*Proceeding Paper*

# Broad-Leaved and Coniferous Forest Classification in Google Earth Engine Using Sentinel Imagery †

## Gordana Kaplan

Institute of Earth and Space Sciences, Eskisehir Technical University, 26555 Eskisehir, Turkey; kaplangorde@gmail.com; Tel.: +90-536-697-5605

† Presented at the 1st International Electronic Conference on Forests—Forests for a Better Future: Sustainability, Innovation, Interdisciplinarity, 15–30 November 2020; Available online: https://iecf2020.sciforum.net.

**Abstract:** Forest structures knowledge is fundamental to understanding, managing, and preserving the biodiversity of forests. With the well-established need within the remote sensing community for better understanding of canopy structure, in this paper, the effectiveness of Sentinel-2 imagery for broad-leaved and coniferous forest classification within the Google Earth Engine (GEE) platform has been assessed. Here, we used Sentinel-2 image collection from the summer period over North Macedonia, when the canopy is fully developed. For the sample collection of the coniferous areas and the accuracy assessment of the classification, we used imagery from the spring period, when the broad-leaved forests are in the early green stage. A Support Vector Machine (SVM) classifier has been used for discriminating forest cover groups, namely, broadleaved and coniferous forests. According to the results, more than 90% of the canopy in North Macedonia is broad-leaved, while less than 10% is conifers. The results in this study show that, with the use of GEE, Sentinel-2 data alone can be effectively used to obtain rapid and accurate mapping of main forest types (conifers-broadleaved) with a fine resolution.

**Keywords:** broad-leaved forest; coniferous forest; remote sensing; Google Earth Engine; sentinel

## 1. Introduction

Large-scale forest maps are needed for efficient forest management. Also, this kind of map can be used for the analysis of impacts such as climate change on vegetation [1]. Forest information of national databases in combination with land covers, are fundamental for understanding, managing, and preserving forest biodiversity. The use of remote sensing data and technologies addresses forest canopy structure. However, it has been established that the use of a small number of multi-spectral bands (5–10 bands) may be challenging for classifying different canopy groups, such as separating coniferous and broad-leaved canopy structures [2]. In recent years, in order to overcome the limitation of middle-resolution multi-spectral satellite imagery, researchers have used different data such as photogrammetric products from unmanned aerial vehicle (UAV), airborne laser scanning data [3], Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) [4], high-resolution RGB imagery [5], full-waveform airborne laser scanning data [6], etc.

However, taking into consideration the higher spatial and spectral resolution of the newest optical satellite mission, Sentinel-2, it is assumed that using Sentinel's data, forest mapping and monitoring can be significantly improved using open-source middle-resolution satellite imagery. In comparison to other open-source remote sensing data, such as Landsat, Sentinel-2 offers three times better spatial resolution with four 10 m, six 20 m, and three 60 m bands, with the opportunity to increase the spatial resolution of all bands to 10 m using pan-sharpening and fusion techniques [7,8]. Also, Sentinel-2 offers three red-edge vegetation bands that significantly improve the vegetation classification [9]. Sen-

tinel-2 data has been used for several forest application studies. The evaluation of different machine learning algorithms for the classification of tree species based on Sentinel-2 data has been made by Wessel et al. [10] over mixed forest in Germany with approximately a 7600 ha area and achieved a significantly high accuracy of 90%. One of the most accurate and secure tree species classification methods is the one with multi-temporal data since the broad-leaved forests have no or low presence of vegetation in the winter days. Multi-temporal Sentinel-2 data have been used for classifying five different tree species in Sweden in an area of over 1500 ha with an accuracy of 88% [11]. Similar studies of the forest districts in Poland [12] and China [13] have been made, using multi-temporal data.

Google Earth Engine (GEE) as a cloud computing platform has been successfully utilized for studying, mapping and monitoring forests. Thus, GEE has been used for investigating the present status of forest stands using time-series Landsat data [14]. Sentinel-1, Landsat-8 and elevation data integration within GEE has been used in the participatory mapping of forest plantations [15], and it has also being used for mangrove forest mapping in China using Landsat and Sentinel-1 time series [16], and forest changes in the Amazon [17].

Having in mind the growing potential of GEE, and the successful use of Sentinel-2 in forest monitoring, in this paper, we use Sentinel-2 data within GEE for broad-leave and coniferous forest classification on national level. As a study area, the Republic of North Macedonia has been selected. Different from the other studies, in this study, we use single image collection from the summer period over the study area, and we implement Supported Vector Machine (SVM) classifier over a large area of approximately 2,571,300 ha.

## 2. Study Area and Methods

### 2.1. Study Area

The Republic of North Macedonia is a landlocked country in the middle of the Balkan Peninsula in Southeast Europe (Figure 1). With its total area of 2,571,300 ha, North Macedonia shares its boundaries with Serbia, Kosovo, Albania, Bulgaria, and Greece. North Macedonia has approximately 2.1 million inhabitants.
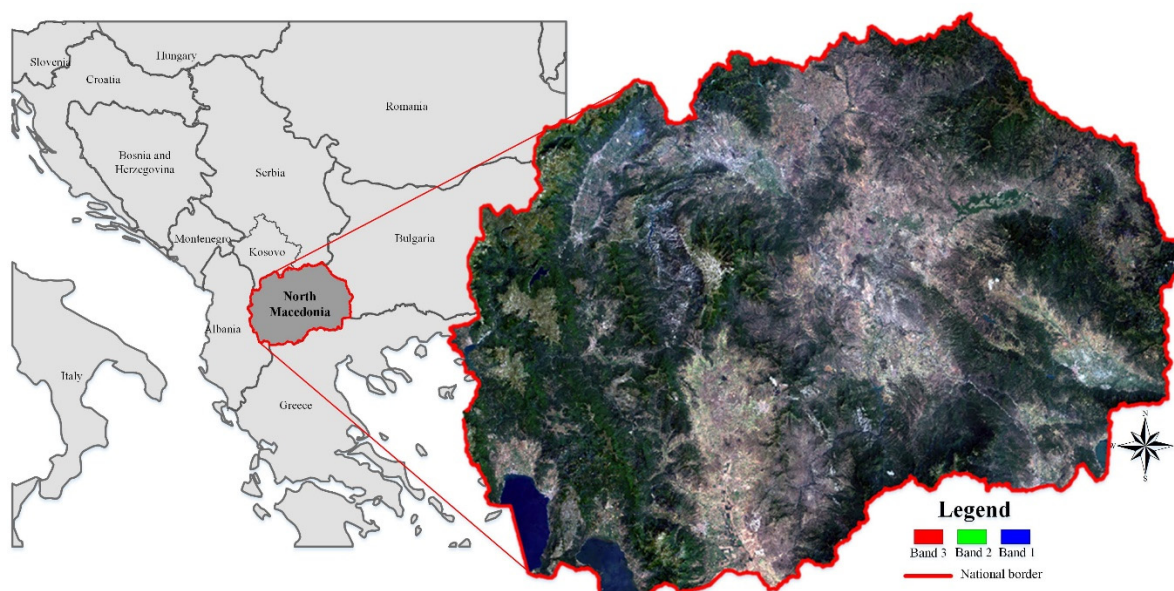


**Figure 1.** Study area, Republic of North Macedonia, Sentinel-2 imagery.

The geography of the country is defined by a central valley formed by the river Vardar framed with Shar and Osogovo Mountains. North Macedonia has three internationally shared large lakes, a large number of river water bodies flowing into the Aegean, Adriatic, and the Black Sea. Vardar is the largest river in the country. The Vardar basin

comprises two-thirds of the territory of North Macedonia, and it plays an integral part in the country's economy and development (Donev 2014).

North Macedonia has a number of national parks, wild mountain massifs covered with dense forest. Mount Korab is the highest peak (2764 m) in both North Macedonia in Albania and it is noted for its rich flora including both confers and broad-leaved forest. Shar Mountain, located in the north part of North Macedonia, is a rich massif, including more than 1800 plant species. The park includes the endemic relict Macedonian Pine, also known as Pinus peuce, which can be also found on Baba Mountain, the third highest mountain in North Macedonia, with a 2601 m peak, called Pelister [18]. According to the state's statistical data, in 2018, approximately 100,000 ha of the country has been covered with forest, and approximately 10% of them were conifers.

*2.2. Methodology*

In order to classify the areas covered with forest and therefore separate the coniferous and broad-leaved forest in the borders of the Republic of North Macedonia, the open-source Sentinel-2 data within the cloud computing platform has been used. Afterwards, using prior knowledge of the study area, where forests are located on the mountainous area, in order to exclude misclassification of the dense agricultural areas, first, the study area was separated as flat areas (Slope < 7°) that are generally used for agricultural and urban settlements, and non-flat areas, consisting of forests, and pastures. For this purpose, we use Sentinel-2 image collection from the summer period, when the vegetated cover is in its full contents. In order to get cloud-free imagery, we used image collection from 1 June 2019–31 August 2019, setting a cloud filter to be less than 5%. All 10 and 20 m Sentinel-2 bands were used for the classification, compiling a dataset of ten spectral bands with a 20 m spatial resolution. It should be noted that minimal sample training data were used for this study, or approximately 40–50 samples per class, for all four classes, namely, coniferous forest, broad-leaved forest, pastures, and water. The training was done over the 20 m Sentinel-2 data, using a Library for SVM (LIBSVM) classifier.

In order to assess the results, accuracy assessment using a confusion matrix, validation overall accuracy and kappa statistics have been calculated. While the validation overall accuracy gives us the accuracy of the classification of the training samples, the kappa statistics give us the accuracy of the classification. Half of training samples have been used for training, while half of the samples have been used for the accuracy assessment [19]. Also, the results have been visually compared with a satellite image from the early spring months, when the broad-leaved forest are still leafless.

The flowchart of the methodology used in this study can be seen in Figure 2.
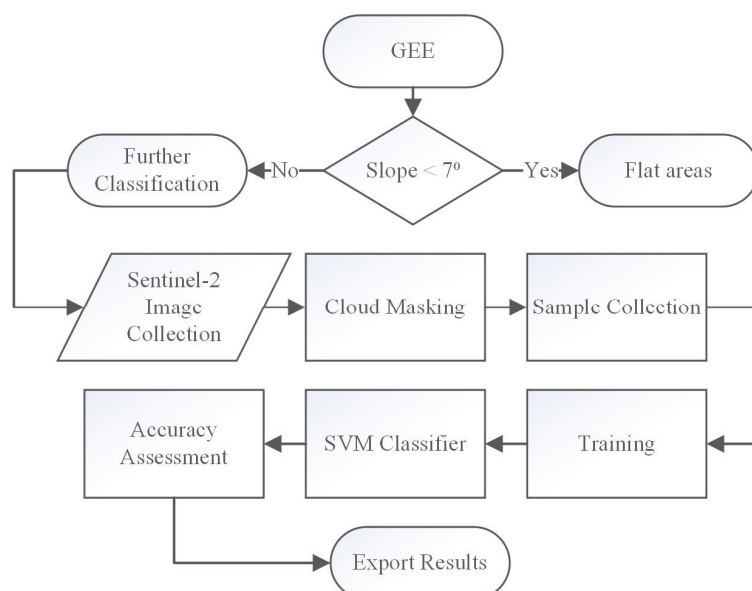
**Figure 2.** Flowchart of the used methodology.

## 3. Results and Discussion

For classifying the non-flat area into four groups, namely, coniferous forest, broad-leaved forest, pastures, and water, approximately 200 sample points were used. However, half of them were used for training and half for assessing the classification. The results of the classification are shown in Figure 3 and Table 1. As seen from the results, the largest area with confer forest is located in the north part for the Republic of North Macedonia, on the border with Greece, on the mountain Nice. According to the statistical report of the national forest, mountain Nice is rich in oak, pine, and other confer tree types.
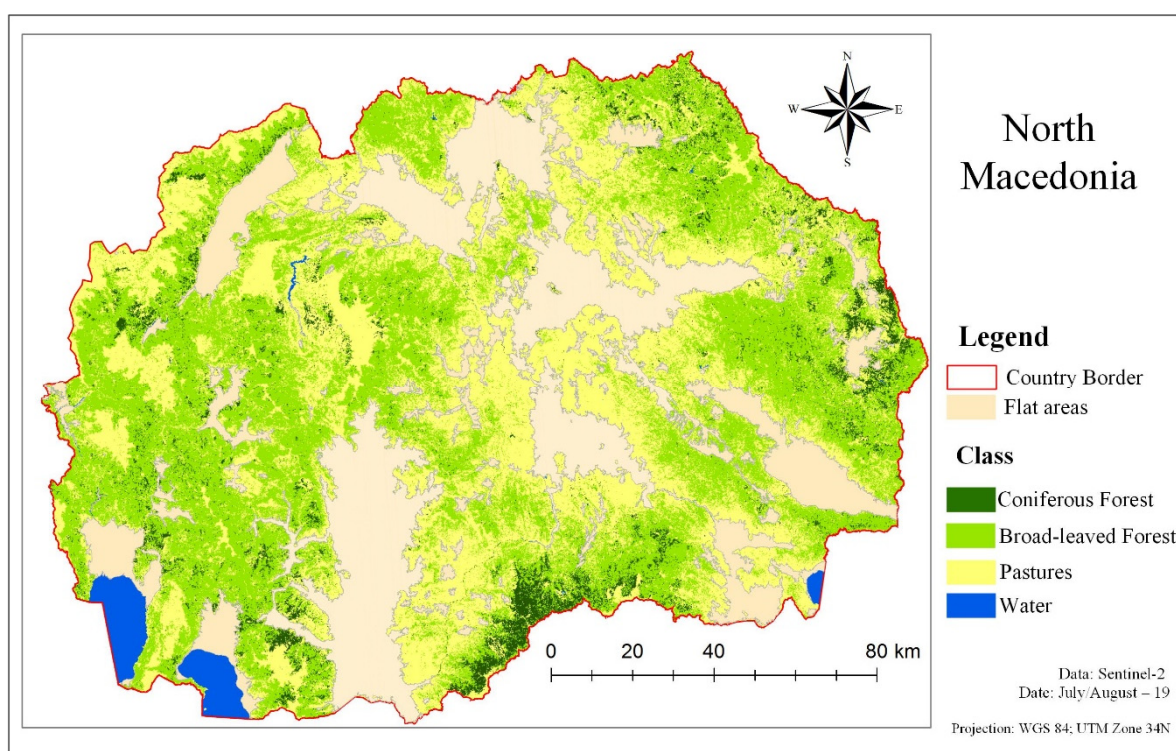


**Figure 3.** Land cover classification results of non-flat areas in North Macedonia.

According to the results, 1,015,526 ha, or approximately 40% of the country, is covered with forest. With 103,667 ha, the coniferous forest covers 11.36% of the total forest area. These results are supported by the national statistical report, where 10% of the total forest cover was confers.

**Table 1.** Classification results represented in ha.

| Class | Area (ha) | Percentage of the Study Area (%) |
|---|---|---|
| Coniferous Forest | 103,667 | 5.3 |
| Broad-leaved forest | 911,859 | 46.6 |
| Pastures | 898,492 | 45.9 |
| Water | 43,287 | 2.2 |
| Total | 1,957,305 | 100 |

In order to assess the accuracy of the classification, half of the samples were used for training while half of the samples were used for accuracy assessment. Both validation overall accuracy, showing the correct classification of the collected samples, and kappa statistics coefficient, showing the accuracy of the classification, have been calculated. While the validation overall accuracy showed 96% fitting of the model, the kappa statistics

for the classification showed a high accuracy of 94%. In addition, a visual comparison was done between the results and satellite imagery form the early spring months, when broad-leaved forests are not still developed and the visibility of the confers is higher, and this comparison also showed good correlation.

The results of this paper are significant since previous work showed high accuracy in classifying different types of forest, confers and broad-leaved, using high-resolution imagery and multi-temporal imagery of additional data to the multi-spectral bands. In this study, apart from the multi-spectral data, slope analysis has been used, based on the study area knowledge. The main purpose of this was to exclude the crop and urban area in order to lower the number of classes, since the main aim of this paper was to separate confers from broad-leaved forests. However, the author believes that similar accuracy results would be achieved if samples from the previously mentioned classes had been collected.

### 4. Conclusions

This study aimed at mapping and classifying coniferous and broad-leaved forest using Sentinel-2 imagery integrated in GEE, using machine learning algorithms such as SVM. This study aimed at high-accuracy classification using minimal knowledge of the study area and thus minimal sample collection. From the presented results, we conclude the following:

- The integration of Senitnel-2 within GEE can be successful for classifying different types of forest, namely, conifers and broad-leaved forest.
- Single image, or image collection from a single season, is sufficient for the accurate classification of different forest types.
- The LIBSVM classifier performs correctly with minimal sample collection over large areas.

The presented study demonstrates the capability of 10 m Sentinel-2 image data to discriminate two main forest types. For future studies, we recommend including more classes to the classification, investigating different seasons, and different forest types.

**Conflicts of Interest:** The author declares no conflict of interest.

### References

1. Päivinen, R.; Van Brusselen, J.; Schuck, A. The growing stock of European forests using remote sensing and forest inventory data. *Forestry* **2009**, *82*, 479–490.
2. Puletti, N.; Camarretta, N.; Corona, P. Evaluating EO1-Hyperion capability for mapping conifer and broadleaved forests. *Eur. J. Remote Sens.* **2016**, *49*, 157–169.
3. Jayathunga, S.; Owari, T.; Tsuyuki, S. Evaluating the performance of photogrammetric products using fixed-wing UAV imagery over a mixed conifer–broadleaf forest: Comparison with airborne laser scanning. *Remote Sens.* **2018**, *10*, 187.
4. Huesca, M.; García, M.; Roth, K.L.; Casas, A.; Ustin, S.L. Canopy structural attributes derived from AVIRIS imaging spectroscopy data in a mixed broadleaf/conifer forest. *Remote Sens. Environ.* **2016**, *182*, 208–226.
5. Nagai, S.; Saitoh, T.M.; Noh, N.J.; Yoon, T.K.; Kobayashi, H.; Suzuki, R.; Nasahara, K.N.; Son, Y.; Muraoka, H. Utility of information in photographs taken upwards from the floor of closed-canopy deciduous broadleaved and closed-canopy evergreen coniferous forests for continuous observation of canopy phenology. *Ecol. Inform.* **2013**, *18*, 10–19.
6. Dostálová, A.; Hollaus, M.; Milenković, M.; Wagner, W. Forest area derivation from sentinel-1 data. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 227–233.
7. Kaplan, G. Sentinel-2 Pan Sharpening—Comparative Analysis. *Proceedings* **2018**, *2*, 345.
8. Kaplan, G. Assessing the effectiveness of PlanetScope synthesized panchromatic bands for spatial enhancement of Sentinel-2 data. *J. Appl. Remote Sens.* **2020**, *14*, 036504.
9. Kaplan, G.; Avdan, U. Evaluating the utilization of the red edge and radar bands from sentinel sensors for wetland classification. *Catena* **2019**, *178*, 109–119.
10. Wessel, M.; Brandmeier, M.; Tiede, D. Evaluation of different machine learning algorithms for scalable classification of tree types and tree species based on Sentinel-2 data. *Remote Sens.* **2018**, *10*, 1419.
11. Persson, M.; Lindberg, E.; Reese, H. Tree species classification with multi-temporal Sentinel-2 data. *Remote Sens.* **2018**, *10,* 1794.

12. Grabska, E.; Hostert, P.; Pflugmacher, D.; Ostapowicz, K. Forest stand species mapping using the Sentinel-2 time series. *Remote Sens.* **2019**, *11*, 1197.

13. Zhou, X.; Li, L.; Chen, L.; Liu, Y.; Cui, Y.; Zhang, Y.; Zhang, T. Discriminating urban forest types from Sentinel-2A image data through linear spectral mixture analysis: A case study of Xuzhou, East China. *Forests* **2019**, *10*, 478.

14. Schmid, J. *Using Google Earth Engine for Landsat NDVI Time Series Analysis to Indicate the Present Status of Forest Stands*; Georg-August-Universität Göttingen: Basel, Switzerland, 2017.

15. Koskinen, J.; Leinonen, U.; Vollrath, A.; Ortmann, A.; Lindquist, E.; d'Annunzio, R.; Pekkarinen, A.; Käyhkö, N. Participatory mapping of forest plantations with Open Foris and Google Earth Engine. *ISPRS J. Photogramm. Remote Sens.* **2019**, *148*, 63–74.

16. Chen, B.; Xiao, X.; Li, X.; Pan, L.; Doughty, R.; Ma, J.; Dong, J.; Qin, Y.; Zhao, B.; Wu, Z. A mangrove forest map of China in 2015: Analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. *ISPRS J. Photogramm. Remote Sens.* **2017**, *131*, 104–120.

17. Brovelli, M.A.; Sun, Y.; Yordanov, V. Monitoring Forest Change in the Amazon Using Multi-Temporal Remote Sensing Data and Machine Learning Classification on Google Earth Engine. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 580.

18. Milevski, I. Morphometry and Land Use on High Mountains in the Republic of Macedonia. In *Sustainable Development in Mountain Regions*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 67–78.

19. Amani, M.; Brisco, B.; Afshar, M.; Mirmazloumi, S.M.; Mahdavi, S.; Mirzadeh, S.M.J.; Huang, W.; Granger, J. A generalized supervised classification scheme to produce provincial wetland inventory maps: An application of Google Earth Engine for big geo data processing. *Big Earth Data* **2019**, *3*, 378–394.