AUTOMATED STRUCTURAL FOREST CHANGES USING LiDAR POINT CLOUDS AND GIS ANALYSES

A. Novo1*, H. González-Jorge2, J. Martínez-Sánchez1, José María Fernández-Alonso3, H. Lorenzo4

1. Geotech-Group, CINTECX, Department of Natural Resources and Environmental Engineering School of Mining Engineering, University of Vigo, 36310, Vigo, Spain - (annovo, joaquin.martinez3@uvigo.es
2. Physics Engineering Lab, School of Aerospace Engineering, University of Vigo, 32004, Ourense, Spain – higiniog@uvigo.es
3. Lourizán Forest Research Center Xunta de Galicia, P.O. Box 127, 36080 Pontevedra, Spain – josemfernandez@uvigo.es
4. Geotech-Group, CINTECX, Department of Natural Resources and Environmental Engineering, School of Forestry Engineering, University of Vigo, 36005 Pontevedra, Spain – hlorenzo@uvigo.es

KEY WORDS: Mapping, Remote Sensing, LiDAR data, GIS analysis, Point cloud processing, Vegetation structure, Forest parameters.

ABSTRACT:

Forest spatial structure describes the relationships among different species in the same forest community. Automation in the monitoring of the structural forest changes and forest mapping is one of the main utilities of applications of modern geoinformatics methods. The obtaining of objective information requires the use of spatial data derived from photogrammetry and remote sensing. This paper investigates the possibility of applying light detection and ranging (LiDAR) point clouds and geographic information system (GIS) analyses for automated mapping and detection changes in vegetation structure during a year of study. The research was conducted in an area of the Ourense Province (NWSpain). The airborne laser scanning (ALS) data, acquired in August 2019 and June of 2020, reveal detailed changes in forest structure. Based on ALS data the vegetation parameters will be analysed. To study the structural behaviour of the tree vegetation, the following parameters are used in each one of the sampling areas: (1) Relationship between the tree species present and their stratification; (2) Vegetation classification in fuel types; (3) Biomass (Gi); (4) Number of individuals per area; and (5) Canopy cover fraction (CCF). Besides, the results were compared with the ground truth data recollected in the study area. The development of a quantitative structural model based on Aerial Laser Scanning (ALS) point clouds was proposed to accurately estimate tree attributes automatically and to detect changes in forest structure. Results of statistical analysis of point cloud show the possibility to use UAV LiDAR data to characterize changes in the structure of vegetation.

1. INTRODUCTION

Forest constitutes the most biologically diverse terrestrial ecosystem on Earth and are imperative for maintaining the balance of terrestrial ecosystems (Dandois, Olano, and Ellis 2015). To promote and support sustainable forest management an accurate monitoring in timely fashion is required (Timilsina et al. 2013). In this context, forest structural parameters (e.g., tree height, volume, and biomass) are key components for effectively quantifying forest structure and are vital for accurately monitoring forest dynamics (Fu et al. 2021). Assessing changes in forest structure over time is crucial for monitoring forest resources. The generation of spatially explicit detailed maps of forest structure, and its dynamics, has multiple implications in forest managing wildfire risk reduction, carbon sequestration assessment, timber resources availability or wildfire habitat analysis may benefit from such high-resolution information. Forest structural diversity is the physical arrangement and variability of the living and non-living biotic elements within forest stands that support many essential ecosystem functions (LaRue et al. 2019). Forest structural diversity arises from the complex interactions of a range of abiotic and biotic factors that influence the growth and the quality of vegetation (Fotis et al. 2018). A wide variety of structural diversity metrics can be estimated using methods that range from traditional forest inventory approaches to next-generation remote sensing techniques (Fahey et al. 2019). The complex and dynamic nature of forest structure has proven to be challenging to measure accurately across scales and forest structure types (Atkins et al. 2018).

The forestry management process is based on the use of a large amount of information which must be stored, managed, analyzed, simulated, and visualized in a dynamic and flexible way. Geographic Information Systems (GIS) and remote sensing are complementary technologies that, when combined, enable to improve monitoring, mapping and management of forest resources (S. E. Franklin 2001). Automation in the monitoring of the structural forest changes and forest mapping is one of the main aspects of applications of modern geoinformatics methods. The obtaining of objective information requires the use of spatial data derived from photogrammetry and remote sensing. Generating the spatial characteristics of vegetation in an automated way undoubtedly provides new possibilities in modelling the structure of vegetation, including defining biometric features and biomass,
The aim of this work is to analyze the potential of automated mapping and detection changes in vegetation structure using LiDAR technology and Geographic Information System (GIS) analyses. In this way, human work would not be necessary for field data collection and the methodology could be extended to large or inaccessible areas, being able to map and analyze the vegetation evolution whose data are necessary in forest management. Firstly, the LiDAR data are collected with a year of difference between first and second flight. Moreover, the ground truth collected by human methods coincide with the second flight date. Then, data are recollected in two square plots of 2 m size within the study area. The LiDAR data are processed to be compared with ground truth data and the results accuracy is calculated in both study plots. Finally, the methodology is extended to the whole study area, and combustible fuel types and their evolution are mapped. In particular, the main contributions of this study are summarized as:

- Design of a methodology to group tree points by height, Prometheus classification, and heights established in ASPRS classification.
- Development of an algorithm to automatically calculate the height distribution of vegetation and their Canopy Cover Fraction (CCF).
- Calculate the accuracy of the methodology compared with the one for the ground truth.

Remote sensing has demonstrated its importance for the characterization of vegetation structure in sparsely dense forests, the greatest challenge being those with medium-high diversity (J. Franklin 2010). Light Detection and Ranging (LiDAR) is a remote sensing technology for characterizing the surface of the earth using a cloud of georeferenced points. LiDAR is a useful tool for the multi-dimensional characterization of forest structure because it has a strong capability to penetrate dense forest canopies and detect understory vegetation, thereby, obtaining high-precision three-dimensional (3D) forest structure information. Over more, there are versatile terrestrial and aerial deployment platforms. Terrestrial laser scanning (TLS) and aerial laser scanning (ALS) have both been shown to be effective at quantifying components of forest structural diversity (LaRue et al. 2020). The development of airborne digital cameras and unmanned aerial vehicles (UAV) has promoted cost-effective methods for enhancing monitoring forest dynamics. In the forest sector, LiDAR has the potential to reduce the need for intensive ground-based measurement of stand structure, making it a valuable tool. LiDAR data have been recently used to quantify complexity and diversity in vegetation structure in a successful way (Atkins et al. 2018; Baxk et al. 2019; Guo et al. 2017; LaRue et al. 2020).

The present manuscript explores the accuracy of airborne LiDAR data to support, automatically map, and monitoring the forest structural changes, in a more efficient way than traditional forest methods do such as human field data collection.

The experimental data for this work were collected using a Phoenix system, which is based on a Velodyne LiDAR model, the Alpha AL3-32. It shows survey-grade centimetric accuracy and intensity calibration. Their 32 lasers emit 700,000 pulses per second and record up to two returns per pulse. The system includes a global navigation satellite system (GNSS) that provides real-time kinematics and post-processing options with an accuracy specification up to 1cm in horizontal and 2.5 cm in vertical positioning. The raw point cloud of the first flight was collected on 30th August 2019 and the second flight was performed on 25th June 2020. Data were collected with a density of 350 points/m² and an average point spacing of 0.05 m. The point cloud collected in 2019 is composed by a total of 1,074,390 points, while the total points of the 2020 point cloud are 705,852.

## 2. MATERIAL AND METHODS

### 2.1 Area of study

The study area is located in the northwest of Spain. It belongs to the Natural Park of Baixa Limia Serra do Xurés, which has been catalogued as an Area for Special Conservation (ASC). The protected areas are ideal settings for research. The subject of study is an area of 0.30 Ha in the central part of the municipality of Lobios (Figure 1). The study area contains two sampling subareas corresponding to the ground truth data. One four square meter plot was established in each subarea to carry out field-based data collection. The climatic type existing in the Baixa Limia is called sub-Mediterranean oceanic temperate, which indicates a certain aridity during summer. This means that a large part of vegetation is adapted to dry periods. Under this climatic type, the potentially dominant vegetation in most of the territory is Quercus pyrenaica and Quercus robur. The main tree species are, Betula alba, Quercus suber, Arbutus unedo, Pinus sp., Ulex sup, Cytisus scoparius and Erica sp. These are several endemic plants, including Portugal laurel and Prunus lusitanica, a species that colonizes the ravines and other areas that have high humidity. The biogeographical location of Baixa Limia greatly favours the diversity of the flora in this territory.

![Figure 1. Location of study area.](image)

### 2.2 Materials

- Analyse the forest changes in the whole study area, supported by height parameters, biomass estimation, individual tree detection, CCF and fuel types of classification.
ALS point clouds were first preprocessed filtering vegetation points using the command Lasground from the LAStools software (Isenburg 2012). This process was done to remove the noise and ground points. Then, data were normalized by using the command Lasheight.

In this work, filed based inventory data are used as ground truth for their comparison with the LiDAR data. The ground truth data were collected in two subareas within the study area. Plot 1 has an average “x” and “y” coordinates of 574037.3 m and 4633978.6 m, respectively. Plot 2 shows an average “x” coordinate of 573999.4 m and “y” coordinate of 4633973.3 m. In both cases EPSG:25829 ETRS89/UTM zone 29N.

That validation data was collected by setting a four-square meters plot. For the computational analyses, a two-meter ratio influence area (buffer) was generated around each plot taking the average “x” and “y”coordinates as reference. The field data collection consisted of the characterization of canopy surface fuel strata. For the canopy strata, top canopy height, height of living crown and canopy closure were recorder. A Haglöf Vertex Hypsometer was used to measure vertical heights. This instrument uses ultrasonic signals to obtain the distance. For the surface fuel stratum, a quadrad sampling method was applied, where height and coverage measurements were taken each 25 cm. Average height of ligneous species (shrubs), coverage of herbaceous species, coverage of shrubs and percentage of plot without vegetation cover was derived from that sampling. Each sampling plot was classified according to Prometheus fuels classification according to measured data and visual inspection.

### 2.3 Data processing

This study was developed using QGIS software (QGIS 2018) and Python language (Van Rossum 2007) for mapping and spatial analysis. The computer on which the data processing was carried out is a DELL G5 5500, with the following technical characteristics: Processor: Intel(R) Core (TM) i7-1070 CPU @ 2.60GHz, installed RAM:16.0 GB and 64-bit operating system, x64-based processor.

Data processing begins with the heigh distribution functions of the point-cloud. The heigh information is synthesizing thought raster layers generation and an algorithm was developed in python language to carry on the transformation of 3D points into the 2D space. The pixel value is related with the z coordinate from the point cloud.

First, the ground points were identified using lasground, while the height of each point above the ground was computed using lasheight. It removes low and hight outliers that are often just noise. Therefore, each point of the point cloud contains its X, Y, and normalized Z coordinates.

First phase of the transformation of LiDAR data into plots was the segmentation of point clouds in circular segments of 2 m of radius. Once the data of each plot were separated, next step was the automatic classification of vegetation points according to the following height intervals, the same as in the field data collection (Figure 2):

- Low vegetation: 0.15 - 0.5 m
- Medium vegetation: 0.5 - 2 m
- Medium- hight vegetation: 2 - 4 m
- High vegetation: > 4m

Once the vegetation was grouped by height the statistical variables were calculated by GIS static analysis. The average height for each established vegetation stratum has been calculated, as well as for the entire vegetation in both years of study. In addition, the subtraction of heights in both years gives the characteristics of evolution and changes in the forest structure both in total area and in study plots.

The CCF has been calculated for each stratum of vegetation and for the entire vegetation in both years of study. The CCF indicates the proportion of ground covered by vertical projection of each vegetation stratum. Figure 3 shows the binary images of CCF calculated in Plot 1 and Plot 2. All the parameters involved in this study were also applied to the whole study area in both years of study. It was divided in cells of 2 m to achieve a better characterisation of the fuels on the vegetation structure.
classification criterion in Prometheus is the type and height of the propagating element, divided into three well differentiated groups: grass, shrub, and tree. The information extracted from the LiDAR data corresponds with the number of points in each generated interval. Moreover, the percentage of vegetation points of the study area have been estimated. An algorithm developed in Python language performs the automated process. Once the percentage of points in each band was calculated the next step was the application of the classification conditions carried out to found out the fuel model. Figure 4 shows the fuel models according to Prometheus in the study area.

![Figure 4. Fuels classified by Prometheus model.](image1)

The next step consists of testing the parameters calculated for study plots with the field data. Besides, the study was complemented with the calculation of the following parameters: biomass estimation and number of trees detected in the study area. The biomass value was estimated by GIS analysis. A digital vegetation model is a normalized surface model in which the ground values are subtracted. A Canopy Height Model (CHM) was computed as a difference between Digital Surface Model (DSM) and Digital Terrain Model (DTM). At first, DTM is created from the ground returns and a DSM from the first returns. They were calculated from the tool las2dem selecting the last pulse, which represents the ground, and the first pulse representing other elevated features on the ground as trees. CHM was generated containing the information of tree heights. To calculate the volume occupied by trees the CHM was multiplied by the area of each pixel (0.10 m × 0.10 m = 0.01 m²). The sum of all volumes (m³) in the study area was calculated using a zone statistics tool.

![Figure 5. CHM of study area in 2019 and 2020](image2)

To estimate the tree points in the study area, a CHM derived from LiDAR was used to detect Individual Tree Crowns (ITC). Two pre-processing steps prepare a watershed segmentation approach: (1) Gaussian filtering and (2) inversion of CHM. The processing toolbox type smooth with gaussian filtering of Orfeo toolbox was used to establish a circular structuring element of a radius of 2 pixels. In the next step, the smoothed CHM was inverted by the toolbox invert grid of SAGA. Finally, the watershed segmentation toolbox of SAGA was used to calculate the geolocation of points. The height of each individual tree was estimated through the CHM using the QGIS point sampling plugin.

![Figure 6. Individual tree crowns detection](image3)

3. RESULTS AND DISCUSSION

Table 1 shows the results obtained in LiDAR data processing and the ground truth between plot 1 and plot 2 in the study area.

|                  | Plot 1 LiDAR | Ground truth | Plot 2 LiDAR | Ground truth |
|------------------|--------------|--------------|--------------|--------------|
| HERBACEOUS HEIGHT (CM) | 21           | 17           | 21           | 26.5         |
| SHRUB HEIGHT (CM)    | 165          | 150          | -            | 57           |
| MAXIMUM HEIGHT (M)  | 9.56         | 6.0          | 10.81        | 8.0          |
| % WITHOUT CCF       | 0.11         | 0            | 0.33         | 0            |
| HERBACEOUS CCF      | 0.03         | 0.05         | 0.14         | 0.9          |
| % SHRUB CCF         | 0.01         | 0.05         | -            | 0.05         |
| FUEL TYPE           | 5            | 5            | 5 - 6        | 5            |

Table 1. Comparison results of LiDAR data and ground truth

The total height in Plot 1 of study detected by LiDAR data in herbaceous stratum was 21 cm in comparison with the 17 cm of the ground truth collected. The result of LiDAR was 4 cm highest in plot 1, while in plot 2 was the contrary, the total height in herbaceous stratum was 21 cm in comparison with the 26.5 cm of the ground truth used. On the one hand, the height shrub detected by LiDAR analyses was 165 cm for shrub in plot 1, in comparison with the 150 cm collected in the field. On the other hand, in plot 2 no shrub stratum was detected in the LiDAR analyses, while the ground truth showed a 57 cm present in the plot 2 study. The maximum height parameters...
showed higher differences between LiDAR data and ground truth, being the Lidar height calculation higher than the ground truth collected by a vertex instrument. In the plot 1 of study the maximum height detected was of 9.56 m, while in the ground truth was 6 m. The maximum height detected in plot 2 was 10.81 m while in ground truth was 8 m. There is a correspondence of approximate 2 m between plots height in both data. Except for herbaceous and shrub height of LiDAR data in plot 2, generally the LiDAR data was higher than the ground truth data.

The percentage of CCF calculated by LiDAR data processing showed a lower result than CCF estimated in the field. The ground truth data showed no percentage without CCF in both study plots, while LiDAR data results detected a 0.11% of plot 1 without CCF and a 0.33% in plot 2. Herbaceous and shrub CCF calculated by LiDAR data showed a lower percentage than in ground truth, besides there is no result of shrub CCF in plot 2 of study.

The comparison in the analysis of Prometheus classification showed the accuracy in the classification of LiDAR height stratum. Results of vegetation classification detected a fuel type 5 in both plots. Moreover, areas with presence of fuel type 6 were detected in plot 2.

These differences between LiDAR data and ground truth could result from the buffer analysis around the plots. The exact same portion of plots has not been extracted, so there is an overestimation in the results. As a result, the most highlighted error detected was in the height shrubs with an absolute error of 15 cm in plot 1 and 57 cm in plot 2, while the highest accuracy in results was in the CCF shrubs with an absolute error of 0.04 % in plot 1 and 0.05 % in both 2.

Focused on the analysis of structural forest changes in whole area of study there was analyzed the average height of each stratum. The vegetation was divided in 4 classes and the following parameters were calculated: CCF, biomass, number of individual trees, area occupied for each fuel type and the average height. These parameters were analyzed in both years of study. To carry on the study of structural changes the area was divided in cells of 2 x 2 m for an exhaustive analysis.

Results of average height are shown in Figure 7 in both years of recorder LiDAR data in this study, 2019 and 2020.

In the first vegetation group, herbaceous group (low vegetation: 0.15 m - 0.5 m) is 0.03 m lower in 2020 than in 2019, although show similar values in both years. With respect to low shrubs (medium vegetation: 0.5 m - 2 m), the average height in 2020 is 6 cm higher than in 2019. In the group of high shrubs (2 m - 4 m), the average height is similar in both years, showing 6 cm higher vegetation in 2019 than in 2020, contrary to the high vegetation group (> 4 m), whose results show an increment of 50 cm in the tree height between the period of 2019 and 2020. The total average height in the study area was also increased, registering a total of 8.95 m in 2019 and 9.58 m in 2020, so the increment of height was of 63 cm.

Figure 8 shows the results of CCF obtained for each stratum. In all groups is appreciable the decrease of CCF in both years, which represents a 5% in tree groups and a 3% in low and higher shrubs groups. The area without CCF is increased in a 6% in 2020.

The analysis of fuel types showed lower cover area in year 2020 than in 2019 with exception of fuel type 6, which covers a total of 10 m² more in 2020. The fuel type 6 covered 178 m² lower in 2020.

The results obtained for biomass parameter shows an increment with a result of 27,644 m³ for 2019 and 28,227 m³ for 2020. A total of 583 m³ in a study area of 0.30 Ha was increased during the study year.

The total number of points which represents the individual trees detection were 834 in 2019 and 573 in 2020, 771 of which are trees in 2019 and 531 in 2020. In conclusion 0.23% of trees increased in the study area during a year.

### 4. CONCLUSIONS

The present study showed a promising approach to characterize, classify, and temporally evaluate forest fuels on a woodland area. Some limitations, however, should be pointed out. The difference obtained between the ground truth in comparison with LiDAR processing is based on the different location of the square plots. This is caused by establishment of the buffer influence area on the average coordinates. Moreover, the difference could be related with the accuracy specification up to 1 cm in horizontal and 2.5 cm in vertical positioning of the UAV data recorder.

CCF results for year 2019 were higher than CCF in 2020. This could be caused by the high number of points. Point cloud of first flight contains a total of 1,074,390 points, while a total of 705,852 points come from the point cloud of second flight (2020).
The point cloud classification has been compared according to Prometheus system, obtaining the same results than in field measures, so it can conclude the accuracy of the method. The maximum error associated in LiDAR processing was in herbaceous and shrubs strata, and it could be originated by the penetration limitation of LiDAR caused by dense canopy cover in the study area.

Results of statistical analysis of point cloud show the possibility to use UAV LiDAR data to characterize changes in the structure of vegetation since the changes in a year also were significant.

The results of this study have special interest for forest management. To generate the spatial characteristics of vegetation in an automated way undoubtedly provides new possibilities in modelling the structure of vegetation, including defining biometric features and biomass, which determines the developmental stage of trees and shrubs forming the succession process.

ACKNOWLEDGEMENTS

A.N. wants to thank University of Vigo through the grant Axudas Predoutorais para a formación de Doutores 2019 (grant number 00VI 1311 6401211) and “Expenses were partially covered by the Travel Award sponsored by the open access journal Sensors published by MDPI.” J.M.-S. wants to thank Ministerio de Ciencia, Innovación y Universidades for the support through the project PID2019-108816RB-I00.

REFERENCES

Albini, Frank A. 1976. Estimating Wildfire Behavior and Effects. Vol. 30. Department of Agriculture, Forest Service, Intermountain Forest and Range....

Atkins, Jeff W, Gil Bohrer, Robert T Fahey, Brady S Hardiman, Timothy H Morin, Atticus E L Stovall, Naupaka Zimmerman, and Christopher M Gough. 2018. “Quantifying Vegetation and Canopy Structural Complexity from Terrestrial LiDAR Data Using the ForestR Package.” Methods in Ecology and Evolution 9 (10). Wiley Online Library: 2057–2066.

Bakx, Tristan R M, Zsofia Koma, Arie C Seijmonsbergen, and W Daniel Kissling. 2019. “Use and Categorization of Light Detection and Ranging Vegetation Metrics in Avian Diversity and Species Distribution Research.” Diversity and Distributions 25 (7). Wiley Online Library: 1045–1059.

Dandois, Jonathan P, Marc Olano, and Erle C Ellis. 2015. “Optimal Altitude, Overlap, and Weather Conditions for Computer Vision UAV Estimates of Forest Structure.” Remote Sensing 7 (10). Multidisciplinary Digital Publishing Institute: 13895–13920.

Fahey, Robert T, Jeff W Atkins, Christopher M Gough, Brady S Hardiman, Lucas E Nave, Jason M Tallant, Knute J Nadehoffer, Christoph Vogel, Cynthia M Scheuermann, and Ellen Stuart-Haëntjens. 2019. “Defining a Spectrum of Integrative Trait-based Vegetation Canopy Structural Types.” Ecology Letters 22 (12). Wiley Online Library: 2049–2059.

Fotis, Alex T, Timothy H Morin, Robert T Fahey, Brady S Hardiman, Gil Bohrer, and Peter S Curtis. 2018. “Forest Structure in Space and Time: Biotic and Abiotic Determinants of Canopy Complexity and Their Effects on Net Primary Productivity.” Agricultural and Forest Meteorology 250. Elsevier: 181–191.

Franklin, Janet. 2010. Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press.

Franklin, Steven E. 2001. Remote Sensing for Sustainable Forest Management. CRC press.

Fu, Xiaoyao, Zhengnan Zhang, Lin Cao, Nicholas C Coops, Tristan R H Goodbody, Hao Liu, Xin Shen, and Xiangqian Wu. 2021. “Assessment of Approaches for Monitoring Forest Structure Dynamics Using Bi-Temporal Digital Aerial Photography Point Clouds.” Remote Sensing of Environment 255. Elsevier: 112300.

Guo, Xuan, Nicholas C Coops, Piotr Tompalski, Scott E Nielsen, Christopher W Bater, and J John Stadl. 2017. “Regional Mapping of Vegetation Structure for Biodiversity Monitoring Using Airborne Lidar Data.” Ecological Informatics 38. Elsevier: 50–61.

Isenburg, Martin. 2012. “LaStools-Efficient Tools for LiDAR Processing.” Available at: Http: Http://Www. Cs. Unc. Edu/~ Isenburg/Lastools/[Accessed October 9, 2012].

LaRue, Elizabth A, Brady S Hardiman, Jessica M Elliott, and Songlin Fei. 2019. “Structural Diversity as a Predictor of Ecosystem Function.” Environmental Research Letters 14 (11). IOP Publishing: 114011.

LaRue, Elizabeth A, Franklin W Wagner, Songlin Fei, Jeff W Atkins, Robert T Fahey, Christopher M Gough, and Brady S Hardiman. 2020. “Compatibility of Aerial and Terrestrial LiDAR for Quantifying Forest Structural Diversity.” Remote Sensing 12 (9). Multidisciplinary Digital Publishing Institute: 1407.

“Phoenix LiDAR Systems.” https://www.phoenixlidar.com/. QGIS. 2018. “QGIS SOFTWARE.” https://www.qgis.org/en/site/

Szostak, Marta. 2020. “Automated Land Cover Change Detection and Forest Succession Monitoring Using LiDAR Point Clouds and GIS Analyses.” Geosciences 10 (8). Multidisciplinary Digital Publishing Institute: 321.

Timilsina, Nilesh, Wendell P Cropper, Francisco J Escobedo, and Joanna M Tucker Lima. 2013. “Predicting Understory Species Richness from Stand and Management Characteristics Using Regression Trees.” Forests 4 (1). Multidisciplinary Digital Publishing Institute: 122–136.

Van Rossum, Guido. 2007. “Python Programming Language.” In USENIX Annual Technical Conference, 41:36.