[13] L. Ma, M. Li, X. Ma, L. Cheng, P. Du, and Y. Liu, “A review of supervised object-based land-cover image classification,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 130, pp. 277–293, Aug. 2017, doi: 10.1016/j.isprsjprs.2017.06.001.

[14] “Forest maps,” < https://gis.ktimanet.gr/gis/forestsuspension > (31 October 2023).