Mapping riparian forest species for biodiversity analysis by fusion of Airborne LiDAR and multispectral satellite imagery

Houssem Njimi [1], Nesrine Chehata [2], Frédéric Revers [3]
[1]UR17DN01 Aviation school of Borj El Amri, Tunisia
Email:houssemnjimi@gmail.com

[2] EA Géoressources & Environnement, Université Michel Montaigne/Bordeaux INP
Email:nesrine.chehata@ensegid.fr

[3] UMR BIOGECO, INRAE
Email:frederic.revers@inrae.fr

1. Introduction

In order to ensure sustainable management of forest resources, the study of the functioning and dynamics of forest resources is essential. Multispectral and 3D LiDAR remote sensing data sources are valuable tools for understanding the relationship between forest structure, biodiversity and microclimate. Remote sensing metrics can be derived either on 3D vegetation structures, biophysical variables or on forest species mapping.

This study focuses on mapping riparian forest species in the canopy strata using a fusion of Airborne LiDAR data and multispectral multi-sources satellite imagery; Sentinel-2 and Pleiades at tree level. The idea is to assess the contribution of each data source in the tree species classification at the considered level and to have first interpretations on relationships with biodiversity taxons (herbaceous, terrestrial invertebrates et vertebrates) which were sampled on each site. Indeed, tree specie composition and mapping are known to modify biodiversity responses.

The fusion was processed at feature-level and decision level. At feature level, LiDAR 2D attributes were derived and combined with vegetation indices. At decision level, LiDAR data was used for 3D tree crown delimitation providing a unique tree or a group of trees that are used as a support for the species classification. Data augmentation techniques were used to increase training samples.

Best results were obtained by the fusion of Sentinel-2 time series and LiDAR data with a Kappa of 0.656 thanks to red-edge based indices that better discriminate vegetation species and the temporal resolution of Sentinel-2 images that allows monitoring the phenological stages helping discriminate the species.

2. Data and methods

2.1 Study Area

The study site is located in south western France. Ciron watershed and its riparian forest is an affluent of the Garonne known for a climatic refuge for beech, on the warm margin of its European range. This riparian forest is made up of an assemblage of species such as Oak, Beech, Locust, Pine, etc. Twenty-eight sites forming a three-dimensional structure gradient are defined. They are located over thirty kilometers along the Ciron and 5 km along an affluent in which the riparian forest is bordered by pine forests (maritime pine) in order to homogenize the potential impact of the surrounding landscape on the biodiversity of flora and fauna in the riverine forest.

2.2 Data

Joint airborne acquisition and in-field observations were conducted in autumnal season with tree foliage on 3rd and 4th October 2019. Trees were measured on canopy and shrub strata leading to more than 31 unbalanced classes. In this study we only focus on canopy classes that are limited to 5 classes. Train and Test data were selected using individual tree crowns generated after the segmentation process. They consist of 165 and 73 samples, respectively.

Complementary data were used exploiting spectral and temporal information (Sentinel-2), very high resolution (VHR) imagery (Pleiades) and geometric information from LIDAR data. 11 Sentinel-2 images were used from January to December 2019, one VHR Pleiades image on the 3rd September 2019 and 632 tiles of 3D point clouds with 50pts\m² density covering the whole site.

Published in: Markus Hollaus, Norbert Pfeifer (Eds.): Proceedings of the SilviLaser Conference 2021, Vienna, Austria, 28–30 September 2021.
Technische Universität Wien, 2021. DOI: 10.34726/wim.1861
This paper was peer-reviewed. DOI of this paper: 10.34726/wim.2009

2.3 Methodology

LiDAR and multispectral data were first co-registered manually exploiting the high accuracy of Lidar data. Each data source was processed separately and then fused at feature and decision levels [1].
First, DSM and DTM were processed at a resolution of 0.25 m. A DHM was then derived leading to a map of canopy height as well as 2D attributes such as Intensity, point density, number of echoes and height range.

3D Lidar point clouds were segmented using PyCrown [2] method. It provides a 3D segmentation of individual trees besides a raster segmentation. It is based on local maximum search (i.e. tree tops) and region growing with regard to user-defined parameters (distance of a crown point from its top and point height w.r.t crown average heights).

11 Sentinel-2 images were used in TOA reflectance. 10 spectral bands of 10-20 m were used and resampled to 0.25 m. Besides 7 vegetation indices [3] were derived for each date based on near infrared and red edge channels: NDVI, GRVI1, CIre, NDVIre3, NDre2, SAVI, MSAVI2. Soil adjusted indices such as SAVI and MSAVI2 were used to better handle non dense tree species. Totally, 10 initial bands and 7 vegetation indices were used per date leading to 187 spectral bands.

Fusion was first processed at feature level by concatenating spectral and geometric LiDAR attributes. Different feature combinations were used to assess the importance of spectral, temporal or spatial information. At decision-level, segmented LiDAR regions were used to derive spectral attributes at an object-level using attributes' mean and standard deviation over each segmented region. Due to few training data, data augmentation techniques were used using Gaussian Noise filtering increasing the samples per twice. Finally, the classification was processed using a Random Forest classifier. Results are evaluated using Overall accuracy, kappa and per class precision and recall.

3. Results and Discussions

Table 1 resumes the obtained classification accuracies with different fusion configurations and measuring data augmentation impact. Table 2 presents the precision and recall values per specie.

Table 1. Comparison of classification accuracies using different fusion configurations with and without data augmentation

	Classification	Kappa	OA
Without data augmentation	Sentinel-2 (single date) + LIDAR	0.481	0.594
	Pleiades + LIDAR	0.434	0.548
	Sentinel-2 (multi dates) + LIDAR	**0.508**	**0.607**
	Sentinel-2 (single date) + Pléiades + LIDAR	0.493	0.595
With data augmentation	Sentinel-2 (mono-date) + LIDAR	0.527	0.620
	Pleiades + LIDAR	0.495	0.598
	Sentinel-2 (multi-dates) + LIDAR	**0.656**	**0.694**
	Sentinel-2 (mono-date) + Pléiades + LIDAR	0.582	0.665

Table 2. Comparison of precision and recall per specie using data augmentation

	Pedonculate Oak	Tauzin Oak	Black Alder	Maritime Pine	Other
Precision	0.47	0.91	0.54	0.73	0.85
Recall	0.56	0.71	0.89	0.8	0.48

The best results are given with the combination of multi dates Sentinel-2 images and LiDAR data showing the importance of temporal and spectral resolution of Sentinel-2 which contributed the most in the classification of forest species.

We actually obtained 0.656 as Kappa and 0.694 as Overall Accuracy. Otherwise, the very high spatial resolution of Pleiades (2m) caused a decrease in the evaluation metrics. This is probably due to detected shadows.

The data augmentation improved the obtained results as it allowed more training samples. The best configuration has also shown that *Tauzin Oak*, *Maritime Pine* and the class *Other* have precision values superior to 0.727 which means that at least 72.7% of these species where correctly labeled. Otherwise, less than 54.2% of *Pedonculate Oak* and *Black Alder* were correctly labeled.

Results have also shown that the classifier under-estimates the *Pedonculate oak* and the class *Other* (low recall values compared to other classes) while it over estimates the *black alder* (high recall value 0.88 and low precision value 0.54)

Figure 1 shows the 2D segmentation result and the corresponding forest species classification using the best fusion configuration on Site 21.

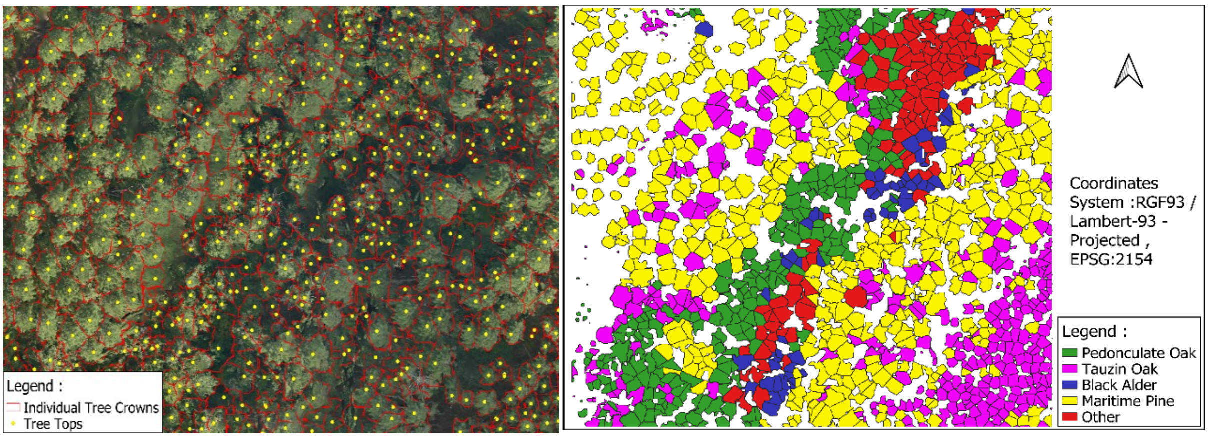

Figure 1: Results on Site 21 : From left to right : Individual Tree Crowns Delineation , Forest Species Classification

4. Conclusions

This study allowed the evaluation of Airborne LiDAR data and multispectral satellite imagery fusion in order to classify riparian forest species. The results revealed the importance of data augmentation, temporal and spectral resolution of Sentinel-2 satellite in the process of classification added to the importance of LiDAR data in the individual tree crown delineation. The best fusion configuration gave respectively 0.656 and 0.694 as kappa and OA values respectively.

Further work will focus on providing spatial metrics from species patterns and measuring their relationships with biodiversity taxons.

Acknowledgements

Authors want to thank the French National Space Agency for their support through the TOSCA FRISBEE project.

References

[1] E. Tusa, A. Laybros, J.-M. Monnet, M. Dalla Mura, J.-B. Barré, G. Vincent*, et al.*, "Fusion of hyperspectral imaging and LiDAR for forest monitoring," *Data Handling in Science and Technology,* vol. 32, pp. 281-303, 2020.

[2] M. Dalponte and D. A. Coomes, "Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data," *Methods in ecology and evolution,* vol. 7, pp. 1236-1245, 2016.

[3] N. Touihri, N. Chehata, H. Chakroun, and T. S. Chahed, "Contribution of Sentinel 2 images spectral bands for mapping forest land cover: Application to Tbeynia site, Jendouba, Tunisia."