

FUSION OF LIDAR AND HYPERSPECTRAL DATA FOR SEMANTIC SEGMENTATION OF FOREST TREE SPECIES

E. Tusa^{1,2,3}*, J. M. Monnet¹, J. B. Barré¹, M. Dalla Mura^{2,4}, J. Chanussot²

¹Univ. Grenoble Alpes, INRAE, LESSEM, F-38402 St-Martin-d'Hères, France - {eduardo.tusa-jumbo, jean-matthieu.monnet, jean-baptiste.barre}@inrae.fr

²Univ. Grenoble Alpes, CNRS, Grenoble INP⁺, GIPSA-Lab, 38000 Grenoble, France.

⁺Institute of Engineering Univ. Grenoble Alpes - {jocelyn.chanussot, mauro.dalla-mura}@grenoble-inp.fr

³Universidad Técnica de Machala, Facultad de Ingeniería Civil, AutoMathTIC, 070210 Machala, El Oro, Ecuador.

⁴Tokyo Tech World Research Hub Initiative (WRHI), School of Computing, Tokyo Institute of Technology, Tokyo, Japan.

Commission III, WG III/4

KEY WORDS: Hyperspectral images, LiDAR, robust PCA, Semantic segmentation, Temperate forests

ABSTRACT:

Hyperspectral images (HI) and Light Detection and Ranging (LiDAR) provide high resolution radiometric and geometric information for monitoring forests at individual tree crown (ITC) level. It has many important applications for sustainable forest management, biodiversity assessment and healthy ecosystem preservation. However, the integration of different remote sensing modalities is a challenging task for tree species classification due to different artifacts such as the lighting variability, the topographic effects and the atmospheric conditions of the data acquisition. The characterization of ITC can benefit from the extraction and selection of robust feature descriptors that solve these issues. This paper aims to investigate the integration of feature descriptors from HI and LiDAR by using the intra-set and inter-set feature importance for the semantic segmentation of forest tree species. A fusion methodology is proposed between high-density LiDAR data - (20 pulses m^{-2}) and VNIR HI - (160 bands and 0.80 m spatial resolution) acquired on French temperate forests along an altitude gradient. The proposed scheme has three inputs: the field inventory information, the HI and the LiDAR data. Our approach can be described in nine stages: polygon projection, non-overlapping pixel selection, vegetation and shadow removal, LiDAR feature extraction, height mask, robust PCA (rPCA), feature reduction and classification. The overall accuracy of tree species classification at pixel-level was 68.9% by using random forest (RF) classifier. Our approach showed that 74.0% of trees were correctly assigned overall, by having conifer species such as Norway Spruce (*Picea abies*) with a producer's accuracy of 97.4%.

1. INTRODUCTION

Spatial distribution of forest tree species has important benefits for sustainable forest management (Ghosh et al., 2014). For instance, scientists are able to study the functioning of forested ecosystems by understanding variables associated to stress, disease patterns, invasive species spread and deforestation (Ghosh et al., 2014, Lee et al., 2016). This information is very relevant to establish useful exploitation policies of forests (Dalponte et al., 2012). Conventional ecological survey methods of tree species mapping require an exhaustive work that thrives in difficult scenarios, relies on small plot-level datasets and implies a significant amount of time, manpower and economic resources (Ghosh et al., 2014, Lee et al., 2015).

Several studies have highlighted the potential of spectral and spatial resolution in remote sensing based tools for monitoring forest ecosystems. For instance, hyperspectral sensors extract radiance information from objects or scenes lying on the Earth surface (Bioucas-Dias et al., 2013) by covering the visible (V), near-infrared (NIR), and shortwave infrared spectral bands in the range from 300 to 2500 nm (Bioucas-Dias et al., 2012). LiDAR data describe the scene using a point cloud that has explicit 3D coordinates (x, y, z), intensity of the returns, return number, number of returns, point classification, among other attributes (Vincent et al., 2017). The advances in the integration of LiDAR and hyperspectral sensors make possible to identify tree species at pixel-level with a high accuracy.

Data fusion is implemented at three levels: observation-level, feature-level and decision-level (Tusa et al., 2020). Considering that tree species identification is a remote sensing topic posed as a supervised approach (Dalponte et al., 2019), recent studies pursue to integrate HI and LiDAR data to improve the classification (decision-level fusion). In (Matsuki et al., 2015), features are extracted from the HI by applying principal component analysis (PCA) for reducing the redundancy within the bands (Liao et al., 2017). In (Lee et al., 2016), the algorithm rPCA is applied for filtering the noise and for selecting relevant features. This study demonstrated that rPCA improved the classification of six tree species over PCA, with an overall accuracy of 61.0% at tree-level. The algorithms support vector machines (SVM) (Lee et al., 2016) and RF (Maschler et al., 2018, Anderson, 2018) have been used for tree species classification with HI. Although SVM performs superior for high dimensional features, it requires a proper parameter-setting that can be time-consuming for large datasets. Alternatively, RF can be a good choice for feature selection (feature-level fusion) and classification (Tusa et al., 2014).

The aim of this study is to integrate HI and LiDAR data for the classification of four forest tree species. The matrix decomposition of rPCA increases the feature representation, which can be reduced by considering the intra-set and inter-set feature importance. This work is divided as follows: section 2. describes the study site and the data acquisition specifications, section 3. presents the contributions based on PCA and rPCA, feature selection and classification based on RF, section 4. discusses the results of feature fusion for the classification and finally, section 5. corresponds to the conclusions and the work perspectives.

*Corresponding author.

2. STUDY SITE

The study area corresponds to the site of Chamrousse, located in Belledonne massif, Northern Alps, France. Trees with a diameter at breast height (DBH) over 7.5 cm were inventoried on seven field plots: four of size of $50 \times 50\text{ m}^2$, two circular plots of 15 m radius, and a plot of $80 \times 100\text{ m}^2$, which is part of a long-term inventory dataset (Fuhr et al., 2017). The forest is dominated by Norway spruce (*Picea abies*; 37.5%) and other conifers: silver fir (*Abies alba*; 31.6%) and mountain pine (*Pinus uncinata*; 10.6%); broadleaves species are mainly represented by European beech (*Fagus sylvatica*; 15.2%).

For tree positions, slope, distance and azimuth relatively to reference poles are recorded. Tree crown extensions were measured during the summer 2018, with ruban tapes in the north, south, east and west directions as the horizontal distance between the trunk center and the vertical projection of the furthest live branch along that direction. Tree positions in the Lambert 93 projected coordinate system are then computed from the angular and distance measurements. As GNSS precision is around a few meters under forest cover, trees positions are displayed over the LiDAR canopy height model. In case of a significant shift, position of the reference pole is adjusted by using the co-registration method described in (Monnet and Mermin, 2014). The field data is summarized in Table 1 by considering seven plots with 893 tree crowns.

The remote sensing data was collected between 21st and 23rd of June 2018. LiDAR acquisition was carried by using a RIEGL LMS Q780 sensor. The 3D point cloud has a mean pulse density of 20 pulses m^{-2} . The HI was collected by using a Hyspex VNIR 1600 sensor. It has 160 bands with a spatial resolution of 0.80 m and a spectral resolution of 4.5 nm in a spectral range between 400 and 1000 nm .

Level	Field measurement	Min.	Max.	Mean
Tree	DBH [cm]	7.5	104.5	26.2
	Tree Height [m]	2.0	38.4	14.8
	Crown area [m ²]	0	161.9	23.3
Plot	Average DBH [cm]	16.3	31.2	25.8
	Number of stems/ha	460	980	640
	BA [m ² /ha]	13.0	64.5	40.8

Table 1: Summary of the field measurements at tree- and plot-level. Minimum (Min), Maximum (Max) and Mean are displayed. BA refers to basal area.

3. METHODOLOGY

The semantic segmentation workflow for forest tree species classification is presented in Figure 1. The proposed scheme has three inputs: the field inventory information, the HI and the LiDAR data. This approach comprises the following steps: polygon projection, non-overlapping pixel selection, spatial filtering, vegetation and shadow mask, LiDAR features, height mask, robust PCA, feature reduction and classification.

3.1 Polygon projection

Each tree crown from the inventory is projected into the xy -plane by fitting an ellipsoidal crown shape. Then, a HI pixel is associated with a crown by verifying if the pixel center coordinate is inside the crown polygon. The pixel center coordinate, (x_c, y_c) , is computed through equations (1)

$$\begin{aligned} x_c &= (2c_i + 1)s_f + x_0 \\ y_c &= y_0 - (2r_i + 1)s_f \end{aligned} \quad (1)$$

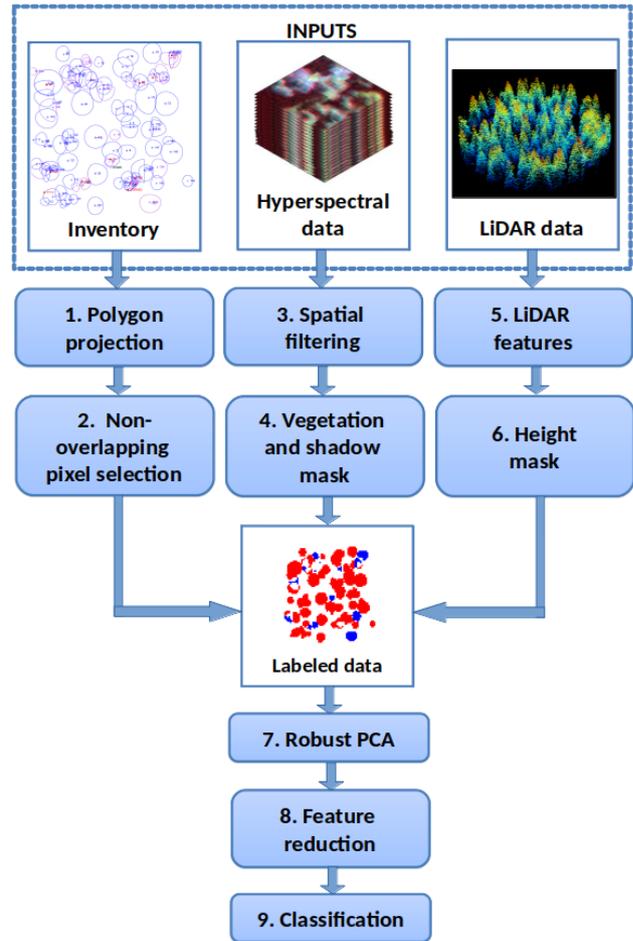


Figure 1: Semantic segmentation flowchart for forest tree species classification by integrating HI and LiDAR data.

where (x_0, y_0) is the xy coordinates for the upper-left pixel in the image, s_f is the half of the spatial resolution of the image, and (r_i, c_i) are the coordinates for the row and the column pixel positions in the image ($r_i \in \{0, \dots, m\}$ and $c_i \in \{0, \dots, n\}$ with an image size $m \times n$).

3.2 Spatial filtering

According to (Laybros et al., 2019), a strategy for reducing local noise and improving class separability can be implemented through the spatial filtering. Conventionally, a mean filtering is computed by moving a windows of 3×3 over each band. In (Dechesne et al., 2017), three circular neighborhoods of radius $r_f \in \{1.0, 3.0, 5.0\}\text{ m}$ are applied to compute the mean value of every pixel in each band. Since the pixel center coordinate (x_c, y_c) is known, the pixels whose centers are inside of the circular neighborhood of each pixel, are selected for averaging the reflectance in each band. Finally, the three mean values of all pixel associated to each r_f are averaged.

3.3 Non-overlapping pixel selection

The crown pixel correspondence becomes challenging when some crown regions overlap among each other. To overcome this issue, those overlapping pixels that are associated with crown trees of different species, are not considered in our ground truth. The purpose is to select a representative set of pure pixels for each species. The overlapping pixels for those trees of the same species, are labeled to the tree with the greatest height and crown area.

This pixel - tree association is relevant when the training and testing sets are defined for the classification.

3.4 Vegetation and shadow mask

The selection of sunlit pixels is advised because shadowed pixels are affected from a low signal-to-noise ratio (SNR) (Torabzadeh et al., 2019). Several strategies have been proposed for shadow removal. In our approach, the criterion explained by (Dalponte et al., 2014), estimated the average of the blue portion of the spectrum in the range of [450, 550] nm. Then, Otsu threshold is calculated over this range. Since the Otsu threshold value removed approximately 50.0% of the pixels, the minimum threshold is computed to preserve around 65.0% of our ground truth.

The normalized difference vegetation index (NDVI) is used to separate forest from non-vegetation (Laybros et al., 2020). In (Dalponte et al., 2018), authors selected a NDVI value greater than 0.5 to consider well-lit, leafy vegetation pixels. In (Alonzo et al., 2016), they used a threshold of 0.6 based on the average NDVI for all crowns. For our approach, a NDVI value of 0.55 is applied for all forest plots.

3.5 LiDAR features

From the LiDAR 3D point cloud, 24 LiDAR features are computed at point-level according to (Dechesne et al., 2017):

- Two vegetation density features: the cumulative sum of local height maxima retrieved from each radius r_f maximum filter within three cylindrical neighborhoods of radius $r_c \in \{1.0, 3.0, 5.0\} m$. The second feature is computed from the number of points classified as ground points over the total number of points within each cylindrical neighborhood of radius r_c . The three outputs of the neighborhoods for every point are averaged.
- Two shape features: the scatter and the planarity are computed from the eigenvalues of the covariance matrix within the three cylindrical neighborhoods of radius r_c .
- 20 statistical features from each LiDAR point using the cylindrical neighborhoods given in r_c . The height and intensity information from the LiDAR data are used to derive statistical features: minimum; maximum; mean; median; standard deviation; median absolute deviation from median; mean absolute deviation from median; mean absolute deviation from mean; skewness; kurtosis; 10th, 20th, 30th, 40th, 50th, 60th, 70th, 80th, 90th and 95th percentiles. All the statistical functions are used for the height, while the mean is used for the intensity only.

After computing the LiDAR features at point-level, these features are rasterized at the resolution of the HI, s_c . The feature values of the points inside the pixel that is centered at the coordinate $\mathbf{x} = (x_c, y_c)$, are weighted according to the function $W(\mathbf{x})$

$$\begin{aligned} W(\mathbf{x}_i) &= W_h(\mathbf{x}_i)W_v(\mathbf{x}_i) \\ W_h(\mathbf{x}_i) &= \exp\left(-\gamma\left\|\frac{\mathbf{x}-\mathbf{x}_i}{w_h}\right\|^2\right) \\ W_v(\mathbf{x}_i) &= \left\|\frac{Z_i-Z_{min}}{w_z}\right\| \end{aligned} \quad (2)$$

where W_h and W_v are the weighting functions of the horizontal and vertical information, respectively. The point \mathbf{x}_i represents the

xy coordinate of each 3D point inside the cylinder, γ is selected to be 0.5 (normal distribution), w_h is an horizontal bandwidth equal to $\sqrt{2}s_c$, Z_i represents the z coordinate of each 3D point inside the pixel and $w_z = Z_{max} - Z_{min}$, corresponds to the difference between the maximum and minimum height 3D point values in the vertical weighting function, W_v . These weighting functions have been used for finding the maximum value of the 3D point cloud distributions for 3D segmentation purposes (Xiao et al., 2019). In this way, the information of the highest points close to the pixel center are given more importance because these are associated to the most illuminated regions in the image.

3.6 Height mask

The purpose of the height mask is to avoid negative effects due to shrubs, grassland and low height objects. A threshold value of $h_{min} = 1.5 m$ (Kandare et al., 2016) is applied over the digital surface model (DSM) for preserving most of the forest regions. At this point, by merging all the criteria of specie purity, pixel illumination and height information, a set of n_{lp} labeled pixels associated with $n_{HI} = 160$ hyperspectral bands and $n_{Li} = 24$ rasterized LiDAR metrics is processed in the next stage.

3.7 Robust PCA

From a given data matrix of size $m \times n$, PCA can generate a low-rank matrix \mathbf{L} that minimizes the error $\mathbf{S} = \mathbf{D} - \mathbf{L}$:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{S}\|_F, \text{ s.t. rank}(\mathbf{L}) \leq r, \mathbf{D} = \mathbf{L} + \mathbf{S} \quad (3)$$

where $r \ll \min(m, n)$ is the target rank of \mathbf{D} and $\|\cdot\|_F$ is the Frobenius norm. Although PCA is widely used as dimensionality reduction technique, it can be affected by large-amplitude noise. rPCA (Lee et al., 2016) is a method proposed to improve the robustness of the PCA for dealing with outliers. This algorithm pursues to recover a low-rank matrix \mathbf{L} from highly noisy measurements $\mathbf{D} = \mathbf{L} + \mathbf{S}$, by separating a sparse matrix \mathbf{S} . Given the matrix \mathbf{D} , the matrices \mathbf{L} and \mathbf{S} are computed by satisfying the following condition:

$$\min_{\mathbf{L}, \mathbf{S}} \text{rank}(\mathbf{L}) + \lambda \|\mathbf{S}\|_0, \text{ s.t. } \mathbf{L} + \mathbf{S} = \mathbf{D} \quad (4)$$

where $\|\cdot\|_0$ is the l_0 -norm, λ is a regularization parameter for balancing the importance between the ranking operator and the sparsity regularization. Equation (4) is reformulated as a convex optimization approach:

$$\min_{\mathbf{L}, \mathbf{S}} \|\mathbf{L}\|_* + \lambda \|\mathbf{S}\|_1, \text{ s.t. } \mathbf{L} + \mathbf{S} = \mathbf{D} \quad (5)$$

where $\|\cdot\|_*$ is the nuclear norm, which is the sum of singular values of \mathbf{L} ; and $\|\cdot\|_1$ is the l_1 -norm. In (Brunton and Kutz, 2019), authors described the programming implementation of the rPCA algorithm. In this stage, the matrices \mathbf{L} and \mathbf{S} are obtained from the HI, \mathbf{D}_{HI} , and the rasterized LiDAR features, \mathbf{D}_{Li} , separately.

3.8 Feature reduction

Considering the findings of (Luan et al., 2014), the sparse matrix \mathbf{S} contains detailed information to establish inter-class and intra-class differences. This is a motivation to focus on the effect of stacking features according to the feature importance given by the RF classifier (Friedman, 2001). Let \mathbf{L}_{HI} and \mathbf{S}_{HI} be the low-rank and sparse matrices from HI, \mathbf{D}_{HI} ; the principal components (PC) are computed to obtain: $\mathbf{PC}_{HI}^D, \mathbf{PC}_{HI}^L$ and \mathbf{PC}_{HI}^S . The same procedure is repeated with the LiDAR information to have six matrices: $\mathbf{D}_{Li}^D, \mathbf{L}_{Li}^L$ and \mathbf{S}_{Li}^S ; and the corresponding PC: $\mathbf{PC}_{Li}^D, \mathbf{PC}_{Li}^L$ and \mathbf{PC}_{Li}^S .

The feature selection was carried out by calculating the RF scores to estimate the intra-set feature importance. Let \mathbf{A} be a feature set of 5 features $[\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5]$, then, the RF scores, s_i , are computed for the i -th feature to obtain $[s_1, s_2, s_3, s_4, s_5]$, where $s_4 > s_3 > s_1 > s_5 > s_2$. In this way, the features are ordered in decreasing magnitude of importance by creating a sorted set of features $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{a}_5, \mathbf{a}_2]$. By considering this order, each feature forms a subset together with the features that have higher importance:

- $[\mathbf{a}_4]$
- $[\mathbf{a}_4, \mathbf{a}_3]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{a}_5]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{a}_5, \mathbf{a}_2]$

Each subset is evaluated in the RF classifier and the subset with the highest F1-score corresponds to the selected feature set. This procedure is applied for every set of features: $\mathbf{D}_{HI}, \mathbf{L}_{HI}, \mathbf{S}_{HI}, \mathbf{PC}_{HI}^D, \mathbf{PC}_{HI}^L, \mathbf{PC}_{HI}^S, \mathbf{D}_{Li}, \mathbf{L}_{Li}, \mathbf{S}_{Li}, \mathbf{PC}_{Li}^D, \mathbf{PC}_{Li}^L$ and \mathbf{PC}_{Li}^S .

The 12 sets of selected features are ordered based on the F1-score as it is described in Table 4. The procedure to stack features from a different set to another is explained as follows. The set with the highest F1-score is our reference set to which new features will be added to increase the overall F1-score. Let $\mathbf{A} = [\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1]$ be our reference set with the highest F1-score, the feature set $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4]$ contains the features to be added to the set \mathbf{A} . The estimation of the inter-set feature importance is achieved by stacking every feature to the reference set:

- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_1]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_2]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_3]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_4]$

Each subset is evaluated in the RF classifier to obtain the F1-scores: $[F1_1, F1_2, F1_3, F1_4]$, where $F1_3 > F1_4 > F1_1 > F1_2$. By considering this ranking, the features are ordered in decreasing F1-score to be stacked sequentially in the reference set and to generate the following subsets:

- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_3]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_3, \mathbf{b}_4]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_1]$
- $[\mathbf{a}_4, \mathbf{a}_3, \mathbf{a}_1, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_1, \mathbf{b}_2]$

The subset of features that increased the F1-score of the reference set are stacked to form a new reference set. This procedure is repeated with the remaining sets.

3.9 Classification

For the tree species classification, tree crowns were selected according to the number of pixels per crown. Thus, those trees that had more than 4 pixels per crown, were considered in our dataset. By applying this criterion, an average of 19.6 pixels per crown was obtained which is close to the average number of pixels per crown selected by (Baldeck and Asner, 2014) (24.9 pixels per crown). After performing the tree crown selection, the number of trees per specie were distributed according to Table 2.

For the RF classifier, the number of decision trees to grow was set to 1000. The training and test sets were defined by randomly separating trees and by associating the pixels with their crowns. The tree crowns were divided into six sets. The training set was formed by five of these six sets, resulting in a test set of 1550 pixels, which is a similar number of samples selected by (Lee et al., 2016) (1537 pixels).

Our dataset showed a problem of class imbalance, which means the number of trees and pixels per class are not evenly distributed (Anderson, 2018). To address this problem, two strategies were implemented. First, those tree species that had less than six trees or had less than 1% of total number of pixels were grouped into a class called “other” species. Second, RF automatically weighted the classes by considering that these are inversely proportional to the frequency that the class has in the data (Albon, 2018). The pixel distribution is described in Table 3.

Abbr.	Genus species	English name	No.
ABAL	<i>Abies alba</i>	European silver fir	139
ACPS	<i>Acer pseudoplatanus</i>	Sycamore maple	5
BEPE	<i>Betula pendula</i>	Silver birch	5
BEsp	<i>Betula sp.</i>	Birch	2
COAV	<i>Corylus avellana</i>	Common hazel	5
FASY	<i>Fagus sylvatica</i>	European beech	73
FREX	<i>Fraxinus excelsior</i>	European ash	1
PIAB	<i>Picea abies</i>	Norway spruce	190
PIUN	<i>Pinus uncinata</i>	Mountain pine	11
POTR	<i>Populus tremula</i>	European aspen	2
SOAR	<i>Sorbus aria</i>	Common whitebeam	1
SOAU	<i>Sorbus aucuparia</i>	Rowan	3

Table 2: Abbreviation, full name and number of tree crowns per species in the study area. The total number of trees is 437

Abbr.	Average pixels/tree	No. pixel	Pixel [%]	Tree [%]
ABAL	18.4	2258	29.9	31.8
FASY	23.4	1707	20.0	16.7
PIAB	20.0	3808	44.5	43.5
PIUN	8.2	90	1.1	2.5
Other	16.4	393	4.6	5.5
Overall	19.6	8556	100	100

Table 3: Abbreviation, number and percentage of trees and pixels for four identified species and one “other” species class

4. RESULTS AND DISCUSSION

The results of feature stacking based on feature importance and F1-score are presented for the HI and the LiDAR data for the fusion of these two modalities. The ground truth of the seven plots is illustrated in Figure 2. Four classes of tree species were discriminated at pixel-level. The tree species PIAB is present in all forest plots, while the class PIUN is in plot 4 only. The classification at tree-level is obtained by applying a majority voting rule to define the species for each crown (Lee et al., 2016).

The ranking of the feature sets based on the F1-score obtained by pixel classification is described in Table 4. Due to the difference between the amount of features from HI and LiDAR data, the spectral descriptors performed better than those features extracted from the LiDAR point cloud. However, the rasterized LiDAR features were not filtered spatially as the HI bands, which clearly considers the information from the neighborhood of each pixel. Feature selection based on RF importance reduced the amount of features and increased the classification performance. According to Table 4, the reference set for our approach is formed by 10 features from D_{HI} . The complete HI set is represented in Figure 3(a) by computing the mean spectral signature of each tree species considering the ground truth from Figure 2.

Rank	Feature	Original set		Reduced set	
		No.	F1 [%]	No.	F1 [%]
1.	D_{HI}	160	49.7	10	54.5
2.	PC_{HI}^D	160	52.5	100	54.0
3.	S_{HI}	160	51.0	82	52.1
4.	PC_{HI}^S	160	51.4	145	52.1
5.	PC_{HI}^L	160	47.8	23	51.9
6.	L_{HI}	160	48.3	8	51.0
7.	S_{Li}	24	49.7	22	50.0
8.	PC_{Li}^D	24	48.6	21	48.5
9.	L_{Li}	24	46.6	19	47.8
10.	D_{Li}	24	47.0	19	47.0
11.	PC_{Li}^L	24	44.3	10	46.5
12.	PC_{Li}^S	24	44.4	20	45.4

Table 4: Ranking of set of features according to the F1-score (F1) and the number of features (No.) in the original and reduced sets, by applying RF scores of importance

Table 5 shows the F1-scores of pixel classification for each tree species by adding the amount of HI features marked in the second row. Since D_{HI} is the reference set with F1-score of 54.5%, the selection of 20 PC from PC_{HI}^D increased 5.2 points of percentage. The effect is clear in the most abundant classes, conifers, such as PIAB and ABAL. The minority classes such as FASY, PIUN and other species; decreased their performance at the limit that the PIUN class is not detected by the model. The addition of 25 features from S_{HI} , the sparse information from HI, improved for the species ABAL, PIAB and PIUN. Although the 18 PC from HI improved the performance for ABAL, FASY and PIAB species, the 8 PC from the low-rank representation, PC_{HI}^L , improved the F1-scores for all species. The contribution of 2 features from L_{HI} increased in 0.01 points of percentage.

Table 6 shows the F1-scores of pixel classification by adding the amount of LiDAR features marked in the second row. The computation of these LiDAR features was explained in the subsection 3.5 and the mean normalized range of each feature by species can be visualized in Figure 3(b). The results are presented for the datasets S_{Li} , PC_{Li}^D and D_{Li} . The datasets associated to the low-rank representation L_{Li} and the PC: PC_{Li}^L and PC_{Li}^S , decreased the overall F1-score. Hence, these features were not considered for the feature fusion. Although the LiDAR information of these three datasets contributed with 1.2 points of percentage, the performance of the PIUN tree species increased 7.4 points of percentage with respect to the F1-score in the HI datasets. LiDAR data can be very useful for describing minority classes, but our feature fusion strategy gave more importance to the overall result than the individual species performance. In total, 4 features from LiDAR information are integrated to the selected features from HI to obtain a new set of 87 features that are described in Figure 3(c). The overall F1-score for the fusion of the selected features from HI and LiDAR data was 68.9%.

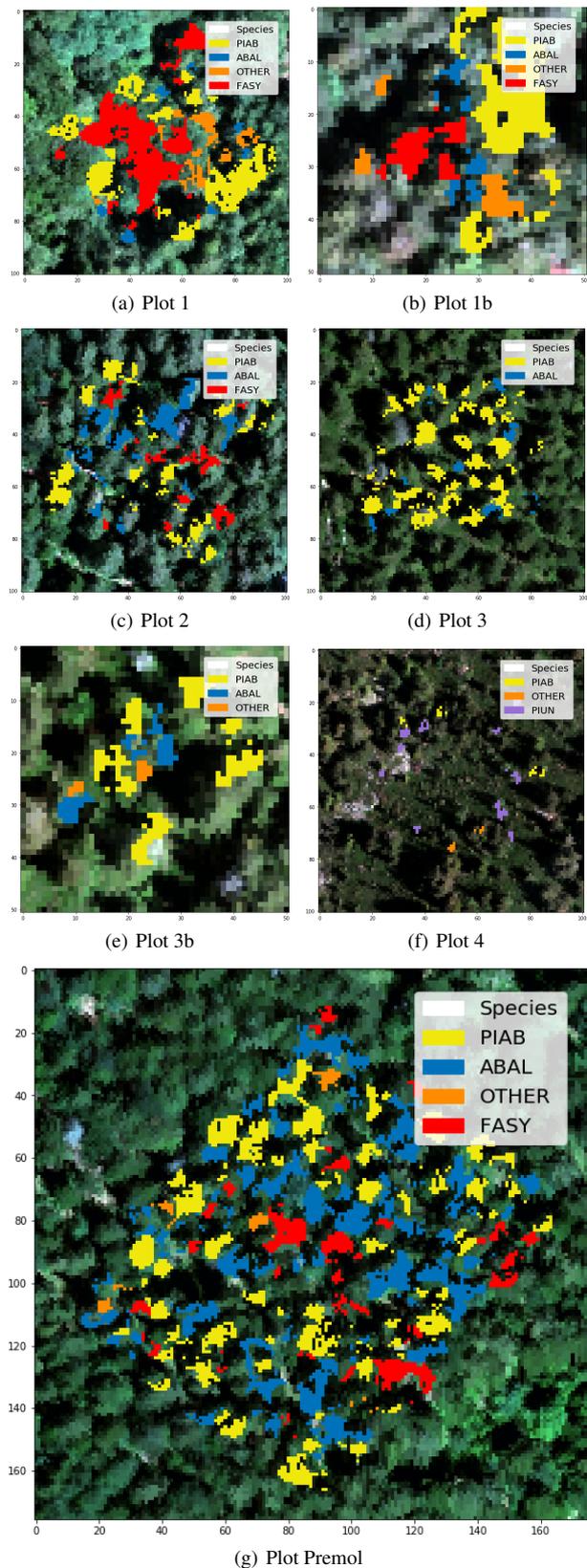
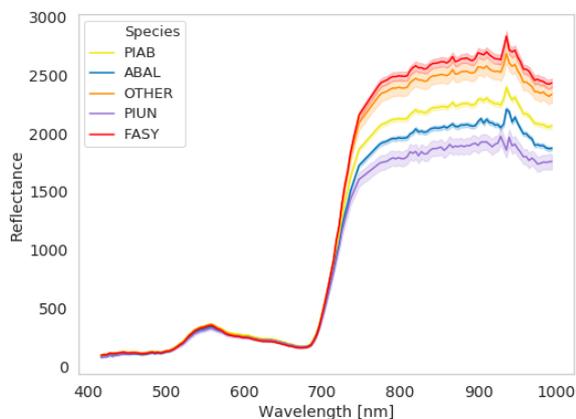
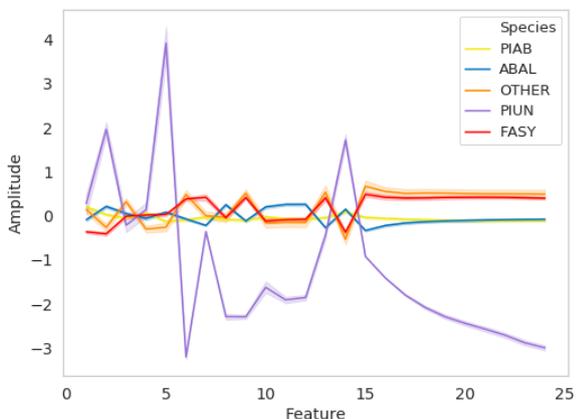


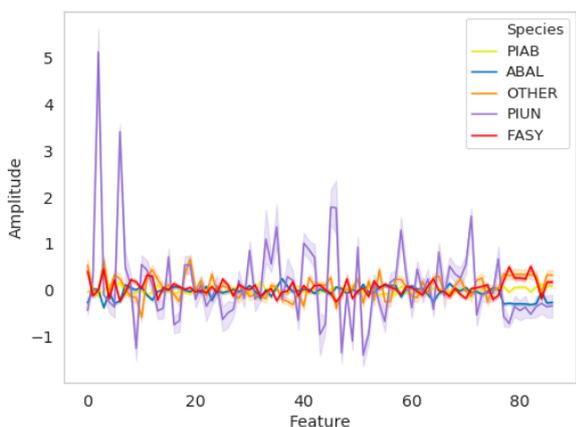
Figure 2: RGB representation of the forest plots corresponding to the Chamrousse site with the following color labels: ● *Abies alba* (ABAL), ● *Fagus sylvatica* (FASY), ● *Picea abies* (PIAB), ● *Pinus uncinata* (PIUN) and ● “other” species



(a) Mean hyperspectral signatures



(b) Mean normalized LiDAR features



(c) Mean normalized selected features

Figure 3: (a) Mean spectral signature of each forest tree species in the range 400 to 1000 nm. (b) Mean normalized LiDAR features of each forest tree species. (c) 87 selected features from HI and LiDAR data from the datasets ranked in Table 4. Legend: ● *Abies alba* (ABAL), ● *Fagus sylvatica* (FASY), ● *Picea abies* (PIAB), ● *Pinus uncinata* (PIUN) and ● “other” species

Specie	D_{HI}	PC_{HI}^D	S_{HI}	PC_{HI}^S	PC_{HI}^L	L_{HI}
No.	10	20	25	18	8	2
ABAL	33.0	43.7	50.3	52.0	55.7	57.9
FASY	53.4	47.8	47.2	48.4	54.7	54.5
PIAB	67.8	71.6	74.3	74.7	77.4	77.1
PIUN	11.1	0	21.1	11.1	38.1	38.1
Other	6.1	3.8	0	0	3.6	0
Overall	54.5	59.7	63.3	64.1	67.6	67.7

Table 5: F1-scores of feature fusion from the HI sets by stacking a number of features (No.) to the reference set D_{HI} : 20 features from PC_{HI}^D , 25 features from S_{HI} , 18 features from PC_{HI}^S , 8 features from PC_{HI}^L and 2 features from L_{HI} . The last column represents the contribution of the HI data for the tree species classification: overall F1-score of 67.7%

Specie	S_{Li}	PC_{Li}^D	D_{Li}
No.	2	1	1
ABAL	59.1	58.7	61.2
FASY	57.6	57.1	57.2
PIAB	77.5	78.0	77.4
PIUN	45.5	45.5	45.5
Other	0	0	3.7
Overall	68.6	68.8	68.9

Table 6: F1-scores of feature fusion from three LiDAR sets by stacking a number of features (No.) to the reference set described in Table 5

Tables 7 and 8 present the species classification at the pixel-level and tree-level, respectively. The overall accuracy at tree-level is 74.0%, which is greater to the accuracy at pixel-level, 68.9%. The species PIAB, the most abundant class, obtained the highest producer’s accuracy among all species, but the user’s accuracy is affected by the wrong classification of the minority species as it is illustrated in Figure 4. Two trees labeled as ABAL (Figure 4(b)) and PIUN (Figure 4(h)) species, are wrongly detected as PIAB. In Figure 4(d), the tree is correctly detected as FASY, but the amount of pixels predicted as PIAB is notorious over the region of interest. The second class after the PIAB tree species with good performance is ABAL, which belongs to the conifers family. At tree-level, ABAL obtained a good producer’s accuracy (70.6%) and the second highest user’s accuracy (85.7%). The FASY tree species has less number of crowns than the class ABAL and the second amount of pixels in our dataset. Broadleaves, such as FASY tree species, are characterized by having the biggest tree crowns with the greatest average amount of pixels per crown (Table 3). FASY performed the highest user’s accuracy at pixel- and tree-level. The minority classes such as PIUN and other species are challenging to be detected because the model predicted them as the majority class PIAB.

The approach in (Lee et al., 2016) based rPCA and SVM provided an overall accuracy for the classification of six tree species of 91.7% and 61.1% at the pixel- and tree-level, respectively. In the test set, they used 1537 pixels associated with 677 crown trees, which means every crown had approximately 2.3 pixels. There is a negative difference of 30.6 points of percentage between the pixel- and the tree-level overall accuracy by having an estimated increase of pixels from 1 to 2.3. Our test set contains 73 tree crowns with 1550 pixels (21.2 pixels per crown). The resulting overall accuracy for the classification of four tree species was 68.9% and 74.0% at the pixel- and tree-level, respectively. The estimated increase of pixels from 1 to 21.2, generated an improvement in the overall accuracy of 5.1 points of percentage between the pixel- and the tree-level performance.

	Species	Ground truth						Producer's accuracy [%]
		ABAL	FASY	PIAB	PIUN	Other	Total	
Classification results	ABAL	174	0	91	0	0	265	65.7
	FASY	55	176	202	0	0	433	40.6
	PIAB	66	4	712	0	3	785	90.7
	PIUN	0	0	12	5	0	17	29.4
	Other species	9	2	38	0	1	50	2.0
	Total	304	182	1055	5	4	1550	
	User's accuracy [%]	57.2	96.7	67.5	100	25.0		68.9

Table 7: Species classification results at pixel-level obtained by the fusion of HI and LiDAR features

	Species	Ground truth						Producer's accuracy [%]
		ABAL	FASY	PIAB	PIUN	Other	Total	
Classification results	ABAL	12	0	5	0	0	17	70.6
	FASY	1	3	6	0	0	10	30.0
	PIAB	1	0	38	0	0	39	97.4
	PIUN	0	0	2	1	0	3	33.3
	Other species	0	0	4	0	0	4	0
	Total	14	3	55	1	0	73	
	User's accuracy [%]	85.7	100	69.1	100	0		74.0

Table 8: Species classification results at tree-level obtained by the fusion of HI and LiDAR features

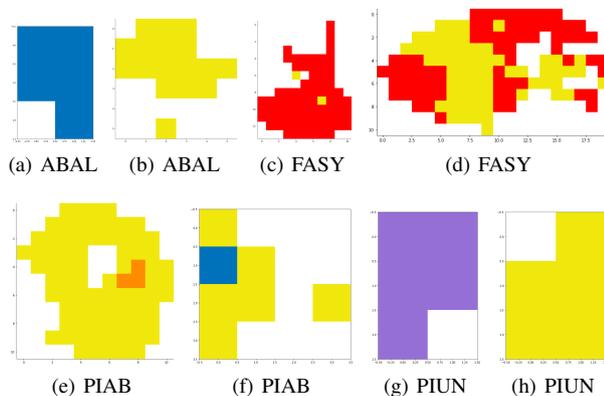


Figure 4: Results of pixel classification for tree-level assessment of 8 tree crowns: (a) ABAL tree correctly detected, (b) ABAL tree detected as PIAB, (c) FASY tree correctly detected, (d) FASY tree correctly detected with an important amount of PIAB pixels, (e) and (f) PIAB trees correctly detected with few pixels from other species, (g) PIUN tree correctly detected and (h) PIUN tree detected as PIAB. Legend: ● *Abies alba* (ABAL), ● *Fagus sylvatica* (FASY), ● *Picea abies* (PIAB), ● *Pinus uncinata* (PIUN) and ● other species

5. CONCLUSION

In this work, we presented a data fusion scheme that integrated features extracted from the HI and the LiDAR data for semantic segmentation. The potential of rPCA and PCA are applied to expand the feature representation through the low-rank and sparse information, and their principal components. The sparse features from the HI contributed to increase the F1-score of the classifier performance by adding 49.4% of the sparse features. The feature selection implements two steps: first, the intra-set feature importance was described by the RF scores; then, the inter-set feature importance was estimated by the overall F1-score for stacking different set of features. This method offers the flexibility of incorporating additional features such as vegetation indexes, which increases the interpretability of the model. The feature reduction process decreased the dimensionality from 1104 to 87 features. The overall accuracy of tree species classification at pixel-level was 68.9%. Our approach showed that 74.0% of trees were

correctly assigned overall, which represented an improvement of 12.9 points of percentage with respect to (Lee et al., 2016). Although these two remote sensing modalities provided valuable information, the most abundant classes, the conifers such as *Abies alba* and *Picea abies*, benefited mainly from the HI information. However, it became a challenging task for dealing with imbalanced classes of our dataset, such as *Pinus uncinata*. Future work is going to be focused in the implementation of tree species detectors by selecting the most relevant features from HI and LiDAR data for each specific tree species.

ACKNOWLEDGMENT

This work was supported by LabEx OSUG@2020 through the Investissements d'avenir under Grant ANR10 LABX56.

REFERENCES

- Albon, C., 2018. Machine learning with python cookbook: Practical solutions from preprocessing to deep learning. O'Reilly Media, Inc.
- Alonzo, M., McFadden, J. P., Nowak, D. J. and Roberts, D. A., 2016. Mapping urban forest structure and function using hyperspectral imagery and LiDAR data. *Urban Forestry & Urban Greening* 17, pp. 135 – 147.
- Anderson, C. B., 2018. The CCB-ID approach to tree species mapping with airborne imaging spectroscopy. *PeerJ* 6, pp. e5666.
- Baldeck, C. A. and Asner, G. P., 2014. Improving remote species identification through efficient training data collection. *Remote Sensing* 6(4), pp. 2682–2698.
- Bioucas-Dias, J. M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N. and Chanussot, J., 2013. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and Remote Sensing Magazine* 1(2), pp. 6–36.
- Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P. and Chanussot, J., 2012. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(2), pp. 354–379.

- Brunton, S. and Kutz, J., 2019. Data-driven science and engineering: Machine learning, dynamical systems, and control. Cambridge University Press.
- Dalponte, M., Bruzzone, L. and Gianelle, D., 2012. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment* 123, pp. 258–270.
- Dalponte, M., Frizzera, L. and Gianelle, D., 2019. Individual tree crown delineation and tree species classification with hyperspectral and LiDAR data. *PeerJ* 6, pp. e6227.
- Dalponte, M., Frizzera, L., Ørka, H. O., Gobakken, T., Næsset, E. and Gianelle, D., 2018. Predicting stem diameters and above-ground biomass of individual trees using remote sensing data. *Ecological Indicators* 85, pp. 367 – 376.
- Dalponte, M., Ørka, H. O., Ene, L. T., Gobakken, T. and Næsset, E., 2014. Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sensing of Environment* 140, pp. 306 – 317.
- Dechesne, C., Mallet, C., Bris, A. L. and Gouet-Brunet, V., 2017. Semantic segmentation of forest stands of pure species combining airborne LiDAR data and very high resolution multispectral imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 126, pp. 129 – 145.
- Friedman, J. H., 2001. Greedy function approximation: A gradient boosting machine. *The Annals of Statistics* 29(5), pp. 1189–1232.
- Fuhr, M., Cordonnier, T., Courbaud, B., Kunstler, G., Mermin, E., Riond, C. and Tardif, P., 2017. Long-term tree inventory data from mountain forest plots in France. *Ecology* 98(4), pp. 1180–1180.
- Ghosh, A., Fassnacht, F. E., Joshi, P. and Koch, B., 2014. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *International Journal of Applied Earth Observation and Geoinformation* 26, pp. 49 – 63.
- Kandare, K., Ørka, H. O., Chan, J. C.-W. and Dalponte, M., 2016. Effects of forest structure and airborne laser scanning point cloud density on 3D delineation of individual tree crowns. *European Journal of Remote Sensing* 49(1), pp. 337–359.
- Laybros, A., Aubry-Kientz, M., Féret, J.-B., Bedeau, C., Brunaux, O., Derroire, G. and Vincent, G., 2020. Quantitative airborne inventories in dense tropical forest using imaging spectroscopy. *Remote Sensing* 12(10), pp. 1577.
- Laybros, A., Schläpfer, D., Féret, J.-B., Descroix, L., Bedeau, C., Lefevre, M.-J. and Vincent, G., 2019. Across date species detection using airborne imaging spectroscopy. *Remote Sensing* 11(7), pp. 789.
- Lee, J., Cai, X., Lellmann, J., Dalponte, M., Malhi, Y., Butt, N., Morecroft, M., Schonlieb, C. B. and Coomes, D. A., 2016. Individual tree species classification from airborne multisensor imagery using robust PCA. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9(6), pp. 2554–2567.
- Lee, J., Cai, X., Schönlieb, C. and Coomes, D., 2015. Mapping individual trees from airborne multi-sensor imagery. In: 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 5411–5414.
- Liao, W., Van Coillie, F., Li, L., Zhao, B., Gao, L., Philips, W. and Zhang, B., 2017. Fusion of multi-scale hyperspectral and lidar features for tree species mapping. In: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 2879–2882.
- Luan, X., Fang, B., Liu, L., Yang, W. and Qian, J., 2014. Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. *Pattern Recognition* 47(2), pp. 495 – 508.
- Maschler, J., Atzberger, C. and Immitzer, M., 2018. Individual tree crown segmentation and classification of 13 tree species using airborne hyperspectral data. *Remote Sensing* 10(8), pp. 1218.
- Matsuki, T., Yokoya, N. and Iwasaki, A., 2015. Hyperspectral tree species classification of Japanese complex mixed forest with the aid of LiDAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8(5), pp. 2177–2187.
- Monnet, J.-M. and Mermin, E., 2014. Cross-correlation of diameter measures for the co-registration of forest inventory plots with airborne laser scanning data. *Forests* 5(9), pp. 2307–2326.
- Torabzadeh, H., Leiterer, R., Hueni, A., Schaepman, M. E. and Morsdorf, F., 2019. Tree species classification in a temperate mixed forest using a combination of imaging spectroscopy and airborne laser scanning. *Agricultural and Forest Meteorology* 279, pp. 107744.
- Tusa, E., Laybros, A., Monnet, J.-M., Mura, M. D., Barré, J.-B., Vincent, G., Dalponte, M., Féret, J.-B. and Chanussot, J., 2020. Chapter 2.11 - Fusion of hyperspectral imaging and LiDAR for forest monitoring. In: J. M. Amigo (ed.), *Hyperspectral Imaging, Data Handling in Science and Technology*, Vol. 32, Elsevier, pp. 281 – 303.
- Tusa, E., Reynolds, A., Lane, D. M., Robertson, N. M., Villegas, H. and Bosnjak, A., 2014. Implementation of a fast coral detector using a supervised machine learning and gabor wavelet feature descriptors. In: 2014 IEEE Sensor Systems for a Changing Ocean (SSCO), pp. 1–6.
- Vincent, G., Antin, C., Laurans, M., Heurtebize, J., Durrieu, S., Lavalley, C. and Dautat, J., 2017. Mapping plant area index of tropical evergreen forest by airborne laser scanning. A cross-validation study using LAI2200 optical sensor. *Remote Sensing of Environment* 198, pp. 254–266.
- Xiao, W., Zaforemska, A., Smigaj, M., Wang, Y. and Gaulton, R., 2019. Mean shift segmentation assessment for individual forest tree delineation from airborne LiDAR data. *Remote Sensing* 11(11), pp. 1–19.