

Czech University of Life Sciences in Prague

Faculty of Forestry & Wood Sciences

Department of Forest Management



**Detection and modeling of forest attributes in forest with different
density using remote sensing and auxiliary data**

Doctoral Dissertation

Author: M.Sc. Azadeh Abdollahnejad
Supervisor: Priv.-Doz. Ing. Peter Surový, PhD.

2018

Author`s Declaration

I confirm that this doctoral dissertation with the title “Detection and modeling of forest attributes in forest with different density using remote sensing and auxiliary data” has

- i. been composed entirely by myself.*
- ii. been solely the result of my work.*
- iii. not been submitted elsewhere.*

The doctoral dissertation implemented under Act no. 111/1998 Sb. on Universities and was done on the basis of literature sources of the Bibliography.

Prague, 2018.

© Copyright by Azadeh Abdollahnejad 2018.

All Rights Reserved.



Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Priv.-Doz. Ing. Peter Surovy, Ph.D. for his continuous support during my entire Ph.D. study and research, for his patience, motivation, and immense knowledge. He was the person who encouraged me to write this dissertation and to achieve that goal. I could not have imagined having a better advisor and mentor for my Ph.D. study.

I would also like to express my gratitude to my colleagues for their support, collaboration and friendship.

I would like to dedicate this work to my parents whose dream for me have resulted in this achievement and without their loving upbringing and nurturing; I would not have been where I am today and what I am today.

This project would have been impossible without the support by the: 1) Internal Grant Agency (IGA) of the Faculty of Forestry and Wood Sciences of the Czech University of Life Sciences (CULS) in Prague:

- I. **2015-2017:** Internal Grant Agency (IGA) of the Faculty of Forestry & Wood Sciences at CULS, Prague (No. B07/15). Title: *New ways in precision forestry: From data to decisions.*
- II. **2016-2017:** Internal Grant Agency (IGA) of Faculty of Forestry & Wood Sciences, CULS, Prague (No. A14/16). Title: *Uses of remote sensing data for modelling forest structure attributes.*
- III. **2017-2018:** Internal Grant Agency (IGA) of Faculty of Forestry & Wood Sciences, CULS, Prague (No. A01/17). Title: *Modern approaches of remote sensing in forest mensuration applied in temperate forests.*

2) Ministry of Agriculture of the Czech Republic (No. QJ1520187) and Project FRAMEADAPT, Title: *Frameworks and possibilities of forest adaptation measures and strategies connected with Climate change.* No. EHP-CZ02-OV-1-019-2014.

Abstract:

This dissertation is based on several scientific articles and deals with the wide possibilities of Remote Sensing (RS) in combination with environmental data in order to increase key parameters for forest inventory. Specifically, paper I investigated the different approaches for crown projection estimation based on literature review. In papers II and III, the main idea was the investigation of indirect measurement of spatial forest parameters such as tree diversity and tree species distribution in stand level, using satellite images and environmental data. For this purpose, in paper II, different sources of data such as edaphic, climatic and topographic layers were used to estimate the diversity of trees in different microclimate conditions. In paper III, Quick-bird satellite images plus environmental data were used in order to classify the tree group species in different spatial conditions. In both papers II and III, data mining algorithms such as k-Nearest Neighbor (k -NN), Support Vector Machine (SVM) and Random Forest (RF) were used because of their potential in combining variety of independent variables simultaneously and also their performance in both classification and regression methods. Papers IV and V, evaluated the potential of Unmanned Aerial Vehicle (UAV) for estimation of individual tree parameters such as height and crown projection. Linear regression was used to calculate the accuracy of estimation. In case of paper IV, the main independent variable was the normalized Canopy Height Model (nCHM) which was the basis of height and crown projection estimation. Paper V, focused more on extrapolating the extracted data from UAV in a larger study area, using spectral correlation between UAV and satellite bands. Finally, in paper VI we developed a novel method for data acquisition based on close-range photogrammetric techniques, to determine the correlation between different environmental variables and regeneration of *Pinus sylvestris* L., in temperate forest conditions. We compared the UAV-based data with fisheye imagery as the standard method for potential solar radiation characterization for botany and forestry purposes.

The result of paper I, showed a summary of different approaches for crown projection determination using indirect methods of measurement. The results of paper II, showed the acceptable performance of data mining algorithms in modelling of tree species diversity in managed forest areas. In addition, its results showed that topographic layers, such as elevation and tangential curvature had the highest impact on diversity of tree species. The results of paper III showed, using environmental data especially

topographic combining with RS data could significantly increase the accuracy of modeling of tree species distribution. Additionally, RF algorithm was the most accurate non-parametric method between the chosen data mining algorithms. Paper IV and V, showed the potential of UAV to estimate tree parameters such as height and diameter with very high accuracy. Paper V showed the possibility of extrapolation of the estimated data, derived from UAV to a larger area using the relation between the Digital Number (DN) values of UAV and satellite data. Paper VI showed that UAV approach can be an accurate replacement method for fisheye camera.

Keywords: Forest parameters prediction, unmanned aerial vehicles, satellite images, linear regression, environmental data, down-scaling, data mining algorithms, solar radiation parameters

Abstrakt:

Tato disertační práce je založena na několika vědeckých člancích a zabývá se širokými možnostmi dálkového průzkumu Země (DPZ) v kombinaci s environmentálními daty pro odvození klíčových parametrů lesních porostů v inventarizaci lesů. Článek I porovnává publikované přístupy odhadu korunových projekcí stromů. Společným tématem článků II a III je ověření možností nepřímého měření prostorových parametrů lesních porostů, jako diverzita a prostorová distribuce dřevin na úrovni porostu na základě satelitních snímků a podpůrných environmentálních dat. Článek II popisuje využití různých datových vrstev, jako edafická, klimatická a topografická data pro odhad diverzity dřevin v různých mikroklimatických podmínkách. Ve článku III byly použity satelitní snímky systému Quick Bird spolu s environmentálními daty pro klasifikaci převládající dřeviny v různých podmínkách. Metodiky článků II i III využívají data mining algoritmy jako k-Nearest Neighbor (k-NN), Support Vector Machine (SVM) a Random Forest (RF), které byly využity z důvodu jejich značného potenciálu pro kombinování množství nezávislých veličin a vhodným vlastnostem pro provádění klasifikačních a regresních analýz. Články IV a V vyhodnocují potenciál bezpilotních letadel (UAV – Unmanned Aerial Vehicle) pro odhad stromových parametrů, jako výška a korunová projekce s využitím regresní analýzy pro vyhodnocení přesnosti odhadů. V případě článku IV byl jako hlavní vysvětlující proměnná pro odvození výšek a korunových projekcí stromů zvolen normalizovaný model výšek korun (nCHM – normalized Canopy Height Model). Článek V je zaměřen na extrapolaci dat získaných z UAV na oblast většího měřítka pomocí spektrální korelace mezi daty získanými pomocí UAV a satelitními snímky. Konečně článek VI popisuje novou metodu sběru dat založenou na fotogrammetrických postupech pro odvození korelace mezi faktory prostředí a zmlazením borovice lesní (*Pinus sylvestris* L.) v podmínkách lesů mírného pásma. Data získaná z UAV jsme porovnali se standardní metodou odvození potenciální solární radiace pomocí hemisférických fotografií používanou v botanickém a lesnickém výzkumu.

Výsledky článku I přináší shrnutí různých přístupů odvození korunové projekce pomocí metod nepřímého měření. Výsledky článku II demonstrují využitelnost data mining algoritmu pro modelování druhové diverzity v hospodářských lesích a zároveň ukazují, že z environmentálních faktorů mají na druhovou diverzitu lesních porostů největší vliv topografické parametry, jako nadmořská výška nebo tečná křivost terénu.

Výsledky článku III dokladují, že environmentální data, zejména topografické parametry, v kombinaci s daty DPZ mohou výrazně zvýšit přesnost odhadu dřevinného složení lesních porostů. Kromě toho přináší porovnání vybraných data mining algoritmů; z neparametrických metod dosáhl nejpřesvědčivějších výsledků algoritmus Random Forest. Články IV a V demonstrují potenciál využití UV pro odhad stromových parametrů jako výška nebo tloušťka s velmi vysokou přesností. Článek V dokladuje možnost extrapolovat odhady parametrů z UAV dat na území většího měřítka pomocí vztahu mezi pixelovými hodnotami UAV a satelitních snímků. Článek VI ukázal, že data získaná pomocí UAV mohou být plnohodnotnou náhradou hemisférických fotografií pro odhad potenciální solární radiace.

Klíčová slova: predikce lesních parametrů, bezpilotní letadlo, satelitní snímky, lineární regrese, environmentální data, down-scaling, data mining algoritmy, solární radiace

“Life is and will ever remain an equation incapable of solution, but it contains certain known factors.”

-Nikola Tesla

“We are stars warped in skin, the light you are seeking has been always within.”

-Romi

Table of Contents

Chapter 1	1
1. Introduction	1
1.1. Statement of the Problem and Motivation	1
1.2 Structure of the Dissertation	2
1.3. Contribution of the Dissertation.....	3
Chapter 2	4
2. Objectives	4
3. Literature Review	6
3.1. Significance of Scaling	6
3.1.1. Basic Concepts.....	6
3.1.2. Up-Scaling	7
3.1.3. Down-Scaling	8
3.2. Remote Sensing	10
3.2.1. Estimation of Forest Attributes Using Space-born Technology	10
3.2.2. UAVs in Forestry Applications	12
3.2.3. UAV-Photogrammetry in Forestry	14
3.3. Environmental Data	15
3.4. Solar Radiation	15
3.4.1. Direct Radiation.....	16
3.4.2. Diffuse Radiation:.....	16
3.4.3. Total Solar Radiation:.....	16
3.4.4. Direct Duration:	16
3.5. Image Segmentation	17
3.5.1. Principal-based Categories.....	17
3.5.2. Functionality-base Category	18
3.5.3. Statistical-base Category.....	20
3.5.4. Post-processing Required After Image Segmentation	20

3.6. Statistics	21
3.6.1. Linear Regression	21
3.6.2. Paired <i>t</i> -test	23
3.6.3. Confidence Interval for the True Mean Difference	24
3.6.4. Non-parametric Algorithms in Forest Parameters Estimation	26
3.6.5. Assessment of Accuracy	29
Chapter 4	31
4. Materials & Methods	31
4.1. Study Area Characterization	31
Chapter 5	33
5. Results	33
Chapter 6	128
6. Discussion	128
7. Conclusion	132
References	134

List of Figures

Figure 1. Illustrates the upscaling and downscaling process in space and the changes in grid density, as well as the relationship between grain and extent in remote sensing.....	9
Figure 2. Two basic categories of UAVs.....	13
Figure 3. Simulation of decision tree principle.....	20
Figure 4. Simple linear regression model	22
Figure 5. Three different hypothesis for mean of data.....	26
Figure 6. Illustration of random forest approach process stages.....	27
Figure 7. Illustration of SVM methodology for classification of unknown pixels	28
Figure 8. <i>k</i> -Nearest Neighbor approach principle.....	29
Figure 9. Locations of the study areas in the Czech Republic and Iran.....	32

List of Abbreviations

AGB	Above Ground Biomass
ALS	Aerial Laser Scanning
ANN	Artificial neural network
DEM	Digital Elevation Model
DN	Digital Number
DOF	Degrees Of Freedom
DSM	Digital Surface Model
DTM	Digital Terrain Model
GCS	Ground Control Station
GIS	Geographical Information System
HP	Hemispherical Photography
ITI	Individual Tree Identification
IWS	Inverse Watershed Segmentation
k -NN	k -Nearest Neighbor
LiDAR	Light Detection and Ranging
LSM	Least Squares Method
MAE	Mean Absolute Error
nCHM	Normalized Canopy Height Model
NFI	National Forest Inventory
RBF	Radial Basis Function
RF	Random Forest
RMSE	Root Mean Square Error
RS	Remote Sensing
SfM	Structure from Motion
SSE	Sum of Square Error
SVM	Support Vector Machine
TLS	Terrestrial Laser Scanning
UAS	Unmanned Aerial Systems
UAV	Unmanned Aerial Vehicle
VHR	Very High-Resolution

Included Articles

- I. **Abdollahnejad A.**, Panagiotidis D., Surový P. 2017: Forest canopy density assessment using different approaches – Review. *Journal of Forest Science*, 63: 107-116. <https://doi.org/10.17221/110/2016-JFS>
- II. **Abdollahnejad, A.**, Panagiotidis, D., Surový, P. 2016: Investigation of a possibility of spatial modelling of tree diversity using environmental and data mining algorithms. *Journal of Forest Science*, 62: 562-570. <https://dx.doi.org/10.17221/97/2016-JFS>
- III. **Abdollahnejad, A.**, Panagiotidis, D., Shataee Joybari, S., Surový, P. 2017: Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. *Forests*, 8: 42. <https://doi.org/10.3390/f8020042>
- IV. Panagiotidis D., **Abdollahnejad A.**, Surový p., Chiteculo V. 2017: Determining tree height and crown diameter from high-resolution UAV imagery. *International Journal of Remote Sensing*, 38: 2392-2410. <https://doi.org/10.1080/01431161.2016.1264028>
- V. **Abdollahnejad, A.**, Panagiotidis, D., Surový, P. 2018: Estimation and Extrapolation of Tree Parameters Using Spectral Correlation between UAV and Pléiades Data. *Forests*, 9: 85. <https://doi.org/10.3390/f9020085>
- VI. **Abdollahnejad, A.**, Panagiotidis, D., Surový, P., Ulbrichová, I. 2018. UAV Capability to Detect and Interpret Solar Radiation as a Potential Replacement Method to Hemispherical Photography. *Remote Sensing*. <https://doi.org/10.3390/rs10030423>

Chapter 1

1. Introduction

1.1. Statement of the Problem and Motivation

Natural resources considered as protector of human life and other living organisms, are essential to be under management. Recognition of phenomena and elements of forming of different sources is the first and most important step in optimized management. Identifying the available sources, monitoring their trends, changes and access to up-to-date data are the key factors in planning, decision-making and management in any field.

Repeated information based on newest satellites and Unmanned Aerial Vehicles (UAVs) improvements, significantly contributed to our understanding of the dynamics, in such complex ecosystems as forests. Volume and Above Ground Biomass (AGB) or any other interested parameter of forest in small spatial extent areas, can be derived with the help of forest inventory data. However, field measurements are typically time-consuming and expensive. On the contrary, in higher spatial extent areas, modelling of forest parameters requires the use of Remote Sensing (RS) information, mainly from practical point of view. Also, RS techniques are able to improve the value of inventoried data with detailed coverages in affordable costs (Tomppo et al. 2008, McRoberts et al. 2010).

Modern techniques of RS can be also used to provide accurate tree height and crown area characteristics, based on individual tree level assessment using a series of algorithms (Carleer et al. 2005) for estimation of crown diameter based on the Individual Tree Identification (ITI) (Edson and Wing 2011), tree height based on smoothing of Canopy Height Model (CHM) and local maxima techniques (Pyysalo and Hyypä 2002, Popescu et al. 2003).

In the past, many studies used different spectral data with variety of spectral and spatial resolution in order to estimate forest parameters in stand or individual tree

-Chapter 1: Introduction-

level. Based on used material, methodology, the results of modeling and characterization of an object, can vary. However, some studies have already addressed the challenges in accurate forest parameters estimation; using only spectral data. Improvement in technology and classifier algorithms allows us to combine different sources of data, such as topographic parameters, soil index, vegetation index and different composition and proceed bands in order to increase the accuracy of modeling.

This dissertation is based on six scientific articles. Each one of them focus on some of the main concepts and challenges of modern RS and it is trying to clarify the potential improvements from methodological point of view for the determination of forest parameters with higher precision. My ambition is the development of innovative algorithms for solving problems of modern RS methods. However, there are a few challenges that need to be addressed a) comparing to traditional methods, how accurate the methods of RS are in indirect estimates of forest parameters under different circumstances (i.e., forest structure) and b) how to overcome the spatial limitation of UAVs or spectral limitation of satellite data, when predicting forest parameters over large forested areas.

1.2 Structure of the Dissertation

The presentation of this dissertation has been divided into the following sections:

1. **Introduction:** Describes the motivation for writing this dissertation by explaining the problems and necessity of using RS data.
2. **Literature review:** Describes the necessary theoretical background of previous studies and surveys which can help to have an overview of the research work.
3. **Material and methods:** Describes the material, data collection and processing which had been used for the fulfillment of this dissertation.
4. **Results:** Presenting the articles.
5. **Discussion:** Presenting the results in comparison with others published researches and explaining the nature and behavior of the data.

6. **Conclusion:** Describes a short summary of main results that can be used as the base of further studies.

1.3. Contribution of the Dissertation

1. It will help to assess the potential of indirect measurement methods (RS techniques) in different purposes of forest management.
2. Introducing modern measurement methods with lower costs and less time consuming than the traditional methods.
3. In general, RS approaches allows managers to maintain up-to-date information of forest parameter changes over some period of time.
4. Used down-scaling approaches will give solution in extrapolating crucial information over large forested areas.
5. The developed methodology will eventually incorporate into the forestry practice for using different sources of data such as: edaphic, climatic, topographic, satellite data, UAV and close-range photogrammetry for estimation of significant parameters.

Chapter 2

2. Objectives

This dissertation contributes to verify which kind of RS data in combination with auxiliary data can provide accurate independent variables and how much these studies can be practical as a basis for forest management plans and strategies.

- ❖ Assessment of the best methods for crown projection estimation based on extensive literature review.
- ❖ Evaluation of effectiveness of environmental factors, such as topographic, edaphic and climatic parameters on spatial distribution of tree diversity.
- ❖ Assessment of performance of spectral and non-spectral (environmental factors) variables combination for modeling the spatial distribution of forest tree species.
- ❖ Evaluation of the performance of nonparametric algorithms in modelling forest attributes in stand level.
- ❖ Study of the capability of UAV for estimation of tree parameters such as height and diameter and crown projection in individual tree level.
- ❖ Extrapolation of estimated parameters derived from UAV by using calibrated satellite images.
- ❖ Investigation of possibility of using UAV data as a replacement method to the Hemispherical Photography (HP), for modeling potential solar radiation components. What extent can the methodology be used as an efficient and accurate source for forest management plans and strategies?

-Chapter 2: Objective-

This dissertation is a completion of multiple and interrelated assignments as suggested in my study plan at the beginning of my doctoral study. More specifically:

Task 1 – Literature review on the state of the art described in chapter 3 and Paper I.

Task 2 – Field data acquisition was done on papers II, III, IV, V and VI.

Task 3 – Large-scale imagery acquisition, supported mainly by paper III and V.

Task 4 – UAV data collection in paper IV, V and VI

Task 5– Pre-processing of spectral and non-spectral data using a series of different software for all case studies (papers II, III, IV, V and VI).

Task 6 – Evaluation of accuracy of estimated data by using statistical indicators (papers II, III, IV, V and VI).

Task 7 – Presenting the results in form of scientific articles in chapter five.

Task 8 – Describing the discussion part analytically in chapter six, where the finding of these studies have been presented and have been compared with previous studies of others.

Chapter 3

3. Literature Review

3.1. Significance of Scaling

3.1.1. Basic Concepts

Pixel: We can define pixel as the smallest controllable element or unit in a raster image, which is represented as a point or dot and which is normally square (Goesele 2004).

Grain: Is nothing but the minimum spatial resolution of the data. In other words, is representing the size of individual units of observation. An easy way of understanding this concept has been given by (Turner et al. 1989), for instance a fine grained map might result structural information into 1 ha, whereas in coarser resolution map we would have structured information into 10 ha.

Extent: Is typically the spatial domain or the size of the study area which the information is available (Delcourt et al. 1983).

Fine Spatial Resolution: Are usually high-resolution images, where small objects in raster image can be detected with much detail (Graf 1999).

Coarser Spatial Resolution: On the contrary, the resolution of an image is coarse when large features are visible. Consequently, these objects are of lower resolution without much of detail (Graf 1999).

Scale: Refers to the largest time interval [T], area [L2] for which the property of interest is considered homogeneous. However, for this time intervals and areas we don't have any information for its variation within, but only for the average value of the property (Bierkens et al. 2000).

Scaling: Is the process of information transferring across different scales (Pelgrum et al. 2000).

Spatial Resolution: Is the concept which refers to the ability of the sensor, to detect the smallest feature that we can observe in an image (Gibson 2010).

Temporal Resolution: Refers to the time which is needed to revisit and acquire data (time step) from the exact same location. It can be high when the time interval of revisiting is short and vice-versa. Therefore, often it is expressed in days (Gibson 2000).

3.1.2. Up-Scaling

Up-scaling or aggregation refers to the conversion of data which has fine spatial resolution to a coarser one. To better illustrate the term, upscale refers to information transfer from a local to a larger scale. Up-scaling increases the scale of the research area (Atkinson 2013). The latest advances have made the use of various RS modalities a powerful approach in quantifying multiple forest attributes, by providing a powerful resource for up-scaling process, understanding to regional and continental domains (Masek et al. 2015). Up-scaling by theory, involves the necessity to discard some level of detail, otherwise the amount of data which remains is either too much in order to be used or simply obscures the relevant emergent phenomena, which are apparently part of only larger spatial scales. Usually, the loss of information in terms of image detail, results reduction in the number of Degrees Of Freedom (DOF).

Scientists arguing that up-scaling is distinct from an averaging process, since it involves applications of additional scale models and therefore is not regarded as a pure. However, up-scaling as a process of averaging, has the potential to be applied in any heterogeneous territory successfully (Wood 2008).

In practical terms, aggregation can be conducted either a) by calculating the output, for instance (E) for each pixel at small scale and averaging the result (so called output up-scaling) or through aggregation of input data at smaller scale and then by calculating the output (E) on the resulting average (input up-scaling). Some of the main characteristics of up-scaling is that a) the number of pixels are increasing as the spatial resolution becomes coarser, b) it leads to a reduction in the variance and the range of data. However, using coarser data can be valuable in some cases since we avoid oversampling the detection of unnecessary variation within a feature

for instance, i.e., forest stand and no single trees (Quattrochi and Goodchild 1997). An example of using effectively RS for up-scaling forest attributes, is given by De Jong et al. (1999) when he tried to develop an index of large scale biomass estimate with the help of book-keeping approach, by using extrapolating measured per area carbon stock density values (based on satellite imagery data of 30 m multispectral spatial resolution) to the spatial extent of the main forest covering types.

3.1.3. Down-Scaling

Down-scaling or disaggregation is the opposite process from up-scaling, it is the transferring of information from large scale to a local or smaller scale. Respectively, down-scaling decreases the scale of the research area (Atkinson 2013).

RS is a valuable source of data for any kind of environmental resource management planning and decision making. Forests are providing all the services which are necessary for our survival, services relevant to society, industry and climate. Nowadays, there is a higher need of understanding the boarder implications of the knowledge which we have in local-scale about ecological processes, in order to sustain these services.

Scaling is a hierarchical scale structure, which can basically reduce not only the complex descriptions in a class, but their redundancy as well (Baatz et al. 2003). One of its basic characteristics, during the interpretation and analysis process is the key factor “spatial resolution” (i.e., same type of objectives appears in a different way when examine them at different scales). Scale research, is crucial in terms of let’s say validation of RS products (Hufkens et al. 2008) and the reliability of the data are highly related and depend on the chosen scaling method as it can be also seen in (Figure 1) (Hay et al. 2001). Scale plays the role of measuring tool or filter, it represents the `window of perception` with which a system is able to viewed and quantified (Hay et al. 2001).

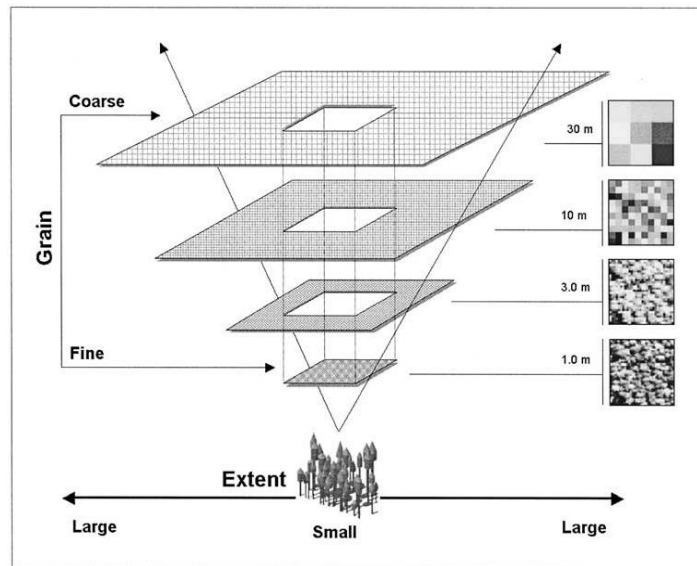


Figure 1. Illustrates the upscaling and downscaling process in space and the changes in grid density, as well as the relationship between grain and extent in remote sensing. Source: Hay et al. 2001.

The intention and the aim of multi-scale studies (up-scaling and down-scaling) is to analyse the relationship between data which acquired in different extent or in different spatial resolution, in effective manner. The purpose of such studies is basically to examine and understand the behaviour of variables, while they are changing scale.

Scale is a real issue when relating different datasets in Geographical Information System (GIS), and relates both the level as well as the extent of detail. Basically, a map for example in GIS is a single scale product, whereas a GIS is able to operate at multi-scales and accommodate multi-scale data. Another way to better comprehend scale is to consider it both as spatial and temporal, were different levels of information relate to different scales. Moreover, we should be in position to answer how to effectively relate the data which acquired at different scales.

For describing forest attributes, independently of which scaling approach is about to be used, there is always a requirement for at least one geospatial map. Either our decision will be based on simulation modelling (Melillo et al. 1993) or traditional approaches like “book-keeping” approach (i.e., Houghton and Hackler 2003), the basic concept of scaling is always to associate the forest type with a particular parameter, where it was measured and then to extrapolate its local value according to the spatial pattern or areal extent of that type across the mapped landscape.

Decision on the optimum scale can be given with the help of geostatistics. The geometrical shape of a variogram can be used in order to indicate the scale of a particular process in the environment (Foody et al. 2004).

Large scale data imagery is in most of the cases nearly impossible to obtain the desired detail which is needed, for instance if we are interested to estimate tree characteristics based on tree level approach. However, even with the latest technological advances, satellites with capability of only a few meters are not in position to give such a necessary information and precision. Attempts to improve the spatial resolution and eventually to be able to extract the precision that we need in low-cost, turned down-scaling into a practical approach. One of the most effective approaches is based on data-fusion techniques, where lower resolution images are enhanced by using information taken with higher spatial resolution sensors (i.e., UAVs). We could describe it as a linking approach between two different level of data acquisition (i.e., UAVs and satellites) by addressing trades-off between them, using for instance UAVs as small calibration measurement for satellite pixels. A similar approach proposed by Du et al. (2014), when he tried to produce a spatially-explicit map of forest biomass derived from MODIS satellite that would matches the data based on forest inventory. The results showed the effectiveness of proportionate down-scaling of using statistical forest data to forest cover pixels (based on the calibrated forest cover proportional maps) for creating forest biomass maps of finer resolution.

3.2. Remote Sensing

3.2.1. Estimation of Forest Attributes Using Space-born Technology

Several studies over the last years have shown the potential of RS methods compared to the traditional field methods for enhancement of inventory data, providing large coverages at low costs (McRoberts and Tomppo 2007, Salajanu and Jacob 2008, Tomppo et al. 2008, McRoberts et al. 2010, Mohammadi et al. 2010, Mohammadi et al. 2011). Nowadays, there are several techniques of RS that can be used for modeling forest terrain and estimate detail tree parameters, among the most popular techniques is Aerial Laser Scanning (ALS) (Hyypä et al. 2000, Næsset et al. 2004, Yu et al. 2015). The ability of ALS relies on the construction of point clouds

-Chapter 3: Literature review-

of crown canopy with high density. Similarly, from below canopy use of Terrestrial Laser Scanning (TLS) (Erikson and Karin 2003, Yu et al. 2011, Newman et al. 2016), can be used for stem modelling (Fritz et al. 2013, Forsman et al. 2016).

Recently, photogrammetry for forest parameters estimation has been a popular technique for RS, it can be used either from the air (aerial photogrammetry i.e., Rosnell and Honkavaara 2012) or ground (terrestrial photogrammetry i.e., Mikita et al. 2016).

However, all the above methods have been proven rather expensive approaches, especially in the case of laser technology. In addition, airborne platforms are not always available to many people and that can be considered as limiting factor. An alternative solution to these kind of limitations is the use of satellite sensors. Several studies in the past used medium spatial resolution satellite sensors i.e., SPOT and Landsat (~10-30 m pixel size), these studies have shown the potential for regional-scale data acquisition of forest attributes (Wulder et al. 2008). In short period of time development of higher spatial resolution satellites, like for instance Quickbird (0.6 to 2.8 m²) were used for determination and distinguishment of species classification with more detailed information. The latter is able to distinguish individual trees compared to the medium resolution imagery.

Latest satellite innovations include Very High-Resolution (VHR) sensors (i.e., Pléiades 1A with panchromatic resolution of 0.5 m, WorldView-2 etc.) So far, only a small number of studies have used VHR for estimation of forest parameters. For example, in their work Ozdemir and Karnieli (2011) used WorldView-2 satellite for estimation of several structural parameters over 29 pine plantations of 30 x 30 m in Israel. Their results showed that VHR recommend for this purpose and specifically can be used efficiently for the extraction of stocking percentage with 110 stems· ha⁻¹ (RMSE = 29%), Basal area (BA) with 1.8 m²· ha⁻¹ and (RMSE = 17%), Vol. with 27 m³· ha⁻¹ (RMSE = 44%), respectively.

In another study, WorldView-2 (in Germany) and Pléiades (In Chile) data sets have been used for estimating AGB (Maack et al. 2015). Their results showed, that combination of spectral derivatives with height metrics had an RMSE = 24%

-Chapter 3: Literature review-

resulting $47 \text{ tons}\cdot\text{ha}^{-1}$ in Germany and $\text{RMSE} = 36\%$ $59 \text{ tons}\cdot\text{ha}^{-1}$ in Chile respectively.

Other works such for instance Immitzer et al. (2016), have shown that combination of VHR satellite sensors with National Forest Inventory (NFI) data in temperate forests of central Europe, can be efficiently used for the construction of wall-to-wall forest cover maps with increased spatial resolution.

3.2.2. UAVs in Forestry Applications

Nowadays, the role of UAVs have become a practical solution for acquisition of forest parameters with high fidelity (Grenzdörffer et al. 2008, Jaakkola et al. 2010, Tahar et al. 2011, Gatziolis et al. 2015). Other synonym definitions have been given to describe these systems, UAVs or drones are basically flying systems that controlled from the ground using Ground Control Station (GCS) and therefore they referred as Unmanned Aerial Systems (UAS) (Colomina et al. 2008). According to Eisenbeiss 2008c, these systems are able to operate either autonomously or semi-autonomously. UAVs can be distinguished in two basic categories (Figure 2):

- a) The fixed-wing b) Rotary-wing (copter)



(a)



(b)

Figure 2. Two basic categories of UAVs, Source (a): <http://aeroexpo.online/prod/sensefly/product-175240-1787.html>

Another categorization of UAVs is the one which has to do with their operation, basically the majority of drones which widely used are based on electric engines. However, for longer flight time, range and heavier cargo, fuel injected engines are more eligible and preferable. Also, the determination of type of engine is usually interwoven with the UAV size, which can be a) very small b) small c) medium and d) large. Some of the main differences between fixed-wing and rotary UAVs are:

- ❖ Fixed-wing, can take-off and land almost autonomously, on the other hand fixed-wing demands catapult for take-off, while for landing process in most of the cases it requires hands for catching it.
- ❖ Rotary-wing can practically take-off and land without demanding lot of space compared to fixed-wing platforms.
- ❖ Rotary-wing has less aerodynamics potential compared to fixed-wing. Essentially, it can fly faster by covering longer ranges.

Wallace et al. 2016, pointed out the significance of small UAVs (weight ~5 kg) as inexpensive systems in forest research as well as the several options for deploying different types of sensors (i.e., from typical consumer grade photo camera to thermal or Light Detection and Ranging (LiDAR) sensors).

Even the price range for a survey-grade UAV is wide, depending on many parameters, it is considering as feasible economically approach.

3.2.3. UAV-Photogrammetry in Forestry

Photogrammetry is a method that permits acquisition of information from an object that has a particular distance. It is simply a technique for recovering exact positions of surface points based on electromagnetic spectrum (Wolf and Dewitt 2000).

Several studies so far pointed out that information of individual trees over large forested areas can be accomplished practically using RS approaches. Prediction of key forest variables can be fulfilled through space and airborne imagery. Spatial resolution and covering area is one of the most significant challenges in remote sensing. Assessment of detail information such as crown diameter for instance derived in individual tree level is beyond the ability of satellites so far. Therefore, UAV platforms equipped with inexpensive and VHR sensors, in combination with low altitude flights can overcome the spatial resolution problems of space-borne imagery. Data acquisition from photo-camera deployed on UAV provides a significant high degree of image overlapping, which is a fundamental concept in photogrammetry (Remondino et al. 2011). Automated photogrammetric systems and Structure from Motion algorithms (SfM) based on computer Vision (Snavely et al. 2008), enables firstly the possibility to retrieve the camera position at the time when the image was captured and secondly to estimate for each raw image an orthographic rendition, allowing reconstruction of 3D objects from 2D images (Rosnell and Honkavaara 2012, Dandois and Ellis 2013). However, with nadir-oriented images based on cameras from UAVs, the majority of points in the point clouds are always positioned near the crown of the trees. This fact is significant, especially in case of lower parts of the crown since the representation of crown sides is always tending to be sparse, containing sizeable gaps, essentially it results difficulties in quantifying lateral space crown competitions. Several studies during the last years, have pointed out the ability of Photoscan© software from Agisoft company for the reconstruction of forest terrain and the extraction of high quality digital surface and terrain models.

Produce of high quality UAV-based Digital Surface Model (DSM) and Digital Terrain Model (DTM), allows detection of individual tree positions and estimations of individual tree crowns (Lisein et al. 2013). Furthermore, photogrammetry based on UAV unlocked a completely new field in the close range applications as a real or better near real time approach (Kerle et al. 2008).

3.3. Environmental Data

Since tree growth, abundance or diversity and distribution of them is directly related to many parameters including the amount of solar radiation, climatic characteristics, water, topography and soil nutrients, it is possible to use local parameters to assess the spatial characteristics of forest trees (Hunter 2003).

Derived properties from Digital Elevation Model (DEM) were used in many studies related to classification or prediction of soil characteristics or vegetation cover or distribution of species in different scales (Wilson et al. 2007)

Topographic parameters, light, temperature, water and other environmental factors, generally has a significant impact on distribution and growth. Under different environmental condition, the density, type, productivity and other forest properties are different. For this reason, different forests exist in different regions.

Elevation changes can have a big influence on local climate and other ecological variables. Noted that climate can change the habits of fauna and flora of a region. Per 100 meters increasing the altitude, temperature decreases 6 °C. Changing the temperature equally affects the ecological factors (such as relative humidity, growing seasons and evapotranspiration).

3.4. Solar Radiation

Incoming global solar radiation is a key factor which influences energy and water balance and thus is fundamental to most biophysical and physical processes (Fu and Richm 2002, Fournier et al. 2003). Identifying factors that influence variation in light availability within forested ecosystems represents an important component in our understanding of the complex determinants of tree seedling regeneration, growth, and increment. In addition, solar radiation is one of the most influential

-Chapter 3: Literature review-

independent variables in predicting the spatial distribution of species groups (Peffer et al. 2003, Abdollahnejad et al. 2017). Within forest stands, variation in vegetation composition, structure, crown dimensions, and foliage distribution create spatial variation in light transmittance, affecting growth dynamics. Characterization of the solar radiation regime and forest canopy structural architecture has undergone considerable evolution since (Evans and Coombe 1959, Anderson 1964) first reported using HP.

Solar radiation is a general term for the electromagnetic radiation emitted by the sun. It is a kind of radiation including visible light, radio waves, gamma rays, and X-rays, in which electric and magnetic fields vary simultaneously. The unit power is the watt (w) which is usually measured per unit area. Solar radiation is typically quoted as W/m^2 - that is Watts per square meter (Boxwell 2012). The main components of solar radiation are defined as follow:

3.4.1. Direct Radiation

Direct solar radiation is used to describe solar radiation traveling on a straight line from the sun down to the surface. Also, it can be described as the amount of electromagnetic radiation which directly reach the surface of the earth.

3.4.2. Diffuse Radiation:

Diffuse radiation describes the electromagnetic radiation that has been scattered by molecules and particles in the atmosphere but that has still made it down to the surface of the ground.

3.4.3. Total Solar Radiation:

Total solar radiation parameter can be calculated by summarizing the direct and diffuse solar radiation.

3.4.4. Direct Duration:

It is a climatological indicator which represents duration of direct incoming solar radiation in given period (usually, a day or a year) for a given location on Earth.

3.5. Image Segmentation

Image segmentation process considered as one of the most significant process in image analysis for the extraction of particular objectives (Zhang et al. 2008). Essentially this process gives to the analysts a better view regarding the areas of interest (Freixenet et al. 2002). In another words, image segmentation is a fundamental tool for many RS applications were one is often interested in reducing the complexity of RS image to a limited number of homogenous classes that may represent, for example different vegetation or land cover types. The classification of pixels into specific classes is commonly based on recognition of their characteristics.

Latest technological innovations in computer vision enhanced the potential of different and more sophisticated algorithms to estimate different kinds of forest key attributes such as crown projection, height, volume etc. with better precision (Tomppo 1987). Image segmentation methods can be divided in variety of categories from different point of view (Table 1).

Table 1. Different categories of image segmentation

	Statistical	Principal	Functionality
1	Parametric	Sub-Pixel-wise	Supervised
2	Non-parametric	Pixel-wise	Un-supervised
3	Non-metric	Shape base or object base	-

3.5.1. Principal-based Categories

In the past, due to the fact that satellite images were of low spatial resolution, it was necessary to use specific algorithms that can work with mixed values in sub-pixel-level. These algorithms can be divided as 1) fuzzy classification (Zadeh 1965), 2) spectral un-mixing (Boardman 1994).

Furthermore, purpose of these algorithms is to break the pixel in subpixels in order to enhance particular areas of interest. Nowadays, other types of algorithms such as object base and pixel base can be used for analysing high-resolution aerial

-Chapter 3: Literature review-

images (space and UAV-born). Some of the base principles of these methods (object base) are;

1) They are able to divide images into segments by using:

- ❖ Region growing algorithms
- ❖ Markovian
- ❖ Watershed
- ❖ Hierarchical algorithms

2) By using other algorithms can specify a label for each segment such as:

- ❖ Support vector machine (SVM)
- ❖ Decision tree
- ❖ Genetic algorithm
- ❖ k -Nearest Neighbour (k -NN)

For identification of the object, object base algorithms use different types of information like 1) spectral, 2) conceptual 3) spatial 4) texture, including tone, shape, size, pattern, texture, shadow, association and compactness. While, in pixel-base methods just the spectral data is critical.

3.5.2. Functionality-base Category

Un-supervised methods: These methods work based on the distance between pixels in multi spectral space. The user does not have any control or influence on classification process and these methods does not need any sample training. Calculation of distance is based on two main ways:

1) Euclidean Distance

$$\begin{aligned} \text{Euclidean distance } d(x_1 - x_2) &= \|x_1 - x_2\| \left\{ \frac{(x_1 - x_2)}{(x_1 - x_2)} \right\}^{\frac{1}{2}} \\ * &= \left\{ \sum_{i=1}^F (x_{1i} - x_{2i})^2 \right\}^{1/2} \end{aligned} \quad (1)$$

2) First Norm (L1)

$$d = (x_1 - x_2) = \left\{ \sum_{i=1}^F |x_{1i} - x_{2i}| \right\} \quad (2)$$

-Chapter 3: Literature review-

The accuracy of euclidean distance is higher than first norm. For evaluation of accuracy in un-supervised methods two following approaches can be used:

- 3) Sum of Square Error (SSE):** By using, the difference between the pixel value and mean of values of its class.
- 4) Using Amount of Class's Heterogeneity:** Comparing mean of classes with total mean and calculating the mean of variance and covariance of classes.

Algorithms such as ISO-Data, k-mean, Chain and single linkage are some examples for un-supervised algorithms.

Supervised methods: Supervised classifier algorithms needs sample training and user can influence the process of classification. Some of the most common algorithms are:

1) Parallelepiped algorithm (Multi alignment):

The accuracy of this method is related to the min and max of values. In the classes that are not completely parallel with axis, accuracy will be low.

2) Minimum Euclidean distance classifier

This method is based on minimum spectral dimension distance between pixel value and mean of values in classes. It means the unknown pixel will belong to the class with the most similar value.

3) Mahalanobis distance algorithms

This algorithm is similar to minimum Euclidean distance classifier, but in addition, it uses also, the variance and covariance matrix in classification.

4) Maximum likelihood

The main pre-condition for this method is normal distribution of training samples for classes.

5) Decision tree

-Chapter 3: Literature review-

This method hierarchically divide the space into smaller part and fit a model for each part each tree includes: a) root, b) node c) leaf (Figure 3). The critical rules of decision in each node are:

- a) Gini coefficient
- b) Entropy index

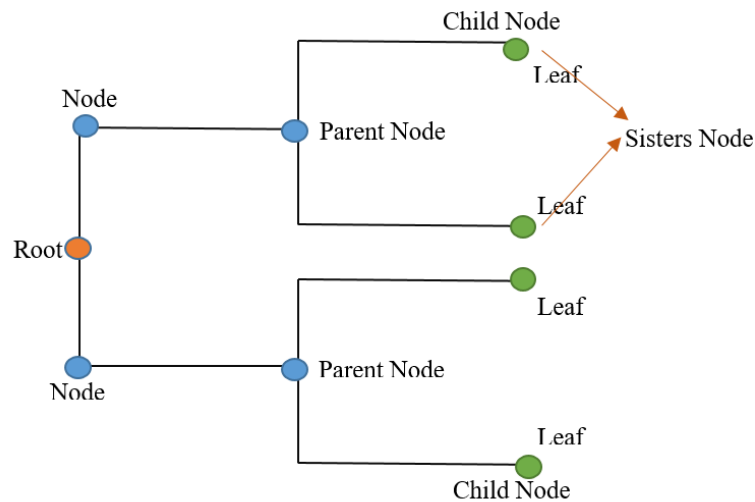


Figure 3. Simulation of decision tree principle

- 6) *Support vector machine (SVM)*
- 7) *Random forest (RF)*

3.5.3. Statistical-base Category

In order to avoid repeating the same topic, this module is analytically described below in the module of statistics, particularly in 3.5.

3.5.4. Post-processing Required After Image Segmentation

After classification three main options can help users to improve the results. Depend on the nature of classified value, users can apply one of the following options on their results:

- 1) *Fusion* : for reducing the dispersion
- 2) *Elimination* : for mixing small clases to other classes with same label
- 3) *Division*: for making classes with smaller variance

3.6. Statistics

In case where one scientific question is to be examined by comparing two or more groups of data, one statistical test can be performed. It means that a null-hypothesis must be considered, which can in principle be rejected. Further, an appropriate test parameter must be established (Du Prel et al. 2009).

First, the scientific question of study and null hypothesis must be formulated. Before fulfilling the study, the test and the level of significance must be determined. Study design, data structure and the scientific question that needs to be answered are the most important factors for the decision for choosing a statistical test (Röhrig et al. 2009).

Two rules are deterministic for the selection of the statistical test:

- 1- The scale of evaluation of the test variable (continuous, binary, categorical)
- 2- The type of study plan (paired or unpaired).

3.6.1. Linear Regression

The aim of regression analysis is to investigate the relative impact of an independent variable on a specific result (Applegate and Crewson 2002). Linear regression tries to find the best fitting line through the measured points. The regression line is made by the predicted y for each x possible value of x . The distance between the points (measured Y) to regression line is the error of prediction (Salehi and Seber 1997) (Figure 4, Formula 3).

$$y_i = a + bx_i + e_i \quad (3)$$

In the above formula (a) is the intercept (on the y axes), (b) is the slope of the regression line and (x) is the independent or explanatory variable. The random error term e_i is assumed to be uncorrelated, with a mean of zero and constant variance (Salehi & Seber 1997).

-Chapter 3: Literature review-

Least Squares Method (LSM) is frequently used and the simplest way to estimate the amount of intercept and slope in this line. LSM consist of minimizing the squared residuals of difference between observed Y and estimated Y . In the end the fitted Y can be used for estimating the amount of Y for all possible amount of independent variables (Neter et al. 1990, Stigler 1978).

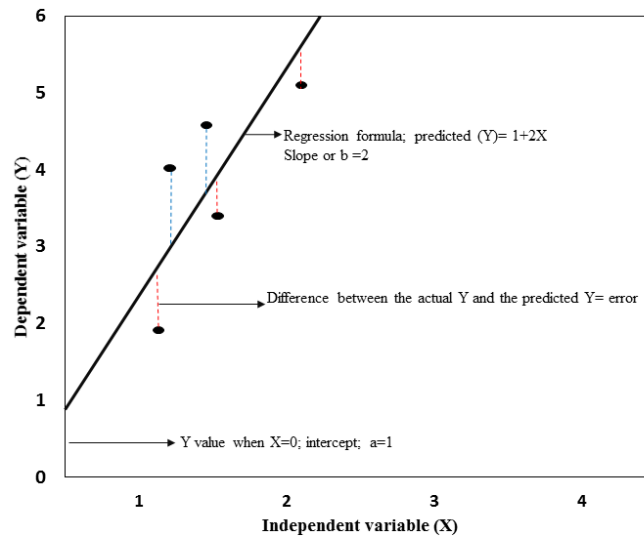


Figure 4. Simple linear regression model. Source:

https://docs.oracle.com/cd/B28359_01/datamine.111/b28129/regress.htm#DMCON005

3.6.1.1. The Criteria for Using Linear Regression

According to (Neter et al. 1990, Nelson 1982), the criteria for using linear regression are:

- ❖ It needs relationship between dependent and independent variables, also the linear regression is very sensitive to effect, scatter plots can be used for linearity assumption.
- ❖ Data must have normal distribution, the best ways for checking the normality is goodness of fit test e.g., the Kolmogorov-Smirnov test. In the case where data are not normally distributed a non-linear transformation might fix this problem.
- ❖ Little or no multicollinearity in the data. Multicollinearity happens when the explanatory variables are not independent from each other. Another important independence assumption consist of the error of the mean has to be

independent from the explanatory variables. Correlation matrix, tolerance, variance inflation factor and condition index can test the multicollinearity. If this issue is found, conducting a factor analysis and rotating the factors to assure independence of the factors in the linear regression analysis can be used for solving the multicollinearity in the data.

- ❖ Linear regression needs little or no auto correlation in the data, this issue happens when the residuals are not independent from each other. Durbin-Watson test and scatter plots can be used for testing the autocorrelation in the data.
- ❖ No homoscedasticity in the data. Goldfeld-Quandt Test and scatter plot can test this issue. In case that homoscedasticity is present, a non-linear correction might fix the issue.

3.6.2. Paired *t*-test

A paired *t*-test is commonly used to compare two mean in two related samples. The null hypothesis consist of equality between two means. The principle of *t*-test is using the difference between the two observations on each pair. *T*-test uses this difference in order to calculate the mean, standard deviation and standard error of mean differences (Snedecor and Cochran 1989, Moore and McCabe 2006). In the end the *t* statistic is calculated which is given by:

$$T = \frac{\bar{d}}{SE} \quad (4)$$

Which: $T = t$ student, \bar{d} = mean difference SE = standard error of mean difference.

One of two outcomes will be considered:

- 1- $t < t_{critical}$ - Tests are accepted, as the differences between the test results are statistically likely to occur.
- 2- $t \geq t_{critical}$ - Tests are not accepted, as the differences between the test results are greater than is statistically likely to occur (Posten 1978).

*Value for *t*-critical is from the table basis on degree of freedom (n-1)

3.6.2.1. The Criteria for Using *t*-test

According to (Posten 1978, Hill and Lewicki 2007, Statistica 2010) the criteria for using *t*-test are:

- ❖ The dependent variable must be continuous (interval or ratio level).
- ❖ The independent variable must consist of two categorical, "related groups" or "matched pairs". "Related groups" represent that the same subjects are present in both groups. In other words, the number of samples in two groups of data must be the same.
- ❖ For sample that was taken randomly or from two different population, independent test must be used (unpaired), even if the number of data points in each set is the same.
- ❖ No significant outliers in the differences between the two groups. Outliers might have negative influence on paired *t*-test and change the differences between the two related groups (whether increasing or decreasing the scores on the dependent variable), and in overall it can decrease the accuracy of the results. Also, the statistical significance of the test can be affected by the outliers.
- ❖ It needs normally distribution of the differences in the dependent variable between the two related groups. Shapiro-Wilk test can be used for testing the normality.

3.6.3. Confidence Interval for the True Mean Difference

The confidence interval associated with a paired *t*-test (Snedecor and Cochran 1989). The confidence interval is an expression of that unreliability, expressing an area around the sample mean that we think the population mean is probably to fall. It would be helpful to calculate the limits of the true difference of mean (in practice 90%, 95%, and 99% intervals are often used in researches, with 95% being the most frequently used. A 95% confidence interval for the true mean difference is:

$$\bar{d} \pm t \times \frac{s_d}{\sqrt{n}} \quad \text{or} \quad \bar{d} \pm t \times SE \quad (5)$$

Where : \bar{d} = mean difference t = (t-distribution on n-1 degree of freedom), s_d = standard deviation and SE = standard error of mean difference.

-Chapter 3: Literature review-

This formula shows that the width of the interval is controlled by two factors:

1. As much as the size of samples (n) be larger the amount of confidence interval will be smaller. In fact, increasing the sample size can cause more precise estimate for mean.
2. Larger standard error afford larger confidence interval. Simply outlier data and noisy data with larger standard deviation give the wider amount of confidence interval.

Confidence interval can be one or two sided. A one-sided confidence interval shows the limitation of the population parameter either from above or below. While, a two-sided confidence interval shows the limitation of the population parameter from above and below (Agresti and Coull 1998). The rejection regions for three possible alternative hypotheses are shown in Figure 5 and Table 2 (p-value = 0.05).

Table 2. Three possible alternative hypothesis for mean source: (Agresti and Coull 1998).

Alternative Hypothesis	Rejection Region
$H_a: \mu \neq \mu_0$	$ T > t_{1-\alpha/2, v}$
$H_a: \mu > \mu_0$	$T > t_{1-\alpha, v}$
$H_a: \mu < \mu_0$	$T < t_{\alpha, v}$

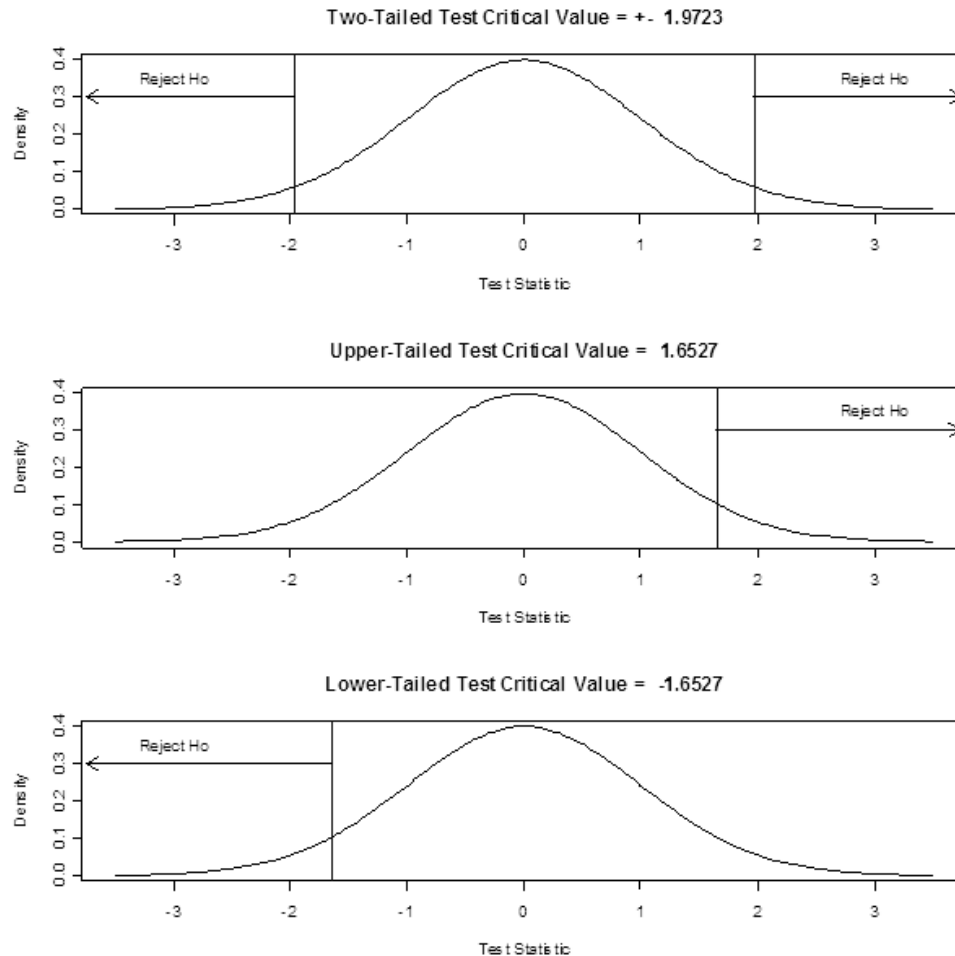


Figure 5. Three different hypothesis for mean of data, Source: <http://www.itl.nist.gov/div898/handbook/eda/section3/eda352.htm>

3.6.4. Non-parametric Algorithms in Forest Parameters Estimation

Even though use of parametric algorithms for modeling RS data has been systematically used for the estimation of several forest parameters (Sivanpillai et al. 2006), the strength of data mining algorithms emerged through several studies and depicted as extremely useful in estimation of forest parameters, mainly because of their flexibility to describe data of non-linear dependencies (McRoberts et al. 2007). Furthermore, fundamentals of data mining algorithms, have some advantages over parametric algorithms, these advantages are:

- ❖ Wide flexibility range of non-linear dependencies.
- ❖ Free of assumptions (Sironen et al. 2010).
- ❖ Additivity, meaning that observations has no dependency to each other.

-Chapter 3: Literature review-

Among the most popular non-parametric algorithms widely used so far, we can distinguish:

- Random Forest (RF)
- Support Vector Machine (SVM)
- K-nearest neighbor
- Artificial Neural Network (ANN)

3.6.4.1. Random Forest

It is considered as one of the most common data-mining algorithm. It is particularly effective for a) solving problems which are referring to classification, for predicting categorical dependent variable (Shataee 2004, Walton 2008, Shataee et al. 2012) and b) solving regression problems for prediction of dependent forest variables (Breiman 2001, Breiman 2002, Breidenbach et al. 2010). Moreover, RF can give different weights according to the importance of layers in modelling process. Finally, both variables and data can randomly sampled in order to construct forest of regression trees (Figure 6).

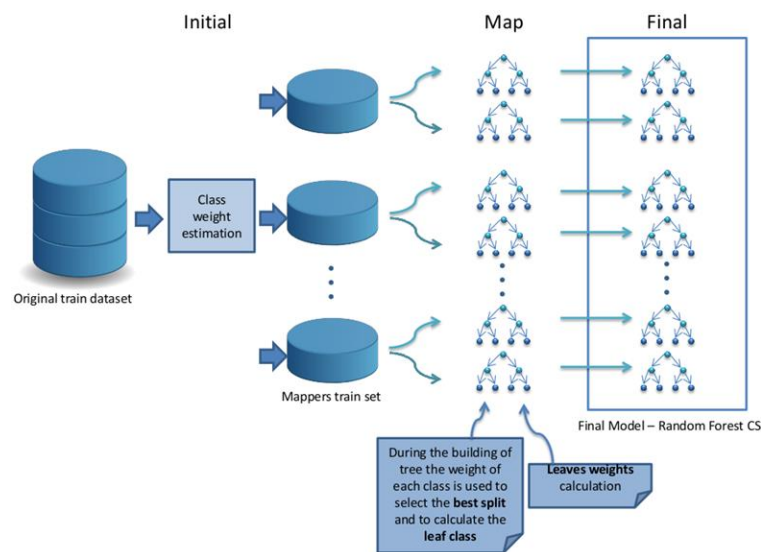


Figure 6. Illustration of random forest approach process stages. Source: Sara del Río et al. 2014.

3.6.4.2. Support Vector Machine

According to (Zhang and Ma 2008, Ostapowicz et al. 2010), SVM is a statistical learning algorithm, which is preferable for both forest classification and for regression modeling. Generally, SVMs focus on the boundary between classes

and map the input space created by independent variables using a non-linear transformation according to a kernel function. Linear, polynomial Radial Basis Function (RBF) and sigmoid are the most commonly used kernel types (Figure 7). The RBF is the most popular kernel, which is used in SVMs (Cortex and Morais 2007, Durbha et al. 2007). According to our literature review, SVM has been used for forest classification (Zhang and Ma 2008, Shafri and Ramle 2009, Ostapowicz et al. 2010). In his work, Sheeren et al. 2016 found SVM, among various non-parametric techniques, to be the best classifier with very close results to other classifiers among them (k -NN and RF).

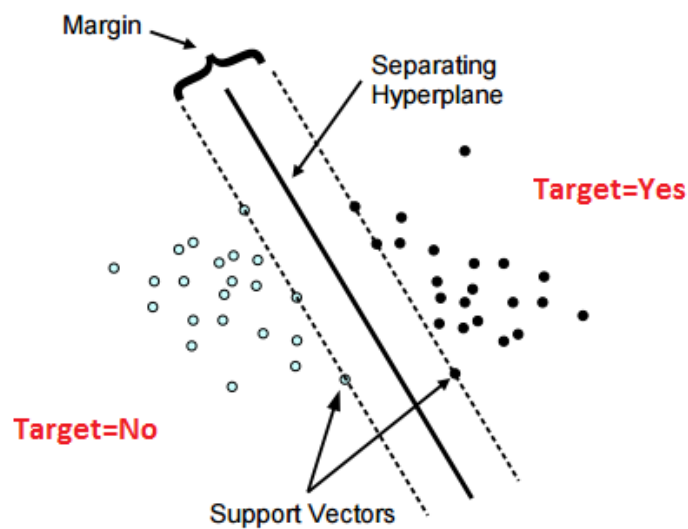


Figure 7. Illustration of SVM methodology for classification of unknown pixels, Source:

<https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf>

3.6.4.3. k -Nearest Neighbor

The k -NN method is one of the simplest and most popular data-mining algorithms used for classification and regression. k -NN is widely used for the estimation of forest description using various topographic and RS data (Franko Lopez et al. 2001, Katila and Tomppo 2001, Ohman and Gregory 2002, Tatjana et al. 2007). In k -NN implementations, three factors should be determined including the number of k , the type of distance measured and weights for nearest neighbors (Figure 8).

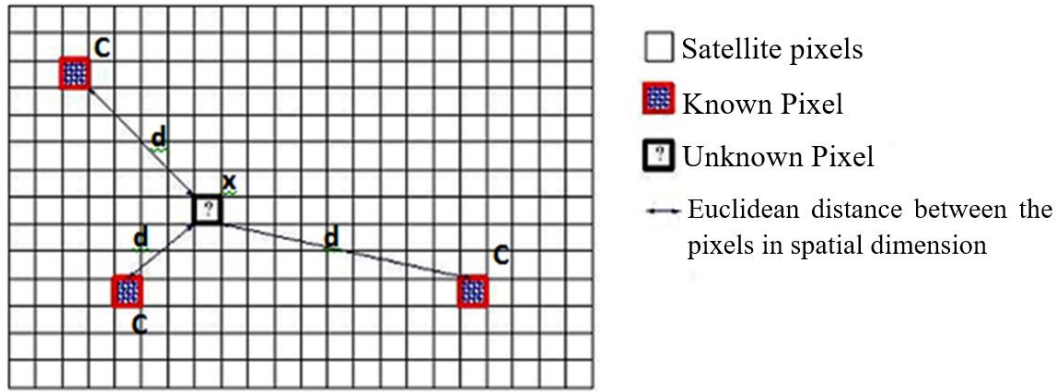


Figure 8. *k*-Nearest Neighbor approach principle.

3.6.5. Assessment of Accuracy

The purpose of accuracy evaluation is measuring the correlation or closeness of estimated map (data) to ground truth data. Availability and level of confidence of results achieve by variety of indicators in different case studies.

3.6.5.1. Regression Accuracy Assessment

The Root Mean Square Error (RMSE), is a commonly used to evaluate the differences between estimated values which derived directly from the mode and the observed values from the environment that is being modelled. These individual differences are also called residuals, and the RMSE basically, serves to aggregate them into a single measure of predictive power.

The RMSE of a model prediction with respect to the estimated variable X_{model} is defined as the square root of the mean squared error (Formula 6):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \quad (6)$$

where X_{obs} is observed values and X_{model} is modelled values at time/place i .

The Mean Absolute Error (MAE) is another common index which can measure the amount of error in estimated value with the same unit of input variable (Hyndman and Koehler 2006). The MAE is given by Formula 7:

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (7)$$

-Chapter 3: Literature review-

As the name suggests, the MAE is an average of the absolute errors, $|e_i| = |f_i - y_i|$ where f_i is the prediction and y_i the true value.

3.6.5.2. Classification Accuracy Assessment

Overall accuracy is one of the most common index for evaluation of the accuracy in case of classification approaches. In calculating this index, just the pixels participate that classified correctly and it ignores the pixels with wrong classified label (Formula 8).

$$\text{Overall accuracy} = \frac{\sum \text{Correct classified pixels}}{\sum \text{Pixels in classification}} \quad (8)$$

Other common indicators described as below (Formula 9, 10).

$$\text{Producer accuracy} = \frac{\sum \text{Correct classified pixels}}{\sum \text{Pixels of X class from ground truth data}} \quad (9)$$

$$\text{Reliability} = \frac{\sum \text{Correct classified pixels}}{\sum \text{Pixels of X class from classification results}} \quad (10)$$

Reliability (also known as user's accuracy) presents the possibility that a pixel of classified map belongs to the same class in reality. In other word, the reliability is the confident level to produced map (Stehman 1997, Powers 2011).

Omission error: Pixels in rows minus the appropriate diagonal cell for the class or group of classes (Formula 11) (Pickles 1995).

$$\text{Omission error} = \frac{\sum \text{non-diagonal elements on each row}}{\text{total pixel on each row}} \quad (11)$$

Commission error: Pixels in columns minus the appropriate diagonal cell for the class or group of classes (Formula 12) (Pickles 1995).

$$\text{Commission error} = \frac{\sum \text{non-diagonal elements on each column}}{\text{total pixel on each column}} \quad (12)$$

Chapter 4

4. Materials & Methods

The used material and methodology for each study area (Figure 9), has been described in details in the section of results. All papers are generally focused on the hypothesis of the potential indirect estimation of forest attributes, using modern approaches of RS for acquisition of forest inventory data. The output of these studies can be used for detail forest management planning and decision-making. Using collected data from both above aerial imagery (UAV and space-born) and below canopy (HP), it was possible to estimate several forest parameters in tree and stand level such as: height-tree positions, DBH, crown diameter, diversity and tree species distribution. The accuracy and overall performance of RS was tested based on the statistical comparison between RS and traditional field methods.

4.1. Study Area Characterization

As it can be seen in figure 9, in two of our papers (II & III) we investigated the forest attributes for stand level in the area called Doctor Bahramnia forestry plan located in north of Iran. The study area is located in the southwest of Gorgan city in Golestan province in Iran (36°43'-36°46'N, 54°21'-54 24'E). The total area is 1,714 ha. The elevation of the study area ranges from 220 to 1,012 m a.s.l. and the slope range is between 0 and 80% and the soil type is brown and grey-brown. The average precipitation is 649 mm. Concerning the aspect, 45% of the total area is facing the west, 42% the north, 10% the east and 3% the south. Concerning the forest type, the majority of the area is covered by *Carpinus-Zelkova*. There were also other types of forest such as: *Zelkova- Carpinus*, *Fagus-Carpinus* and *Fagus*.

In paper IV, two research plots from the Czech Republic were included in this work; Plot 1 is located southeast of Prague (N50.01577, E14.75998) and Plot 2 is located south of the city (N49.81452, E14.700590). Both plots were located on flat terrain and were 625 m² (25 X 25 m) in size. Plot 1 had 48 trees with predominantly Norway spruce (*Picea abies* L.), with a few European larch (*Larix decidua* Mill.) and Scots pine (*Pinus sylvestris* L.) individuals scattered throughout. Plot 2 had 39 trees with mainly Norway

-Chapter 4: Materials and methods-

spruce and Scots pine trees, with scattered individuals of European larch and Silver birch (*Betula pendula* Roth.).

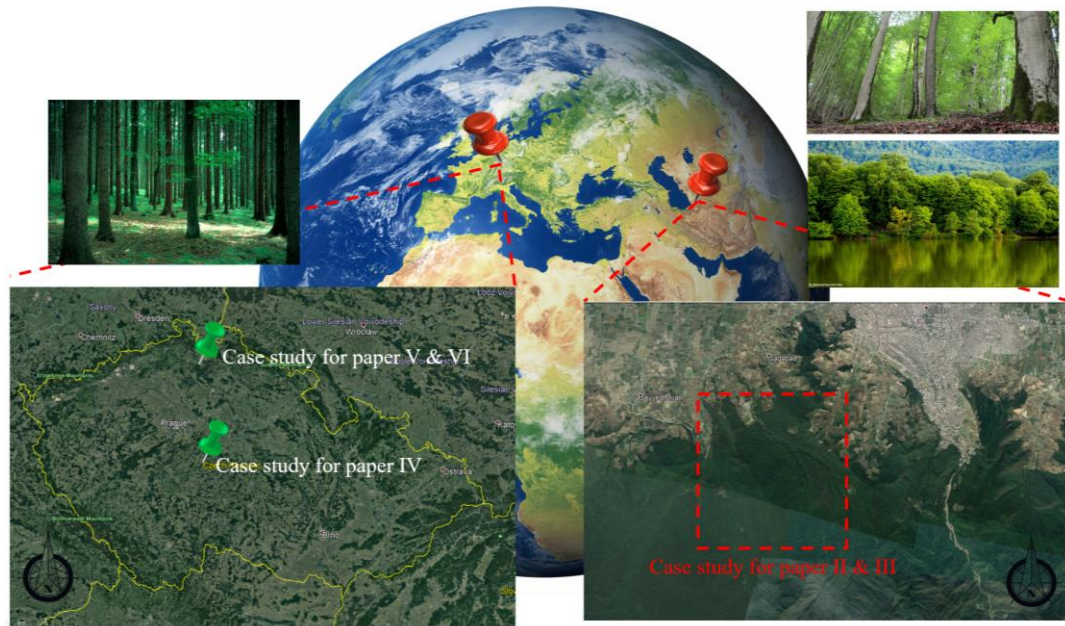


Figure 9. Locations of the study areas in the Czech Republic and Iran, Source: Partially made using Google Earth.

Paper V and VI, the studies were fulfilled on the Doksy territory, lies on the shores of Lake Mácha in Northern Bohemia in the Czech Republic. Geologically, the area is characterized by sandstone pseudokarst in the late stages of development, and the soils are either sandy or peaty, with shallow, peaty basins over rocky sandstone hummocks and sporadic volcanic hills. The areas were located northeast of the city of Doksy (-718000, -991250 NW to -717000, -991950 SE in the local coordinate system S-JTSK/Krovak East North,). The research was carried out in a 140-year-old *Pinus sylvestris* L. (Scots pine) monoculture natural stand established on sandy soils (68%). The vegetative period tends to be rather warm and dry. The mean annual air temperature is 7.3 °C and the average annual maximum temperature is 31.5 °C. The mean annual precipitation is 635 mm, with only 354 mm during the growing season.

Chapter 5

5. Results

Paper I



REVIEW



Forest canopy density assessment using different approaches – Review

AZADEH ABDOLLAHNEJAD*, DIMITRIOS PANAGIOTIDIS, PETER SUROVÝ

*Department of Forest Management, Faculty of Forestry and Wood Sciences,
Czech University of Life Sciences Prague, Prague, Czech Republic*

*Corresponding author: abdollahnejad@fld.czu.cz

Abstract

Abdollahnejad A., Panagiotidis D., Surový P. (2017): Forest canopy density assessment using different approaches – Review. *J. For. Sci.*, 63: 107–116.

Crown canopy is a significant regulator of forest, affecting microclimate, soil conditions and having an undeniable role in a forest ecosystem. Among the different materials and approaches that have been used for the estimation of crown canopy, satellite based methods are among the most successful methods regarding cost-saving efforts and different kinds of options for measuring the crown canopy. Different types of satellite sensors can result in different outputs due to their various spectral and spatial resolution, even when using the same methodologies. The aim of this review is to assess different remote sensing methods for forest crown canopy density assessment.

Keywords: satellite sensors; vegetation index; classification methods; pixel based; spectral analyses

Recently, measuring the crown canopy of a forest has been part of the inventory schedule. The tree crown size determines, among others, carbon sequestration, shading, filtering of fine air particulates, risk of wind-breaking and tree growth. The dependence of the crown size on resource supply, species and tree age complicates an accurate evaluation of the space requirement of a tree, its size-dependent functions and services in forested areas. Two important factors that affect measuring the crown canopy are: definition of crown canopy and the technique used to estimate the crown canopy (KORHONEN et al. 2006).

There are three different methods for measuring or estimating the crown canopy in a forest: (i) ground measurement at the study area (SARVAS 1953; RAUTIAINEN et al. 2005; KORHONEN et al. 2006), (ii) statistical approaches, if the information such as basal area or DBH and number of stems is available, (iii) remote sensing data like aerial photo-

graphs (PITKÄNEN 2001; CULVENOR 2003), satellite data (IVERSON et al. 1989; GEMMELL 1999) or laser scanner data (NÆSSET et al. 2004). From among these the satellite based models are the most common approach for measuring the crown canopy and they can be divided into two main categories.

Remote sensing methods

In these methods, mainly different kinds of algorithms or enhancement functions are applied to a satellite image in order to resolve more clear bands like: soil, atmosphere or vegetation indicators, texture analysis, tasselled cap transformation, etc. Slicing, image arithmetic (BOLES et al. 2004; MATSUSHITA et al. 2007), segmentation and multi-spectral image classification (WANG 1990; SEONG, USERY 2001) are the most common approaches in this category. Although supervised classification is

the most complete one, there are some disadvantages of this approach. Requirement for training area establishment for estimation is one of those impediments. Training area establishment is time consuming, difficult to fulfil and sometimes it cannot give right or enough information.

Biophysical response modelling

The International Tropical Timber Organization developed a new method to solve the problems of remote sensing methods. The advantage of this approach is that it does not need any training samples during the process of so called forest canopy density (FCD) mapping model known as Rikimaru's approach. The FCD mapping model uses the crown canopy density as an important factor for assessing the crown status.

The main purpose of this paper is to review different approaches to estimation and classification of crown canopy density as well as the possibility of remote sensing methods for providing the needed material.

Description of the individual methods

Main bands and analysed spectral data. Remote sensing has been widely used with varying degrees of success to quantify spatially forest structure characteristics such as crown cover, tree density, tree diameter, basal area, tree height, tree age, biomass, and leaf area index. Nowadays using a wide range of software with different enhancement options makes an opportunity for researchers to detect and clarify their interest variables easier. Image contrast enhancement, linear principle component analysis and tasselled cap transformations and texture analysis are some of the common processing methods for image enhancement in remote sensing software such as PCI Geomatica, ERDAS, IDRISI, etc. Pixel based (the most common approach) (SHAO et al. 1996) and object based approaches (SHATAEE 2003) have been used for the classification of crown canopy in related researches.

The development of robust object based methods suitable for medium to high resolution satellite imagery provides a valid alternative to "traditional" pixel based methods (BAATZ et al. 2004; BENZ et al. 2004). The object oriented classification involves segmenting an image into objects (groups of pixels). There are two main methods for object based approaches:

- (1) Direct method such as (i) region growing technique – can be employed to a group of adjacent pixels with similar spectral values into individual objects (GAO et al. 2006), (ii) edge detection technique – can be used to identify discontinuities (object boundaries or edges) throughout the image, these boundaries can be used to build polygons for the object based classification (CARLEER, WOLFF et al. 2006);
- (2) Indirect method: here, the imagery is supplemented with another spatial data, often digital vector map data. The objects characterised by the vector polygons are assigned land cover values derived from the imagery.

Also, some researchers (TUCKER 1979; HUANG et al. 2001) used the main bands of satellites (Landsat 7 ETM+) for the classification of crown canopy.

Moreover, there are different image classification procedures used for different purposes by various researchers (ERNST, HOFFER 1979; TUCKER 1979; BUTERA 1983; LO, WATSON 1998; LIU et al. 2002; OZESMI, BAUER 2002; DEAN, SMITH 2003; PAL, MATHER 2003).

These techniques are distinguished in two main ways: (i) unsupervised classification which requires no training data (GHAZANFARI 1996; SHIRIAN 1997; MIRAKHORLO 2003; HOSSEINI et al. 2004), (ii) supervised classification including maximum likelihood (BUTERA 1983; LEE, PARK 1992; Yi et al. 1994; CLARK et al. 2005), minimum distance to mean (HUGUENIN et al. 1997), Mahalanobis, Fisher classifier, parallelepiped (HINES et al. 1993) and Bayesian formulation based classification (SCHISTAD-SOLBERG et al. 1994) (Fig. 1). These algorithms were used by different studies such as RAMTIN NIA (1997), ABBASI (2001), and HOSSEINI et al. (2004). Recently, non-parametric algorithms are widely used for the classification of forest attributes like crown canopy. Some of those data mining classifier algorithms have been commonly used in recent studies such as:

- (i) *k*-Nearest neighbour (*k*-NN) method is one of the simplest and most popular data-mining algorithms used for classification and regression. *k*-NN is widely used for the estimation of forest description using various topographic and remote-sensing data (BREIMAN 2001, 2002). In *k*-NN implementations, three factors should be determined including the number of *k*, the type of distance measure and weights for the nearest neighbours;
- (ii) Support vector machine classification: this algorithm is suitable for both classification and regression techniques based on the statistical

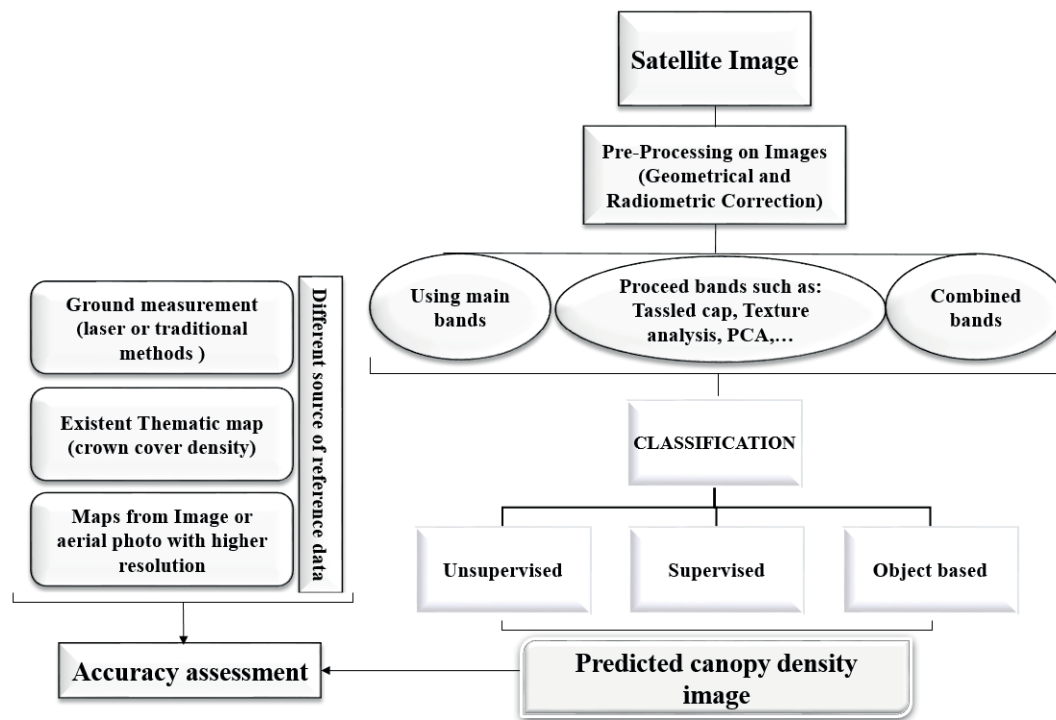


Fig. 1. The main stages of image classification in different methods using different reference data
PCA – principal component analysis

learning theory (WALTON 2008). Generally, support vector machine (SVM) focuses on the boundary between classes and maps the input space created by independent variables using a non-linear transformation according to a kernel function. Linear, polynomial radial basis function (RBF) and sigmoid are the most commonly used kernel types. The RBF is the most popular kernel, which is used in SVMs (CORTEZ, MORAIS 2007; DURBHA et al. 2007). According to our literature review, SVM has been used for forest classification (ZHANG, MA et al. 2008; OSTAPOWICZ et al. 2010). In his work, SHEEREN et al. (2016) found out the SVM among various non-parametric techniques to be the best classifier with very close results to other classifiers among them (*k*-NN and random forest);

(iii) Random forest (RF) is a new algorithm to the field of data mining, designed to produce accurate predictions that do not overfit the data (BREIMAN 2002). RF can also be used for regression-type problems (to predict a continuous dependent variable) and classification problems (to predict a categorical dependent variable). Implementation of RF depends on the regularization of decision tree and stopping parameters. The decision tree model parameters include the maximum number of trees that must be grown in the forest and the number of variables (*k* predictor or independent

variables in each node for predicting dependent values) that are randomly selected in each node. Alternatively, choosing a small number of predictor variables may downgrade the prediction performance, because this can exclude variables that may account for most of the variability and trends in the data (StatSoft, Inc. 2010). The stopping parameters or control parameters are used to stop running the algorithm when satisfactory results have been achieved (SHATAEE et al. 2012). In some studies such as GARZÓN et al. (2008) and SHATAEE et al. (2012), the RF has been used for the prediction of forest attributes.

Vegetation indicators. Vegetation indicators are based on the spectral reflection of plants (red and near infra-red range). There are three categories for vegetation indicators (Table 1):

(i) Mean vegetation indices: almost most of these indicators have been used for measuring the frequency of plants and biological characteristics of crown cover. These indicators just use red and infra-red bands. Normalized difference vegetation index (NDVI) (ROUSE et al. 1973), ratio vegetation index (also known as the simple ratio) (BIRTH, McVEY 1968), green normalized difference vegetation index (BUSCHMANN, NAGEL 1993) and green difference vegetation index (SRIPADA et al. 2006) are some of the most common indicators in this category;

Table 1. Categories of vegetation indices

Vegetation index category	Indices	Explanation
Mean	NDVI	$(\text{NIR} - \text{red})/(\text{NIR} + \text{red})$
	RVI	NIR/red
	GDVI	$\text{NIR} - \text{green}$
	GNDVI	$(\text{NIR} - \text{green})/(\text{NIR} + \text{green})$
	GEMI	$\dot{\eta} \times (1 - 0.25 \times \dot{\eta}) - [(\text{red} - 0.125)/(1 - \text{red})]$
Atmospherically resilient	GARI	$\text{NIR} - [\text{green} - (\text{blue} - \text{red})]/\text{NIR} \times [\text{green} - (\text{blue} - \text{red})]$
	VIg	$(\text{green} - \text{red})/(\text{green} + \text{red})$
	VARIg	$(\text{green} - \text{red})/(\text{green} + \text{red} - \text{blue})$
	SAVI	$[(\text{NIR} - \text{red})/(\text{NIR} + \text{red} + L)] \times (1 + L)$
Soil-adjusted	MSAVI2	$[2 \times \text{NIR} + 1 - \sqrt{(2 \times \text{NIR} + 1)^2 - 8 \times (\text{NIR} - \text{red})}]/2$
	GSAVI	$[(\text{NIR} - \text{green})/(\text{NIR} + \text{green} + L)] \times (1 + L)$
	TSAVI	$[a \times (\text{NIR} - a \times \text{red} - b)]/[a \times \text{NIR} + \text{red} - (a \times b) + X \times (1 + a^2)]$

NDVI – normalized difference vegetation index, RVI – ratio vegetation index, GDVI – green difference vegetation index, GNDVI – green normalized difference vegetation index, GEMI – global environmental monitoring index, GARI – green atmospherically resilient index, VIg – vegetation index green, VARIg – vegetation atmospherically resilient index green, SAVI – soil-adjusted vegetation index, MSAVI2 – modified soil-adjusted vegetation spectral index, GSAVI – green soil-adjusted vegetation index, TSAVI – transformed soil-adjusted vegetation index, NIR – near infra-red band, $\dot{\eta} = (2 \times (\text{NIR}^2 - \text{red}^2) + 1.5 \times \text{NIR} + 0.5 \times \text{red})/(\text{NIR} + \text{red} + 0.5)$, $L = 0.5$, a – the slope of the soil line, b – the intercept of the soil line, X – adjustment factor

(ii) Atmospherically resilient vegetation indices: these indicators use the blue or green bands besides the red and infra-red bands in order to solve the dependence of vegetation indices on atmospheric effects. Global environmental monitoring index (PINTY, VERSTRAETE 1992), green atmospherically resilient index (GITELSON et al. 2002), vegetation index green (GITELSON et al. 2002), and vegetation atmospherically resilient index green (GITELSON et al. 2002) are some of these indices;

(iii) Soil-adjusted vegetation indices: by using one parameter called L these indicators try to decrease the soil effect on NDVI index. The L factor is an adjustment parameter, the amount of this factor for the area with low density 1, for the area with intermediate density 0.5 (the most common amount) and for the area with high crown cover 0 is considered. Soil-adjusted vegetation index (HUETE 1988), modified soil-adjusted vegetation spectral index (MSAVI2) as it is described in Eq. 19 in QI et al. (1994), transformed soil-adjusted vegetation index, the median soil line values of which reported in BARET and GUYOT (1991) are $a = 1.2$ and $b = 0.04$, and green soil-adjusted vegetation index (SRIPADA et al. 2006) are some indicators from this category.

Different studies arrived at different indices as the best vegetation index based on the density of their case study. Sensitivity of indicators to the amount of crown canopy and soil or gap area percentage

can cause different results in different studies. For example NDVI and MSAVI2 in the area with lower or without crown canopy density have higher accuracy (ABDI et al. 2009).

Biophysical response modelling (FCD model). FCD is used as an important variable for the characteristics of forest status. FCD is based on the growth component and it can illustrate the degree of degradation (RIKIMARU et al. 1999). FCD model shows the growth phenomena of forests, which is quantitative analysis. The degree of forest density is expressed in percentages: i.e. 10, 20, 30, 40% and so on.

Also, this model makes it possible to monitor changes in the forest crown canopy density over time. This method also makes it possible to monitor the transformation of forest conditions over time and it can assess the progress of reforestation activities.

Based on four different indicators, FCD can calculate the percentage of canopy density for each pixel. These indices are: (i) advanced vegetation index (AVI), (ii) bare soil index (BI), (iii) canopy shadow index or scaled shadow index (SSI), (iv) thermal index (TI) (RIKIMARU et al. 2002; GODINHO et al. 2016) (Table 2). The principle of this method is shown in Fig. 2. There are direct and indirect relationships between forest canopy and FCD components (Table 3).

In this method like in other approaches, first of all the geometric accuracy and spectral quality of

Table 2. Indicators of forest canopy density model

	Formula	Explanation and practical uses
Advanced vegetation index (AVI)	$AVI = [(B4 + 1)(256 - B3)(B4 - B3)]^{\frac{1}{3}}$	The use of power degree on NDVI enables AVI to be more sensitive to forest density and physiognomic vegetation indices.
Bare soil index (BI)	$BI = \frac{(B4 + B2) - B3}{(B4 + B2) + B3}$	It is the index prepared for analysing soils, in other words it can be used to identify the difference between agricultural and non-agricultural vegetation.
Canopy shadow index (SI)	$SI = \sqrt{(256 - B2)(256 - B3)}$	Evaluates the different shadow patterns, based on the structure, age, species distribution etc., by affecting the spectral responses each time.
Thermal index	-	Source of info is the thermal band of thematic mapper sensor (band 6).

B2 – green band, B3 – red band, B4 – near infra-red band, NDVI – normalized difference vegetation index

images must be checked. Then all the bands (except the thermal band) must be normalized by using Eqs 1 and 2:

$$A = \frac{(Y1 - Y2)}{(X1 - X2)} = \frac{(Y1 - Y2)}{(M - 2S) - (M + 2S)} \quad (1)$$

where:

- A – linear transformation,
- Y1 – maximum value of standardized value,
- Y2 – minimum value of standardized value,
- X1 = M - 2S,
- X2 = M + 2S,
- M – mean of values,
- S – standard deviation.

$$Y = AX + (-AX1 + Y1) \quad (2)$$

where:

Y – normalized data,

X – original value.

According to many researchers who used this method, alone or in comparison with other approaches, the FCD model can be a feasible and accurate approach to the estimation of forest crown canopy density (Table 4).

Detection and separation of bare soil reflectance and plant reflection in a high density forest (more than 70%), no need of training samples during the crown cover density classification are the advantages of FCD model. Also, according to BONYAD (2005), TAEFI (2006), MOEINAZAD TEHRANI et al. (2008), and PAKKHESAL and BONYAD (2013) the low accuracy of classification in a low and middle density forest can be the disadvantage of FCD model.

It is worth mentioning that the elimination of individual pixels after classification by using a low pass filter (3 × 3 or 5 × 5) can increase the accuracy of classification – average increment 5% (BONYAD 2005; PAKKHESAL and BONYAD 2013).

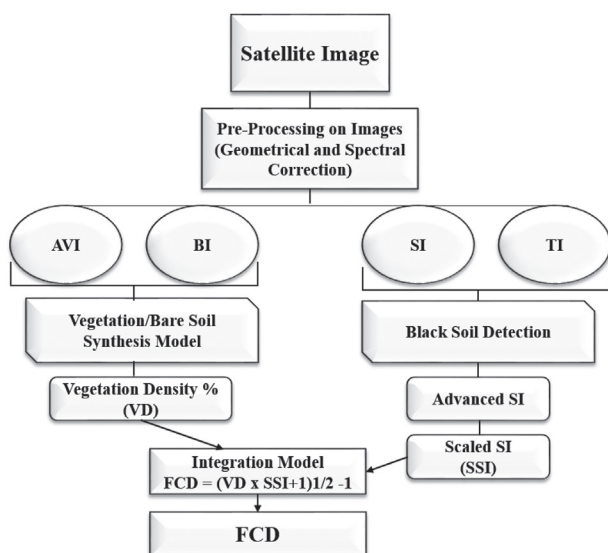


Fig. 2. Methodological flow chart demonstrating different stages of the forest canopy density (FCD) model

AVI – advanced vegetation index, BI – bare soil index, SSI – scaled canopy shadow index, SI – canopy shadow index, TI – thermal index

Table 3. Relationship between forest canopy and forest canopy density (FCD) parameters

Index	High FCD	Low FCD	Grassland	Bare land
AVI	high	mid	high	low
BI	low	low	low	high
SI	high	mid	low	low
TI	low	mid	mid	high

AVI – advanced vegetation index, BI – bare soil index, SI – canopy shadow index, TI – thermal index

Table 4. Some researches on the forest canopy density model using different images

Satellite/sensors	References	Resolution (m)
Landsat (TM)	RIKIMARU et al. (2002)	
Landsat (TM, ETM+)	JAMALABAD and AKBAR (2004)	
Landsat (TM)	NANDY et al. (2003)	30
Landsat (GeoCover)	HADI et al. (2004)	
Landsat 7 ETM+	JOSHI et al. (2006)	
IRS (LISS-III)	AZIZI et al. (2008)	25
SPOT, ALI	MAHBOOB and IGBAL (2012)	
Landsat 7 (ETM+)	PAKKHESAL and BONYAD (2013)	
Landsat 7 ETM+	SHAHVALI-KOUHSHOUR et al. (2012)	30
Landsat (TM)	DEKA et al. (2013)	
Landsat (TM)	BANERJEE et al. (2014)	
Landsat 5 (TM) – visible and NIR band	GODINHO et al. (2016)	
Landsat 5 (TM) – thermal band	GODINHO et al. (2016)	120

TM – thematic mapper, ETM+ – enhanced thematic mapper plus, IRS – Indian remote sensing satellite, LISS-III – linear imaging self-scanning sensor 3, SPOT – Earth observation satellite, ALI – advanced land imager, NIR – near infra-red

DISCUSSION

Main categories of classification

In this study we tried to show the different methods used for crown canopy assessment. All these methods can be divided into two main categories: pixel and object based approaches. The principle of pixel based approaches is based on spectral data derived from pixel cells. As an alternative to the essentially pixel based analysis, the object based method attempts to identify groups of pixels that form discrete objects on the basis of characters that might include overall shape or texture as well as their spectral similarity. Object based methods avoid the need for the complete classification of the whole image, where one has specific interest in one component.

In fact, object oriented methods use segments that are regions specified by one or more yardsticks of homogeneity in one or more dimensions (HAY, CASTILLA 2008). Using different dimensions like spatial dimensions (distances, neighbourhood, topologies, etc.) is crucial to object oriented methods, making them the most popular methods in recent times, as compared to the usage of pixel based methods (CONCHEDDA et al. 2008).

Biophysical response modelling (FCD method)

Like all models, the FCD model has some advantages and some disadvantages. It is disability to achieve high accuracy in very dense areas and

having to use pixel based principles are the biggest disadvantages of this method. On the other hand, modelling with high accuracy and no need of training area establishment (i.e. ground truth) are the most important criteria of FCD that can reduce the time and cost of modelling. So far, many studies have used this approach and almost all of them had acceptable results.

JAMALABAD and AKBAR (2004) used three sets of thematic mapper and enhanced thematic mapper plus (ETM+) of 1991, 1998 and 2002. FCD from each data set was classified into 5 classes (class 1 – water and clouds, class 2 – no forest, class 3 – low forest, 5–40%, class 4 – middle forest, 41–70%, class 5 – dense forest, 71–100%). The result that came from using ETM+ 2002 was with overall accuracy of 83% and Kappa coefficient 0.78. Finally they used the FCD results of images from the same season in 1991 and 1998 in order to prepare the change detection map for their study area.

AZIZI et al. (2008) tested the FCD model using a geometrically corrected image coming from Indian remote sensing satellite (IRS) imagery 2007 of an old growth forest of the north forest division of Iran. The overall accuracy of the IRS images was 84.4% and the Kappa coefficient was 0.783.

After geometric correction (RMSE = 0.5 pixel) of the images and spectral range normalization of the first five bands of Landsat 7, SHAHVALI-KOUHSHOUR et al. (2012) calculated the four main indicators of FCD using these indicators. They were able to create an advanced shadow index and vegetation density index maps. They used different numbers of classes (3, 4 and 6) to classify the FCD result.

Because their study area was highly heterogeneous and dense, the results showed that the class 3 layer had the highest overall accuracy (62%) and Kappa coefficient 0.30.

MAHBOOB and IQBAL (2012) used an Earth observation SPOT 5 satellite (2.5 m) and advanced land imager ALI (30 m) for forest crown closure to calculate the forest crown canopy density in Ayubia National Park, Pakistan. A diverse variety of tree species like coniferous and broadleaved tree species are present in this natural environment. Results showed that the crown canopy of the study area using SPOT imagery was between 20 and 65%, and 45–65% with reference to ALI imagery. It was also concluded that SPOT imagery gave better results because of the higher spatial resolution compared to ALI imagery. On the other hand, SPOT was unable to detect the built up and landslide areas and gave them high values, whereas ALI imagery, having a higher spectral resolution compared to SPOT data, was able to detect these areas and give them low values.

PAKKHESAL and BONYAD (2013) conducted their study on Landsat ETM+ images in order to classify the crown canopy classes in the Shafarud Area of Guilan. First they prepared a forest density map which included different density classes (bare, 5–25, 25–50, 52–75 and 75–100%). They used the four different indicators of FCD (AVI, BI, SSI, TI) in order to calculate the percentage of canopy density for each pixel. One thematic map came from an orthomosaic aerial photo and was used for the evaluation of FCD accuracy. Results of maximum likelihood classification showed that the FCD map results were close to ground reality (overall accuracy 71% and Kappa coefficient 0.61). Also, the matrix of errors showed that the FCD method is not applicable to areas with high and medium density but it can provide high accuracy for dominant trees and area with low density (lower than 5%).

In overall, the spatial and spectral resolution of images acquired by different sensors brings several advantages to natural resource managers and academic researchers for the classification, monitoring, and management of natural ecosystems. Some of the basic requirements for successful remote sensing-based monitoring can be listed briefly as follows: (i) availability of required digital data sources (i.e., imagery, maps, and any other forms of data), (ii) collecting up-to-date from the interest area, (iii) applying the image processing methodology with respect to the specific characteristics of the study area, (iv) producing accurate and useful outputs (e.g., maps and statistics).

Selecting an appropriate method and material with respect to the study area status, aim of study and commercial acceptance can help the managers and researchers to access easier to their purposes. In case of crown canopy classification, the crown density of the study area and the aim of the study have a significant role which can influence the total accuracy of canopy modelling.

References

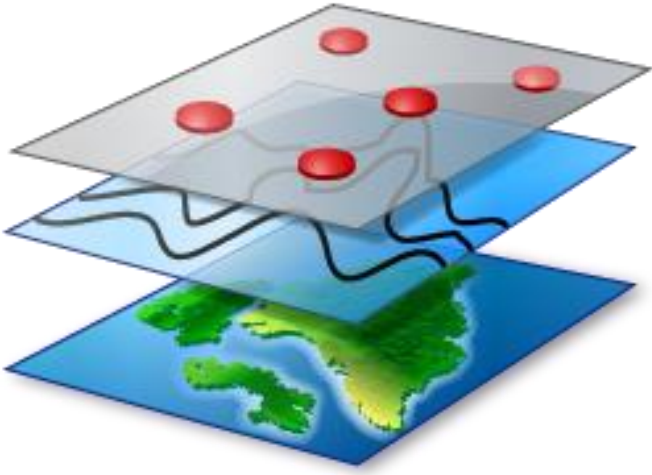
- Abbasi M. (2001): Investigating the possibility of *Fagus orientalis* stand type mapping using ETM+ data in Chalir district – Nowshahr. [MSc Thesis.] Tehran, University of Tehran: 114.
- Abdi O., Akbari H., Sosani J., Shirvani Z. (2009): Comparison of vegetation indicators for determination of crown canopy of Zagros forest using ETM+ data. In: Abbasi A. (ed.): Geomatic Conference, Tehran, May 10–11, 2009: 12.
- Azizi Z., Najafi A., Sohrabi H. (2008): Forest canopy density estimating, using satellite images. In: Chen J., Jiang J., Peled A. (eds): Proceedings of the 21st ISPRS Congress, Commission VIII, Beijing, July 3–11, 2008: 1127–1130.
- Baatz M., Benz U., Deghani S., Heynen M., Höltje A., Hofmann P., Lingenfelder I., Mimler M., Sohlbach M., Weber M., Willhauck G. (2004): Definiens Enterprise Image Intelligence Suite 7 for Windows – installation and administration guide. Available at <http://www.ecognition.com/download/WindowsInstallationGuide.pdf> (accessed Mar 7, 2006).
- Banerjee K., Panda S., Bandyopadhyay J., Jain M.K. (2014): Forest canopy density mapping using advance geospatial technique. International Journal of Innovative Science, Engineering & Technology, 1: 358–363.
- Baret F., Guyot G. (1991): Potentials and limits of vegetation indices for LAI and PAR assessment. Remote Sensing of Environment, 35: 161–173.
- Benz U., Hofmann P., Willhauck G., Lingenfelder I., Heynen M. (2004): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 58: 239–258.
- Birth G.S., McVey G.R. (1968): Measuring colour of growing turf with a reflectance spectrophotometer. Agronomy Journal, 60: 640–649.
- Boles S.H., Xiao X.M., Liu J.Y., Zhang Q.Y., Munkhtuya S., Chen S.Q., Ojima D. (2004): Land cover characterization of Temperate East Asia using multi-temporal vegetation sensor data. Remote Sensing of Environment, 90: 477–489.
- Bonyad A. (2005): Multitemporal satellite image database classification for land cover inventory and mapping. Journal of Applied Sciences, 5: 835–837.
- Breiman L. (2001): Random forests. Machine Learning, 45: 5–32.

- Breiman L. (2002): Using models to infer mechanisms. IMS Wald lecture 2. Available at <https://oz.berkeley.edu/users/breiman/wald2002-2.pdf> (accessed Sept 19, 2002).
- Buschmann C., Nagel E. (1993): *In vivo* spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing*, 14: 711–722.
- Butera M.K. (1983): Remote sensing of wetlands. *IEEE Transactions on Geosciences and Remote Sensing*, 21: 383–392.
- Carleer A.P., Wolff E. (2006): Region-based classification potential for land-cover classification with very high spatial resolution satellite data. In: Lang S., Blaschke T., Schöpfer E. (eds): *Proceedings of the 1st International Conference on Object-based Image Analysis*, Salzburg, July 4–5, 2006: 1–6.
- Clark M.L., Roberts D.A., Clark D.B. (2005): Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment*, 96: 375–398.
- Conchedda G., Durieux L., Mayaux P. (2008): An object-based method for mapping and change analysis in mangrove ecosystems. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63: 578–589.
- Cortez P., Morais A. (2007): A data mining approach to predict forest fires using meteorological data. In: Neves J., Santos M.F., Machado J.M. (eds): *Proceedings of the EPIA 2007 – Portuguese Conference on Artificial Intelligence*, Guimarães, Dec 3–7, 2007: 512–523.
- Culvenor D.S. (2003): Extracting individual tree information. A survey of techniques for high spatial resolution imagery. In: Wulder M.A., Franklin S.E. (eds): *Remote Sensing of Forest Environments: Concepts and Case Studies*. Boston, Kluwer Academic Publishers: 255–277.
- Dean A.M., Smith G.M. (2003): An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. *International Journal of Remote Sensing*, 24: 2905–2920.
- Deka J., Tripathi O.P., Khan M.L. (2013): Implementation of forest canopy density model to monitor tropical deforestation. *Journal of the Indian Society of Remote Sensing*, 41: 469–475.
- Durbha S.S., King R.L., Younan N.H. (2007): Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment*, 107: 348–361.
- Ernst C.L., Hoffer R.M. (1979): *Digital Processing of Remotely Sensed Data for Mapping Wetland Communities*. LARS Technical Report 122079. West Lafayette, Purdue University: 119.
- Gao Y., Mas J.F., Maathuis B.H.P., Zhang X.M., Van Dijk P.M. (2006): Comparison of pixel-based and object oriented image classification approaches – a case study in a coal fire area, Wuda, Inner Mongolia, China. *International Journal of Remote Sensing*, 27: 4039–4051.
- Garzón M.B., Dios R.S., Ollero H.S. (2008): Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, 11: 169–178.
- Gemmell F. (1999): Estimating conifer forest cover with thematic mapper data using reflectance model inversion and two spectral vegetation indices in a site with variable background characteristics. *Remote Sensing of Environment*, 69: 105–121.
- Ghazanfari H. (1996): Investigation of the application of satellite data for classifying forest types in the forests managed by Mazandaran wood and paper company. [MSc Thesis.] Gorgan, University of Gorgan: 124.
- Gitelson A.A., Kaufman Y.J., Stark R., Rundquist D. (2002): Novel algorithms for remote estimation of vegetative fraction. *Remote Sensing of Environment*, 80: 76–87.
- Godinho S., Guiomar N., Machado R., Santos P., Sá-Sousa P., Fernandes J.P., Neves N., Pinto-Correia T. (2016): Assessment of environment, land management, and spatial variables on recent changes in montado land cover in southern Portugal. *Agroforestry Systems*, 90: 177–192.
- Hadi F., Wikantika K., Sumarto I. (2004): Implementation of forest canopy density model to monitor forest fragmentation in Mt. Simpang and Mt. Tilu Nature Reserves, West Java, Indonesia. In: Sumarto I. (ed.): *Proceedings of the 3rd FIG Regional Conference*, Jakarta, Oct 3–7, 2004: 9–20.
- Hay G.L., Castilla G. (2008): Geographic object-based image analysis (GEOBIA): A new name for a new discipline. In: Blaschke T., Lang S., Hay G.J. (eds): *Object-Based Image Analysis – Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Berlin, Heidelberg, Springer-Verlag: 93–112.
- Hines M.E., Pelletier R.E., Crill P.M. (1993): Emission of sulfur gases from marine and freshwater wetlands of the Florida Everglades: Rates and extrapolation using remote sensing. *Journal of Geophysical Research*, 98: 8991–8999.
- Hosseini S.Z., Khajeddin S.J., Azarnivand H., Khalilpour S.A. (2004): Land use mapping using ETM+ data (case study: Chamestan area, Iran). In: Altan O. (ed.): *Proceedings of the 20th ISPRS Congress, Commission VII*, Istanbul, July 12–23, 2004: 391–393.
- Huang C., Yang L., Wylie B., Homer C. (2001): A strategy for estimating tree canopy density using Landsat 7 ETM+ and high resolution images over large areas. Available at http://landcover.usgs.gov/pdf/canopy_density.pdf (accessed Mar 1, 2008).
- Huete A.R. (1988): A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25: 295–309.
- Huguenin R.L., Karaska M.A., Blaricom D.V., Jensen J.R. (1997): Subpixel classification of bald cypress and tupelo gum trees in Thematic Mapper imagery. *Photogrammetric Engineering & Remote Sensing*, 63: 717–725.
- Iverson L.R., Cook E.A., Graham R.L. (1989): A technique for extrapolating and validating forest cover across large

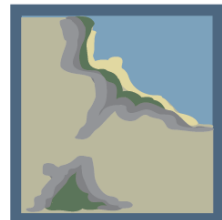
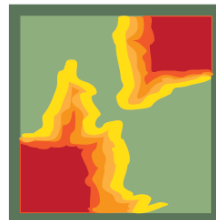
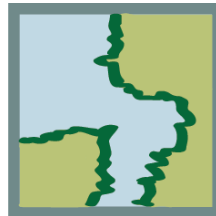
- regions: Calibrating AVHRR data with TM data. *International Journal of Remote Sensing*, 10: 1805–1812.
- Jamalabad M.S., Akbar A.A. (2004): Forest canopy density monitoring, using satellite images. In: Altan O. (ed.): *Proceedings of the 20th ISPRS Congress, Commission VII, Istanbul, July 12–23, 2004*: 244–249.
- Joshi C., Leeuw J.D., Skidmore A.K., Duren I.C., Van Oosten H. (2006): Remotely sensed estimation of forest canopy density: A comparison of the performance of four methods. *International Journal of Applied Earth Observation and Geoinformation*, 8: 84–95.
- Korhonen L., Korhonen K.T., Rautiainen M., Stenberg P. (2006): Estimation of forest canopy cover: A comparison of field measurement techniques. *Silva Fennica*, 40: 577–587.
- Lee J.K., Park R.A. (1992): Application of geoprocessing and simulation modelling to estimate impacts of sea level rise on the northeast coast of Florida. *Photogrammetric Engineering & Remote Sensing*, 58: 1579–1586.
- Liu X.H., Skidmore A.K., Oosten V.H. (2002): Integration of classification methods for improvement of land-cover map accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56: 257–268.
- Lo C.P., Watson L.J. (1998): The influence of geographic sampling methods on vegetation map accuracy evaluation in a swampy environment. *Photogrammetric Engineering & Remote Sensing*, 64: 1189–1200.
- Mahboob J., Iqbal F. (2012): Forest crown closure assessment using multispectral satellite imagery. *African Journal of Agricultural Research*, 7: 5033–5042.
- Matsushita B., Yang W., Chen J., Onda Y., Qiu G. (2007): Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: A case study in high-density cypress. *Forest Sensors*, 7: 2636–2651.
- Mirakhorlo K.H. (2003): Land cover/land use mapping of the Northern forests of Iran using Landsat ETM+ data. *Iranian Journal of Forest and Poplar Research*, 1: 1–11. (in Persian)
- Moeinazad Tehrani S.M., Darvishsefat A.A., Namiraniyan M. (2008): Evaluation of FCD model for estimation of forest density using Landsat 7 imagery (case study: Chalus Forest). *Iranian Journal of Forest and Poplar Research*, 16: 124–138. (in Persian)
- Næsset E., Gobakken T., Holmgren J., Hyyppä H., Hyyppä J., Maltamo M., Nilsson M., Olsson H., Persson A., Söderman U. (2004): Laser scanning of forest resources: The Nordic experience. *Scandinavian Journal of Forest Research*, 19: 482–499.
- Nandy S., Joshi P., Das K. (2003): Forest canopy density stratification using biophysical modelling. *Journal of the Indian Society of Remote Sensing*, 31: 291–297.
- Ostapowicz K., Lakes T., Kozak J. (2010): Modelling of land cover change using support vector machine. In: Painho M., Santos M.Y., Pundt H. (eds): *Proceedings of the 13th AGILE International Conference on Geographic Information Science, Guimarães, May 10–14, 2010*.
- Ozesmi S.L., Bauer M. (2002): Satellite remote sensing of wetlands. *Wetlands Ecology and Management*, 10: 381–402.
- Pakkhesal E., Bonyad A.E. (2013): Classification and delineating natural forest canopy density using FCD Model (case study: Shafarud area of Guilan). *Iranian Journal of Forest and Poplar Research*, 21: 99–114. (in Persian)
- Pal M., Mather P.M. (2003): An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment*, 86: 554–565.
- Pinty B., Verstraete M.M. (1992): GEMI: A non-linear index to monitor global vegetation from satellites. *Vegetation*, 101: 15–20.
- Pitkänen J. (2001): Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. *Canadian Journal of Forest Research*, 31: 832–844.
- Qi J., Chehbouni A., Huete A.R., Kerr Y.H. (1994): A modified soil-adjusted vegetation index. *Remote Sensing of Environment*, 48: 119–126.
- Ramtin Nia K. (1997): Preparing forest cover type map using digital satellite information in Kheiroudkenar Forest – Nowshahr. [MSc Thesis.] Tehran, University of Tehran: 96. (in Persian)
- Rautiainen M., Stenberg P., Nilson T. (2005): Estimating canopy cover in Scots pine stands. *Silva Fennica*, 39: 137–142.
- Rikimaru A., Roy P., Miyatake S. (2002): Tropical forest cover density mapping. *Tropical Ecology*, 43: 39–47.
- Rikimaru A., Utsuki Y., Yamashita S. (1999): Basic study of the maximum logging volume estimation for consideration of forest resources using time series FCD (forest canopy density) model. In: Harris R. (ed.): *Asian Conference on Remote Sensing, Manila, Nov 16–20, 1998*: 6.
- Rouse J.W., Haas R.H., Schell J.A., Deering D.W. (1973): Monitoring vegetation systems in the Great Plains with ERTS. In: Fraden S.C., Marcanti E.P., Becker M.A. (eds): *3rd ERTS-1 Symposium, NASA SP-351, Washington, D.C., Dec 10–14, 1973*: 309–317.
- Sarvas R. (1953): Measurement of the crown closure of the stand. *Communicationes Instituti Forestalis Fenniae*, 41: 1–13.
- Schistad-Solberg A.H., Jain A.K., Taxt T. (1994): Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 32: 768–778.
- Seong J.C., Usery E.L. (2001): Assessing raster representation accuracy using a scale factor model. *Photogrammetric Engineering & Remote Sensing*, 67: 1185–1191.
- Shahvali-Kouhshour A., Pir-Bavaghar M., Fathi P. (2012): Forest cover density mapping in sparse and semi dense forests using forest canopy density model (case study: Marivan forests). *Journal of RS and GIS for Natural Resources*

- (Journal of Applied RS and GIS Techniques in Natural Resource Science), 3: 373–383.
- Shao G., Zhao S., Shugart H., Wang S., Schaller J. (1996): Forest cover types derived from Landsat thematic mapper imagery for Changbai Mountain area of China. *Canadian Journal of Forest Research*, 26: 206–216.
- Shataee S.H. (2003): Investigation of the possibility of forest type mapping using satellite information (a case study of Kheiroudkenar forest – Nowshahr). [Ph.D. Thesis.] Tehran, University of Tehran: 155.
- Shataee S.H., Kalbi S., Fallah A., Pelz D. (2012): Forest attribute imputation using machine-learning methods and ASTER data: Comparison of k -NN, SVR and RF regression algorithms. *International Journal of Remote Sensing*, 19: 6254–6280.
- Sheeren D., Fauvel M., Josipović V., Lopes M., Planque C., Willm J., Dejoux J.F. (2016): Tree species classification in temperate forests using Formosat-2 satellite image time series. *Journal of Remote Sensing*, 8: 734.
- Shirian R. (1997): Vegetation mapping of Golestan National Park using GIS and Landsat TM data. [MSc Thesis.] Gorgan, University of Gorgan: 99.
- Sripada R.P., Heiniger R.W., White J.G., Meijer A.D. (2006): Aerial colour infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal*, 98: 968–977.
- StatSoft, Inc. (2010): Statistica. Electronic textbook. Available at <http://www.Statsoft.com> (accessed Nov 20, 2010).
- Taefi M. (2006): Evaluation and Optimization of FCD Model in Order to Estimate Forest Canopy Density Using Merger Data Method and Image Index. Tehran, Nasir al-Din Tusi University: 95.
- Tucker C.J. (1979): Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8: 127–150.
- Walton J.T. (2008): Sub pixel urban land cover estimation: Comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering & Remote Sensing*, 74: 1213–1222.
- Wang F. (1990): Fuzzy supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 28: 194–201.
- Yi G.C., Risley D., Koneff M., Davis C. (1994): Development of Ohio's GIS-based wetland inventory. *Journal of Soil and Water Conservation*, 49: 23–28.
- Zhang R., Ma J. (2008): An improved SVM method P-SVM for classification of remotely sensed data. *International Journal of Remote Sensing*, 29: 6029–6036.

Received for publication September 21, 2016
Accepted after corrections January 31, 2017



Paper II



Investigation of a possibility of spatial modelling of tree diversity using environmental and data mining algorithms

A. ABDOLLAHNEJAD, D. PANAGIOTIDIS, P. SUROVÝ

*Department of Forest Management, Faculty of Forestry and Wood Sciences,
Czech University of Life Sciences Prague, Prague, Czech Republic*

ABSTRACT: Biological diversity is the basis for a wide array of goods and services provided by forests. The variety of forest trees and shrubs plays a vital role in the daily life of forest communities. The purpose of this study is to investigate the possibility of modelling the diversity of tree species by characteristics of topography, soil and climate, using data mining algorithms k -NN, RF and SVM in Dr. Bahramnia forestry plan in the north of Iran. Based on the basal area factor for each species in a total of 518 sample plots, diversity indices such as species richness, evenness and heterogeneity were calculated for each plot. Topographic maps of primary and secondary properties were prepared using the digital elevation model. Categories of the soil and climate maps database of Dr. Bahramnia forestry plan were extracted. Modelling rates of tree and shrub species diversity using data mining algorithms, with 80% of the sampling plots were taken. Assessment of the model accuracy, using 20% of samples and evaluation criteria, was conducted. Results showed that topographic features, especially elevation, had the highest impact on the species diversity index. The modelling results also showed that Camargo evenness index had lowest root mean square error (RMSE) (0.14) and RMSE% (24.35), compared to other indicators of diversity. In addition, the results of the comparison between the algorithms showed that the random forest algorithms were more accurate in modelling the diversity.

Keywords: topographic characteristics; suborder soil; climate; non-parametric algorithms; richness; evenness indicators

Diversity or taxonomy is the middle level of hierarchical biodiversity classification and its purpose is to assess the diversity of plant and animal species within certain areas (VAN DER MAAREL 2005). The main concern is to compare taxonomic groups in different geographical areas. Diversity is a significant part of biodiversity and can be divided into two categories: the first is richness and the second is evenness (LUDWING, REYNOLDS 1988; MAGURAN 1996; KREBS 1999). Richness is one of the basic indicators for measuring diversity in terms of the region and it has a direct, scientific effect on diversity, meaning that the higher the number of species, the higher the diversity. On the other hand, evenness is an indicator that shows the distribution of trees in different classes of species (EJTEHADI et

al. 2010). When comparing two different populations that have the same richness, the population that has higher evenness has higher diversity. On the contrary, when comparing two different populations that have the same evenness but not the same richness, the population with higher richness has higher diversity. In places where these two components (richness and evenness) are different, identifying areas with higher diversity becomes a difficult task. In order to solve this problem, non-parametric methods are used through the combination of richness and evenness components (ARDESTANI et al. 2010).

Forest diversity in the northern part of Iran is one of the richest forest ecosystems of the temperate forests. In order to achieve sustainable devel-

Supported by the Czech University of Life Sciences Prague, Project No. B07/15.

opment, further study of ecological and environmental factors which affect this ecosystem and the diversity of this forest is crucial (FALLAHCHAY, MARVIE MOHAJER 2005; MARVIE MOHAJER 2006). Some of the ecological factors which can affect biodiversity are elevation, aspect, slope, climate, and human activities (EJTEHADI et al. 2010). So far there are different kinds of studies that have been conducted in different territories, trying to predict or investigate diversity distribution related to topographic factors (POURBABAEE 1998; MARVIE MOHAJER 2005; GRACIA et al. 2007; ISMAILZADEH, HOSSEINI 2007; GHANBARI 2008; SAATCHI et al. 2008; KYMASI 2012), edaphic (QOMIOGHLI et al. 2006; EJTEHADI et al. 2010; KYMASI 2012) and climatic factors (MEHDINYA et al. 2006; PARMENTIER 2011; GIXHARI et al. 2012). One of the main purposes of modelling research is to clarify the most appropriate method, regarding the spatial prediction of forest characteristics, based on sampling methods (KINT et al. 2003).

Spatial analyses belong among the non-classical methods for data processing in order to estimate the information on unmeasured areas (WHITTAKER 1977). The principle of these models is based on the hypothesis that environmental factors can control the spatial distribution of plants (GHANBARI et al. 2011). Different kinds of models such as remote

sensing techniques, geostatistics, generalized linear regression, neural networks, nearest neighbours, decision trees and their variants, for example random forest, have been used to predict the biological characteristics of forests. According to FRANKLIN (1998) and SHATAEE et al. (2012), nearest neighbours, support vector machine and random forest, which can do both classification and regression, are the most common algorithms among data mining algorithms. Many researchers used these algorithms to model the quantity of forest characteristics (ISMAIL, MUTANGO 2010; O'SULLIVAN et al. 2010; PARMENTIER 2011; YAZDANI 2011; SHATAEE et al. 2012). Moreover, the advantages of non-parametric algorithms rely on the fact that they are not sensitive to a high number of independent variables as inputs for modelling.

The objective of this study is to investigate the possibility of modelling the spatial distribution of tree diversity using topographic, edaphic and climatic factors through three different types of algorithms.

MATERIAL AND METHODS

Study area. The study area is located in the southwest from Gorgan city in Golestan province in Iran ($36^{\circ}43' - 36^{\circ}46'N$, $54^{\circ}21' - 54^{\circ}24'E$). The total area is

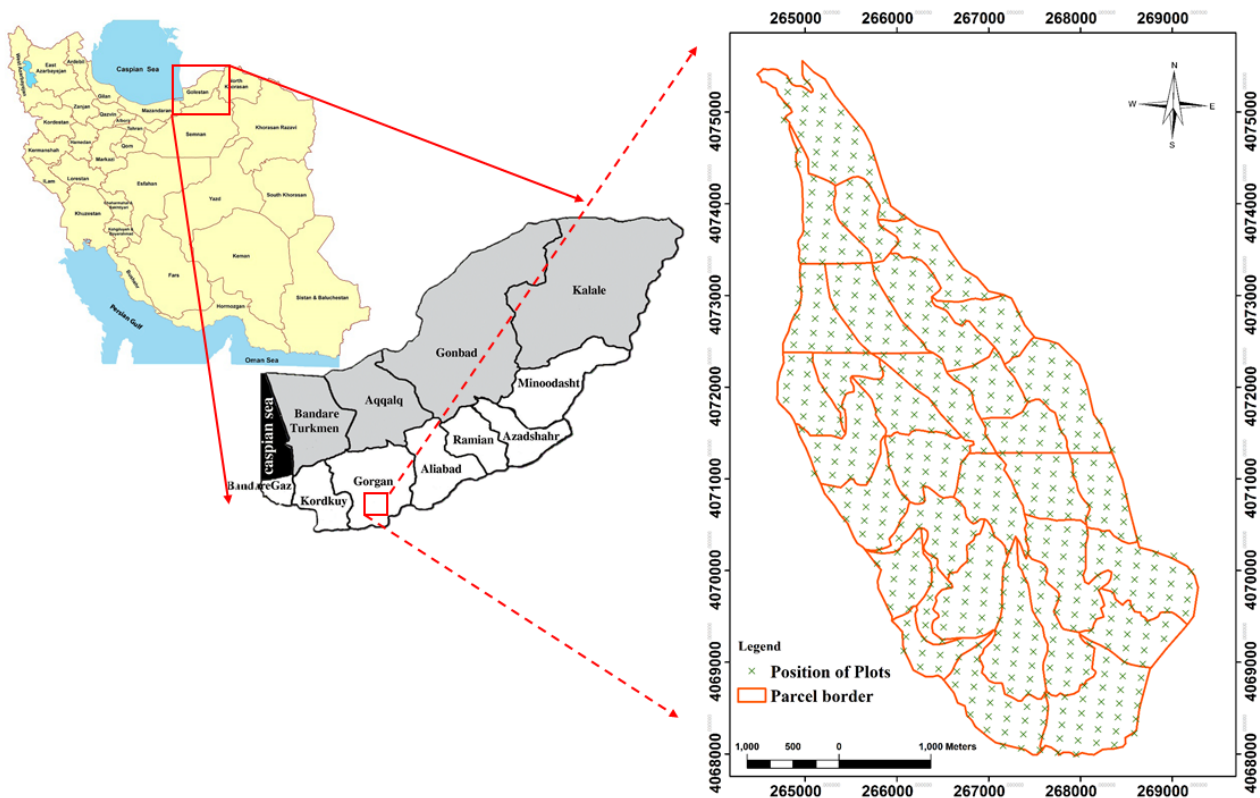


Fig. 1. Location of plots in Dr. Bahramnia's forestry district, Golestan province, northern Iran

Table 1. Primary topographic attributes that were computed by terrain analysis from digital elevation model data in this study (BEVEN, KIRKBY 1979; MOORE et al. 1993; WILSON, GALLANT 2000)

Characteristic	Definition	Importance
Altitude	elevation	vegetation, climate conditions, solar energy
Slope	gradient	flow rate, precipitation, vegetation, flow velocity, soil conditions
Aspect	slope azimuth	evapotranspiration, species distribution (fauna and flora), solar energy
Specific catchment area	used to estimate saturation excess overland flow	runoff volume and rate, soil characteristics, water viscosity, geomorphological conditions
Profile curvature	slope profile curvature	runoff acceleration, erosion/deposition percentage, geomorphological conditions
Tangential curvature	plan curvature multiplied by slope	an alternative measure of local flow conditions and divergence
Plan curvature	contour curvature	soil and water content, soil characteristics

1,714 ha (Fig. 1). The elevation of the study area ranges from 220 to 1,012 m a.s.l. and the slope range is between 0 and 80% and the soil type is brown and grey-brown. The average precipitation is 649 mm.

In regard to the aspect, 45% of the total area is facing the west, 42% the north, 10% the east and 3% the south. Concerning the forest type, the majority of the area is covered by *Carpinus-Zelkova*. There were also other types of forest such as: *Zelkova-Carpinus*, *Fagus-Carpinus* and *Fagus*.

Ground Data. To calculate the diversity indicators, we used the information on measured trees from 518 permanently visible sample plots with a radius of 17.5 m that were used in a systematic network of 150 × 200 m grids. The geographic position of plot centres and forest attributes such as diameter, height, crown diameter, name of species and tree health status were recorded on inventory forms. We included all trees with diameters greater than 10 cm for measuring the diversity indicators. Based on the basal area factor for each species in each plot, we were able to extract and calculate the amount of diversity indicators using the ecological methodology software by KREBS (1999).

We then constructed a digital elevation model – DEM (cell size = 30 m) for the study area using the topographic map (1:25000 scale and with 10 m contour interval). To check the DEM quality, we first used the hillshade tool to create a shaded relief from DEM, by considering the illumination source angle and shadows in order to be able to visually identify large errors (noise and sudden change in values). Then, we used the contour tool to recreate the contours from DEM (15 m contour interval), and finally we compared the original contours with new contours which came from interpolation.

Using a variety of software such as ArcGIS (Version 9.3, 2008) and TAS (Version 1.0, 2003), we

used DEM to construct primary and secondary topographic characteristic maps (Tables 1 and 2). Climatic maps such as: average annual precipitation (Eq. 1), average annual temperature (Eq. 2) and average annual evaporation (Eq. 3) were made using DEM in ArcGIS software and the formula for the Ghare Sou area based on information of meteorological stations for our study area during two decades (BYROODYAN 1990):

$$Y = \frac{282X^2 - 285,000X + 18.10^7}{X^2 - 1,000X + 45.10^4} \quad (1)$$

where:

Y – average annual precipitation (mm),

X – elevation.

$$T = -0.006X + 17.75 \quad (2)$$

where:

T – average annual temperature (°C).

$$ETP = 651 - 0.092X \quad (3)$$

where:

ETP – average annual evaporation (mm).

The zonal statistics algorithm was used for extracting the suborder soil factor from KARDGAR (2012) and other layers such as topographic and climatic factors using a buffer layer around the centre of the plots (17.5 m radius).

Diversity indicators. There are two different methods for measuring diversity: (i) numerical indicators, (ii) non-numerical or parametrical indicators. Numerical indicators present one number as a result. These types of indicators using richness component or evenness component or both together, can be divided into three categories: (i) richness indicators, (ii) evenness indicators, (iii) heterogeneity indicators. In this study we choose five dif-

Table 2. Secondary topographic attributes that were computed by terrain analysis from digital elevation model data in this study (BEVEN, KIRKBY 1979; MOORE et al. 1993; WILSON, GALLANT 2000)

Characteristics	Definition	Importance
Stream power indices (SPI)	$SPI = A_s \tan \beta_R$ where: A_s – specific catchment area, β_R – local slope angle.	It is a measure of erosive power of flowing water, predicts tangential concavity and net deposition in areas of profile concavity and net erosion in areas of profile convexity.
	$LS = (M + 1) \left(\frac{A_s}{22.13} \right)^m \left(\frac{\sin \beta}{0.0896} \right)^n$ where: LS – length-slope factor, $M = 0$, A_s – specific catchment area ($m^2 \cdot m^{-1}$), $m = 0.4$, β ($^\circ$) – slope gradient, $n = 1.3$.	It is the Revised Universal Soil Loss Equation in certain circumstances, predicts locations of net erosion and net deposition areas.
Topographic wetness index	$\ln \frac{a}{\tan b}$ where: a – local upslope area draining through a certain point per unit contour length, b – local slope in radians.	For steady-state flow conditions, it describes the spatial distribution of the saturation zone for runoff generation, soil transition, slope gradient.
Radiation indices	$R_{ne} = (1 - \alpha) R^\downarrow + \sigma (\epsilon_a T_a^4 - \epsilon_s T_s^4)$ where: R_{ne} – estimated net radiation ($W \cdot m^{-2}$), α – albedo (dimensionless), R^\downarrow – incoming short wave solar radiation ($W \cdot m^{-2}$), σ – Stefan-Boltzmann constant ($5.67 \times 10^{-8} W \cdot m^{-2} \cdot K^{-4}$), ϵ_a – atmospheric emissivity (dimensionless), determined according to equation $\epsilon_a = \phi (e_a / T_a) 1/7$ (BRUTSAERT 1975), ϕ – empirical coefficient, e_a – air vapor pressure (kPa), T_a – air temperature ($^\circ K$), ϵ_s – surface emissivity (dimensionless), T_s – surface temperature ($^\circ K$).	The three main terms account for direct-beam, diffuse, and reflected irradiance. A variety of methods are used by different authors to calculate these individual components. The methods vary tremendously in terms of sophistication, input data, and accuracy.

ferent kinds of indicators for measuring diversity based on previous research, study of changing the tree and shrub diversity in different environmental conditions which was conducted by ABDOLLAH-NEJAD and SHATAEE (2014).

(1) Richness indicators. These indicators are the simplest and oldest method for measuring the diversity and they are based on the number of species (s) and total number of individuals in the sample (N), one of these indicators is so called Menhinick's index – D_{Mn} (WHITTAKER 1977), as Eq. 4:

$$D_{Mn} = \frac{s}{\sqrt{N}} \quad (4)$$

(2) Evenness indicators. For evenness indicators we used two different types of index:

(i) Camargo index (E). Richness and scarce species cannot influence the Camargo index (CAMARGO 1992), as Eq. 5:

$$E^i = 1 - \left(\sum_{i=1}^S \sum_{j=i+1}^S \left[\frac{|P_i - P_j|}{S} \right] \right) \quad (5)$$

where:

P_i – number of i species,

P_j – number of j species,

S – total number of species.

(ii) Smith and Wilson index (E_{var}). This index, suggested by SMITH and BASTOW WILSON (1996), is based on the variance of species frequency, as Eq. 6:

$$E_{var} = 1 - \left[\frac{2}{\pi \arctan \left\{ \frac{\sum_{i=1}^S (\log_e(n_i) - \sum_{j=1}^S \log_e(n_j)/S)^2 / S}{S} \right\}} \right] \quad (6)$$

where:

n_i – number of i species,

n_j – number of j species,

S – total number of species in all the samples.

(3) Heterogeneity indicators.

(i) Shannon-Wiener index (H'). This index is the most common index and it can be calculated using SHANNON and WEAVER (1949), as Eq. 7:

$$H' = - \sum_{i=1}^s P_i \ln P_i = - \sum_{i=1}^s (P_i)(\log_2 P_i) \quad (7)$$

where:

P_i – percentage of i species,

s – number of species.

(ii) Simpson index (D). This is the most popular and the first non-parametric index for diversity and it is more sensitive to evenness than to richness. It can be calculated using SIMPSON (1949), as Eq. 8:

$$1 - D = 1 - \sum_{i=1}^s P_i^2 \quad (8)$$

where:

P_i – percentage of i species,

s – number of species.

Calculation of diversity indicators. For each plot, we calculated the cross-sectional area of each tree using the diameter, and then according to the tree species we were able to calculate the total basal area of each species in a sample plot using the ecological methodology software by KREBS (1999).

Applied algorithms. The prediction of tree diversity was based on 80% of the samples (414 plots) with different independent variables, such as topography, climate, edaphic factors and several different combinations of these variables.

Three different kinds of data mining algorithms were used to predict the diversity of tree species in the whole study area (all measured plots), and for our calculations we used the Statistica software (Version 7.0.61, 2006).

k -Nearest Neighbour (k -NN) is the most common algorithm based on training samples. The hypothesis of this algorithm is that all the samples are located in a space with n dimensions and it specifies the neighbours, based on the standard Euclidean distance.

The meaning of k is the number of nearest neighbours. In order to find the optimal number of k , we used the cross-validation method with k ranging from 1 to 50. To measuring the metric distance between the known (neighbour plots) and unknown plots (estimated plots using neighbours data), we used weighted Euclidean distance as the most appropriate option, providing comparison with other options that the software has.

Another algorithm that we used for estimating the diversity of species was Random Forest (RF) algorithm. To apply RF algorithm, we used 400 decision trees and we considered 5 as the minimum and 100 as the maximum number of nodes for each decision tree.

The final algorithm that we used was the so-called Support Vector Machine (SVM). The option that we used considers two kernel functions where one is the so-called type-1 regression and the other is the Radial Basis Function (RBF). Moreover, for improving the accuracy of the model prediction, we used the cross-validation method (YAZDANI 2011; SHATAEE et al. 2012).

Assessing the accuracy of modelling. The purpose of this assessment was mainly to investigate the ability of the models that we used for estimating diversity using training samples. We used relative root mean square error (RMSE), RMSE%, BIAS and BIAS% in order to evaluate the accuracy of the algorithm results.

RESULTS

The results of the inventory plots showed that the majority of the populations comprised 8 different species. In addition, the number of species as described in Table 3 had a high variance between the plots. As an example, based on inventory data we observed that in some plots there was only one spe-

Table 3. Statistical table displays the species distribution based on inventory data

Species	Cross section area (cm ²)			Basal area (cm ²)	Frequency	Species occurrence by plot
	minimum	maximum	average			
<i>Fagus orientalis</i> Lipsky	78.50	18,859.62	2,514.16	3,539,937.87	1,408	243
<i>Carpinus betulus</i> Linnaeus	78.50	20,096	1,568.18	5,087,192.50	3,244	440
<i>Quercus castaneifolia</i> C.A. Meyer	78.50	9,498.5	689.99	234,597.25	340	38
<i>Alnus subcordata</i> C.A. Meyer	78.50	18,859.62	795.95	506,226.87	636	102
<i>Acer velutinum</i> Boissier	176.60	1,256	330.35	11,892.75	733	247
<i>Zelkova carpinifolia</i> (von Pallas) C. Koch	78.50	5,671.62	330.20	200,430.12	36	6
<i>Diospyros lotus</i> Linnaeus	78.50	5,671.62	330.20	200,430.12	607	194
<i>Parrotia persica</i> (de Candolle) C.A. Meyer	78.50	13,266.50	703.41	3,014,812.12	4,286	426
Other species	176.62	3,316.62	488.11	83,955.75	172	34

Table 4. Results of the application of data mining algorithm in a case study (richness-evenness indicators)

Independent variables	Algorithms	Index											
		Menhinick's				Camargo				Smith and Wilson			
		RMSE	RMSE%	BIAS	BIAS%	RMSE	RMSE%	BIAS	BIAS%	RMSE	RMSE%	BIAS	BIAS%
Topography	<i>k</i> -NN	0.24	31.12	-0.18	-23.72	0.24	43.05	-0.10	-17.5	0.30	68.73	-0.205	-45.53
	SVM	0.26	33.19	-0.11	-13.43	0.24	42.45	-0.11	-19.63	0.28	64.42	-0.205	-49.7
	RF	0.25	31.12	-0.16	-20.3	0.14	24.35	-0.13	-23.63	0.24	51.8	-0.16	-35.88
Soil	<i>k</i> -NN	0.38	47.99	-0.06	-6.93	0.37	65.19	-0.29	-81.25	0.47	90.18	-0.207	-139.50
	SVM	0.30	37.92	-0.18	-24.02	0.24	43.22	-0.14	-26.53	0.31	72.30	-0.205	-41.55
	RF	0.25	32.37	-0.17	-21.018	0.23	39.13	-0.13	-23.86	0.27	60.03	-0.14	-30.53
Climate	<i>k</i> -NN	0.25	31.99	-0.17	-21.3	0.24	42.65	-0.11	-19.90	0.30	70.62	-0.205	-46.21
	SVM	0.25	31.74	-0.18	-23.59	0.24	42.82	-0.11	-18.59	0.28	64.81	-0.205	-49.46
	RF	0.26	32.90	-0.17	-21.76	0.17	30.01	-0.12	-21.77	0.28	59.50	-0.164	-34.88
All variables	<i>k</i> -NN	0.24	30.63	-0.18	-23.59	0.24	43.06	-0.10	-18.19	0.27	63.32	-0.205	-45.53
	SVM	0.27	33.92	-0.11	-12.43	0.25	44.09	-0.11	-20.30	0.30	69.11	-0.205	-47.81
	RF	0.28	34.33	-0.21	-26.85	0.16	24.74	-0.13	-23.65	0.26	54.67	-0.16	-35.46
10 affecting layers	<i>k</i> -NN	0.24	30.96	0.17	-21.67	0.24	42.89	-0.10	-18.04	0.28	65.82	-0.205	-44.60
	SVM	0.27	34.29	-0.11	-12.51	0.24	42.53	-0.11	-19.47	0.30	68.75	-0.205	-47.99
	RF	0.26	33.28	-0.01	-22.19	0.15	26.32	-0.14	-24.15	0.25	54.12	-0.162	-34.25

k-NN – *k*-Nearest Neighbour algorithm, SVM – Support Vector Machine algorithm, RF – Random Forest algorithm, RMSE – root mean square error

cies (low richness) while in other plots we found up to seven different species (high richness).

Using different combinations of independent variables: (i) topographic, (ii) edaphic, (iii) climatic ones, we created continuous maps for the whole study area where the results of accuracy evaluation by using 20% of the plots (testing plots) showed that the RF algorithm had the highest accuracy ~99%

of indicators, compared to the other two algorithms. More specifically, the results of RF showed RMSE% = 24.35 for Camargo index, RMSE% = 51.81 for Smith and Wilson index, RMSE% = 34.02 for Simpson index and finally RMSE% = 34.69 for Shannon-Wiener index. Menhinick's index had the highest accuracy using the *k*-NN algorithm with RMSE% = 30.63 (Tables 4 and 5).

Table 5. Results of the application of data mining algorithm in a case study (heterogeneity indicators)

Independent variables	Algorithms	Index							
		Shannon-Wiener				Simpson			
		RMSE	RMSE%	BIAS	BIAS%	RMSE	RMSE%	BIAS	BIAS%
Topography	<i>k</i> -NN	0.52	45.84	-0.904	-78.11	0.20	46.45	-0.07	-15.7
	SVM	0.51	45.17	-0.904	-80.99	0.19	44.84	-0.09	-20.69
	RF	0.46	38.07	-0.89	-77.14	0.17	37.05	-0.11	-25.66
Soil	<i>k</i> -NN	1.27	90.01	-0.904	-237.78	0.41	93.67	-0.32	-21.7
	SVM	0.56	49.57	-0.904	-73.42	0.21	49.23	-0.14	-35.65
	RF	0.49	41	-0.88	-73.62	0.19	40.30	-0.12	-26.73
Climate	<i>k</i> -NN	0.52	46.21	-0.909	-79.95	0.19	43.77	-0.085	-18.47
	SVM	0.50	44.57	-0.907	-76.50	0.19	84.43	-0.09	-20.92
	RF	0.44	36.54	-0.89	-74.83	0.17	36.28	-0.10	-23.26
All variables	<i>k</i> -NN	0.51	45.06	-0.904	-77.90	0.19	44.64	-0.08	-17.24
	SVM	0.52	46.25	-0.905	-82.57	0.21	47.28	-0.13	-32.37
	RF	0.43	34.69	-0.902	-80.28	0.16	34.02	-0.13	-15.42
10 affecting layers	<i>k</i> -NN	0.50	44.09	-0.904	-78.31	0.20	44.96	-0.07	-15.42
	SVM	0.51	44.95	-0.904	-81.11	0.20	46.91	-0.14	-35.86
	RF	0.48	39.17	-0.900	-77.76	0.17	37.44	-0.12	-26.69

k-NN – *k*-Nearest Neighbour algorithm, SVM – Support Vector Machine algorithm, RF – Random Forest algorithm, RMSE – root mean square error

Feature Selection and Variable Screening algorithms were used to find the 10 most important predictors that highly influence the variable of interest (Table 6). The results from estimation using these layers showed that the RF algorithm had the highest accuracy of modelling evenness and heterogeneity indicators and *k*-NN had the highest prediction accuracy of richness index (RMSE% = 30.96).

DISCUSSION AND CONCLUSIONS

According to the results of Menhinick's index (richness index) and heterogeneity indicators, we found out that modelling based on all independent variables had the highest accuracy. These results also showed that evenness maps which were made by topographic variables had the highest accuracy. From the conclusions above it is evident that richness and heterogeneity indicators are more connected to environmental factors like soil and climate, compared to the evenness indicators. According to the results of Table 6 it is clear that topographic factors, specifically elevation, had the most significant influence on the distribution of tree diversity. Changes in elevation can cause different environmental conditions such as temperature, precipitation, moisture, solar radiation, air pressure etc. and eventually they can affect the spatial distribution of trees. Moreover, due to the difficulties of human activities (selective cutting) at higher elevations these areas are closer to natural patterns (different frequency of species). Many studies like PEFFER et al. (2003), MARVIE MOHAJER (2006), GRACIA et al. (2007), GUOYU (2011), SHIRZAD and TABARI (2011), KYMASI (2012), and MOMENI MOGHADDAM et al. (2012) have shown that elevation is the most important environmental factor affecting the spatial distribution of species. This conclusion is also verified in our study. Additional research of BALE et al. (1998) and GHANBARI et al. (2011) showed that topographic factors like slope and aspect can also influence the spatial distribution of tree species. Different amounts of solar radiation, exposure, air-flow pressure, water flow sources, density of cloud cover and fog in different aspects can cause different environmental conditions that can affect the distribution and combination of tree species.

In addition, a varying amount of soil drainage in different percentages of slope can affect soil conditions and create an altered habitat for tree species.

Solar radiation is one of the secondary topographic characteristics that cause a significant impact on the spatial distribution of trees. According

Table 6. The most significant predictors (independent variables) in modelling diversity indicators

Independent variables	<i>P</i> -value
Elevation (most affecting factor)	0.0009
Tangential curvature	0.0848
Slope percentage	0.0417
Solar radiation	0.1223
Suborder soil	0.1254
Temperature	0.1824
Evaporation	0.1824
Aspect	0.2584
Specific catchment area	0.3373
Precipitation	0.3463

to Table 6, solar radiation had the greatest affect among the secondary topographic characteristics (PEFFER et al. 2003; SAATCHI et al. 2008; GHANBARI et al. 2011). Overall, using topographic layers can yield reasonable outputs in the modelling of spatial distribution of plants and diversity (ZIMMERMANN, KIENAST 1999).

In our study, according to Table 6, soil characteristics such as soil nutrients, soil depth, soil humidity and types of soil were some of the most important factors in the distribution of tree species because they influence site conditions (MCKENNEY, PEDLAR 2003; STEPHENSON et al. 2006).

Regarding the algorithms that were used in this paper, the results comparing different algorithms showed that RF had the highest accuracy compared to both *k*-NN and SVM. The ability of RF in determining important coefficients, weighting the independent variables and its non-necessity of decision tree structure pruning are all factors that enhance the functionality and effectiveness of this algorithm. On the contrary, *k*-NN and SVM algorithms use the same proportions of weighting for all independent variables (KERNES, OHMANN 2004). Comparing this with other algorithms, data mining algorithms are easier to comprehend; they need little data preparation (no need to normalize data) and can handle numerical and categorical data, in addition to that, a large amount of data can be analysed by data mining algorithms in a reasonable time. As a result, this enables researchers to study the effects of different independent variables on one factor simultaneously. Also, data mining algorithms such as RF can compare the effectiveness of independent variables and consider different weights for independent layers (input data) in the modelling process. Data mining algorithms with this potential for classification and regression of forest attributes at a high level of accuracy and robustness can be used as a promising approach in a wide range of forest research.

In conclusion, the above forestry plan was under management for many years, which affected the structure and combination of tree species and disconnected the links between the current forests and environmental factors. In other words, this type of managed forest appears to be completely different from a natural forest and as a result there is no pure connection between this forest and environmental conditions. In future work, we suggest this type of research for unmanaged forests which are not influenced by human activities and are more connected to topographic and edaphic factors.

References

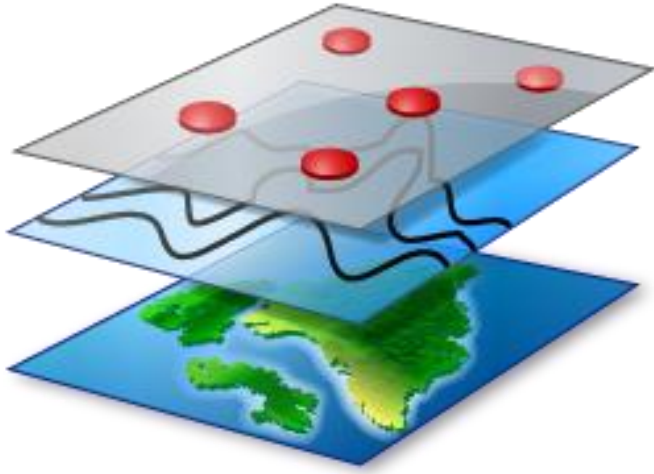
- Abdollahnejad A., Shataee S.H. (2014): The study of tree and shrub species diversity changes in the parameters of a physiographic, soil and vegetation. District one of DR. Bahramnia forestry plan. *Journal of Wood and Forest Science and Technology*, 21: 61–84. (in Persian)
- Ardestani E., Basir M., Torkesh M., Borhani M. (2010): Indicators for assessment of pasture species diversity in four places in Isfahan province. *Journal of Rangeland*, 4: 43–46. (in Persian)
- Bale C.L., Williams J.B., Charley J.L. (1998): The impact of aspect on structure and floristics in some Eastern Australian site. *Forest Ecology and Management*, 110: 363–377.
- Beven K.J., Kirkby M.J. (1979): A physically based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, 24: 43–69.
- Brutsaert W. (1975): On a derivable formula for longwave radiation from clear skies. *Water Resources Research*, 11: 742–744.
- Byroodyan M. (1990): *Weather and Climatology (Ghare Sou River Watershed Studies)*. Gorgan, Agriculture Publication: 300. (in Persian)
- Camargo J.A. (1992): New diversity index for assessing structural alterations in aquatic communities. *Bulletin of Environmental Contamination and Toxicology*, 48: 428–434.
- Ejtehadi H., Sepehry A., Akafi H.R. (2010): *Methods of Measuring Biodiversity*. Mashhad, University of Mashhad: 288. (in Persian)
- Fallahchay M.M., Marvie Mohajer M.R. (2005): Ecological role of elevation on tree diversity of Siahkal forest in the north of Iran. *Iranian Journal of Natural Resources*, 58: 89–101. (in Persian)
- Franklin J. (1998): Predicting the distribution of shrub species in Southern California from climate and terrain-derived variables. *Journal of Vegetable Science*, 9: 733–748.
- Ghanbari F. (2008): Predicting the spatial distribution of forest allometric growth properties using geostatistics and GIS. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 160. (in Persian)
- Ghanbari F., Shataee S.H., Mohseni A., Habashi H. (2011): Using a logistic regression model to predict the spatial characteristics of topography and forest type. *Iranian Journal of Forest and Poplar Research*, 19: 27–41. (in Persian)
- Gixhari B., Ismaili H., Vrapu H., Elezi F., Dias S., Sulovari H. (2012): Geographic distribution and diversity of fruit tree species in Albania. *International Journal of Ecosystems and Ecology Sciences*, 2: 355–360.
- Gracia M., Montané F., Piqué J., Retana J. (2007): Overstory structure and topographic gradient determining diversity and abundance of understory shrub species in temperature forest in central Pyrenees (NE Spain). *Forest Ecology and Management*, 242: 391–397.
- Guoyu L. (2011): Topography related spatial distribution of dominant tree species in a tropical seasonal rain forest in China. *Forest Ecology and Management*, 262: 1507–1513.
- Ismail R., Mutango O. (2010): Comparison of regression tree ensembles: Predicting *Sirex noctilio* induced water stress in *Pinus patula* forest of KwaZulu-Natal, South Africa. *International Journal of Applied Earth Observation and Geoformation*, 12: 45–51.
- Ismailzadeh A., Hosseini M. (2007): Relationship between ecological groups of plants with biodiversity indicators of plants in Afratakhteh cache for *Taxus bacata*. *Journal of Ecology*, 43: 21–30. (in Persian)
- Kardgar N. (2012): Accuracy assessment of soil maps in Dr. Bahramnia forestry plan. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 150. (in Persian)
- Kernes B.K., Ohmann J.L. (2004): Evaluation and prediction of shrub cover in coastal Oregon forests (USA). *Catena*, 55: 341–365.
- Kint V., van Meirvenne M., Nachtergale L., Geudens G., Lust N. (2003): Spatial methods for quantifying forest stand structures development: A comparison between nearest neighbour indices and variogram analysis. *Forest Science*, 49: 36–49.
- Krebs C.J. (1999): *Ecological Methodology*. 2nd Ed. Menlo Park, Addison-Wesley Educational Publishers, Inc.: 620.
- Kymasi F. (2012): Spatial distribution of tree and shrub species diversity in forests in Golestan province using GIS. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 170. (in Persian)
- Ludwing J.A., Reynolds J.F. (1988): *Statistical Ecology: A Primer on Methods and Computing*. New York, John Wiley & Sons: 202.
- Maguran A.E. (1996): *Ecological Diversity and Its Measurement*. Princeton, Chapman & Hall: 179.
- Marvie Mohajer M.R. (2006): *Silviculture*. Tehran, Tehran University Press: 387. (in Persian)
- McKenney D.W., Pedlar J.H. (2003): Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada. *Forest Ecology and Management*, 175: 497–507.

- Mehdinya T., Ejtehadi H., Sepehri A. (2006): Physiographic variables and the correlation between rainfall and vegetation communities present in the watershed of the Babol, Mazandaran province using geographic information systems. *Journal of Agricultural Sciences and Natural Resources*, 13: 99–107. (in Persian)
- Momeni Moghaddam T., Sagheb Talebi K.H., Akbarinia M., Akhavan M., Hosseini S.M. (2012): Impact of physiographic and edaphic factors on some of qualitative and quantitative characteristics of *Juniperus* trees. Case study: Layn region – Khorasan. *Iranian Journal of Forest*, 4: 143–156. (in Persian)
- Moore I.D., Gessler P.E., Nielsen G.A., Petersen G.A. (1993): Terrain attributes: Estimation methods and scale effects. In: Jakeman A.J., Beck M.B., McAleer M. (eds): *Modelling Change in Environmental Systems*. London, John Wiley & Sons: 189–214.
- O'Sullivan S., Keady E., Keane S., Irwin O'Halloran J. (2010): Data mining for biodiversity prediction in forests. In: Coelho H., Studer R., Wooldridge M. (eds): *Proceedings of the 19th European Conference on Artificial Intelligence*, Lisbon, Aug 16–20, 2010: 289–294.
- Parmentier I. (2011): Predicting alpha diversity of African rainforests: Models based on climate and satellite-derived data do not perform better than a purely spatial model. *Journal of Biogeography*, 38: 1164–1176.
- Peffer K., Pebesma E.J., Burrough P.A. (2003): Mapping alpine vegetation using vegetation observation and topographic attributes. *Landscape Ecology*, 18: 759–776.
- Pourbabae H. (1998): Biodiversity of wooden plants in the forests of Gillan. [Ph.D. Thesis.] Tehran, Tarbiat Modares University: 264. (in Persian)
- Qomioghli A.S., Hosseini M., Mataji A., Jalali G.H. (2006): Biodiversity of wooden plants on different soil in two different plant communities. *Journal of Biology*, 20: 200–207. (in Persian)
- Saatchi S., Buermann W., ter Steege H., Mori S.A., Smith T.B. (2008): Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment*, 112: 2000–2017.
- Shannon C.E., Weaver W. (1949): *The Mathematical Theory of Communication*. Urbana, University of Illinois Press: 163.
- Shataee S.H., Kalb S., Fallah A., Pelz D. (2012): Forest attribute imputation using machine-learning methods and ASTER data: Comparison of *k*-NN, SVR and random forest regression algorithms. *International Journal of Remote Sensing*, 33: 6254–6280.
- Shirzad M.A., Tabari M. (2011): Effect of some environmental factors on diversity of woody plants in *Juniperus excelsa* habitat of Hezarmasjed mountains. *Iranian Journal of Biology*, 24: 800–808. (in Persian)
- Simpson E.H. (1949): Measurement of diversity. *Nature*, 163: 688.
- Smith B., Bastow Wilson J. (1996): A consumer's guide to evenness indices. *Oikos*, 76: 70–82.
- Stephenson C.M., MacKenzie M.L., Edwards C., Travis J.M.J. (2006): Modeling establishment probabilities of an exotic plant, *Rhododendron ponticum*, invading a heterogeneous woodland landscape using logistic regression with spatial autocorrelation *Ecological Modeling*, 193: 747–758.
- van der Maarel E. (2005): *Vegetation Ecology*. London, Blackwell Publishing: 273.
- Whittaker R.H. (1977): Evolution of species diversity in land communities. *Evolutionary Biology*, 10: 1–67.
- Wilson J.P., Gallant J.C. (2000): *Terrain Analysis: Principles and Applications*. New York, John Wiley & Sons: 520.
- Yazdani S. (2011): Quantitative estimation of forest characteristics using QuickBird images. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 129. (in Persian)
- Zimmermann N.E., Kienast F. (1999): Predictive mapping of alpine grassland in Switzerland: Species versus community approach. *International Journal of Vegetable Science*, 10: 469–482.

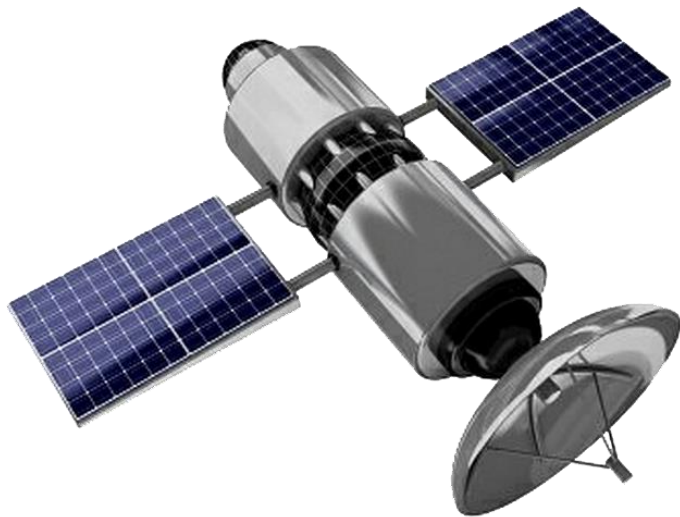
Received for publication August 1, 2016
Accepted after corrections October 18, 2016

Corresponding author:

Ing. AZADEH ABDOLLAHNEJAD, Czech University of Life Sciences Prague, Faculty of Forestry and Wood Sciences, Department of Forest Management, Kamýcká 1176, 165 21 Prague 6-Suchdol, Czech Republic;
e-mail: abdollahnejad@fld.czu.cz



Paper III



Article

Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data

Azadeh Abdollahnejad ^{1,*}, Dimitrios Panagiotidis ¹, Shaban Shataee Joybari ² and Peter Surový ¹

¹ Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha 165 21, Czech Republic; panagiotidis@fld.czu.cz (D.P.); surovy@fld.czu.cz (P.S.)

² Faculty of Forest Forestry, Gorgan University of Agriculture and Natural Resources, Basij Square, 386 Gorgan, Iran; shataee@yahoo.com

* Correspondence: abdollahnejad@fld.czu.cz; Tel.: +420-774-844-679

Academic Editors: Rodney J. Keenan and Timothy A. Martin

Received: 9 November 2016; Accepted: 10 February 2017; Published: 14 February 2017

Abstract: Modelling the spatial distribution of plants is one of the indirect methods for predicting the properties of plants and can be defined based on the relationship between the spatial distribution of vegetation and environmental variables. In this article, we introduce a new method for the spatial prediction of the dominant trees and species, through a combination of environmental and satellite data. Based on the basal area factor (BAF) frequency for each tree species in a total of 518 sample plots, the dominant tree species were determined for each plot. Also, topographical maps of primary and secondary properties were prepared using the digital elevation model (DEM). Categories of soil and the climate maps database of the Doctor Bahramnia Forestry Plan were extracted as well. After pre-processing and processing of spectral data, the pixel values at the sample locations in all the independent factors such as spectral and non-spectral data, were extracted. The modelling rates of tree and shrub species diversity using data mining algorithms of 80% of the sampling plots were taken. Assessment of model accuracy was conducted using 20% of samples and evaluation criteria. Random forest (RF), support vector machine (SVM) and k-nearest neighbor (*k*-NN) algorithms were used for spatial distribution modelling of dominant species groups using environmental and spectral variables from 80% of the sample plots. Results showed physiographic factors, especially altitude in combination with soil and climate factors as the most important variables in the distribution of species, while the best model was created by the integration of physiographic factors (in combination with soil and climate) with an overall accuracy of 63.85%. In addition, the results of the comparison between the algorithms, showed that the RF algorithm was the most accurate in modelling the diversity.

Keywords: spatial distribution; dominant tree species; topography; climate; soil; QuickBird; non-parametric algorithms

1. Introduction

The forests of Iran cover an area of about 12.4 million ha, comprising 7.4% of the country's area [1]. Of the five vegetation regions, the most important according to forest density, canopy cover and diversity, is the Hyrcanian (Caspian) region [2]. In forest ecosystems, trees and shrubs are living either independently (individually) or in association with each other, where some species are dominant over other species based on different biotic and non-biotic factors, comprising a stand with an area larger than 0.5 hectare or group species with an area smaller than 0.5 hectare.

Forest stand types or dominant trees and shrubs species mapping is one of the most important ways to manage and protect plant communities. Information about dominant tree species is required to assess forest resilience and vulnerability to any threat, for instance, drought and pathogens [3]. Investigating and using other methods with less cost and time can be an alternative approach for

field mapping. One of the alternative ways of mapping is the use of environmental data which are spatially related with forest tree grouped species or stands. Plant spatial distribution modelling is one of the indirect methods for predicting the properties of plants and can be defined as the relationship between the spatial distribution of vegetation and environmental variables. Recognition of the relationship between environmental factors and the plant species distribution plays an important role in environmental planning and ecosystem management [4,5]. Therefore, the appearance and stability of each plant species is generally influenced by environmental factors and the relationship between them and one or more environmental factors that have the greatest impact on their establishment.

In recent decades, airborne or space-borne remote sensing (RS) images, have been given the ability to record the electromagnetic radiation of vegetation at various wavelengths. Tree species identification using remote sensing is a classic topic in forestry [6]. The RS data with different spatial, spectral, and radiometric resolutions are presented as another alternative way for estimating and modelling vegetation attributes [2]. The estimated results were often differentiated according to chosen algorithms and studied area conditions [7]. However, some studies have already addressed the challenges in accurate tree species classification; Shataee et al. [8] and Mohammadi et al. [9] reported that using only spectral data could not provide useful information. Despite the emergence of very high resolution (VHR) sensors and their significance in the classification of dominant tree species, the limited number of spectral bands did not permit precise species discrimination [10]. Several other studies also investigated the possibility of using airborne hyperspectral imagery [11–13] LiDAR data or even a combination of multiple other techniques, for instance, auxiliary (environmental) data for prediction of tree species [14–16]. It is clear that differences in biochemical properties of different tree species and structural parameters of trees such as the texture and surface of the leaves, are better preserved using hyperspectral information, allowing continuous sampling of the electromagnetic spectrum. Moreover, LiDAR-based information about canopy structure to hyperspectral responses is also crucial for improving tree species classification of dominant tree species [17]. However, the operational use of these kind of data is still a challenging task, mainly because of their high cost and limited availability.

To overcome the above limitations, we attempted a different and more radical approach toward the prediction of dominant tree species, through a combination of satellite and auxiliary (environmental) data. Spatial distribution is directly dependent on both competition and environmental conditions such as the solar radiation, climate, water, topography and available nutrients. If these factors and their behaviors in relation to the distribution of species could be determined, it will be possible to achieve the prediction of dominant tree and shrub species distribution models [18]. Therefore, it is possible to investigate spatial distribution of these features by the environmental variables [19,20]. Until now, many studies have been conducted to predict the distribution of vegetation with physiographic, climate or soil factors. In many studies, physiographic factors have been reported as the most influential and were used to classify and predict vegetation characteristics as well as species distribution at different scales [21]. For instance, the use of elevation and slope [22,23], aspect [23,24] and secondary topography variable of potential solar radiation [22,24] were reported as the most important factors influencing the distribution of tree species. Additionally, edaphic factors such as soil type, moisture, soil temperature and soil nutrient availability to plants, are the most important factors affecting the distribution of tree species [25–27]. On the other hand, climatic factors can increase, decrease, or change the distribution of tree species depending on the species type. Several studies that investigated the relationship between climatic factors with the distribution of tree species, identified rain [28,29] and temperature [23,29] as the most important factors affecting the plant communities.

Also, previous studies have shown that the use of auxiliary data in combination with spectral data can improve the prediction results [30,31]. The use of high and very high spatial resolution satellite imagery (such as QuickBird IKONOS and Landsat) can result in high-precision imputation, especially for tree species and plot estimation [2]. Also, many studies concluded that, adding satellite-based spectral data in combination with environmental data can improve the prediction of qualitative and quantitative characteristics of the forests. Mohammadi et al. [31] showed an

improvement in the classification results of forest types by combining spectral and auxiliary data by determining the prior probability and creating spatial models of occurrence of forest types. With their results, many studies verified the potential for classification improvements by adding auxiliary data. As an example, Wheatley et al. [32] used topographic data with Landsat thematic mapper (TM) images to improve the accuracy of land cover maps. Also, Saatchi et al. [33] used the moderate-resolution imaging spectroradiometer (MODIS), quick scatterometer (QSCAT), the shuttle radar topography mission (SRTM), and the tropical rainfall measuring mission (TRMM) together with elevation and solar radiation data to model the potential distribution of tree species and diversity. In addition, Wang et al. [34] combined TM and auxiliary data, including digital elevation model (DEM), slope and moisture percentage in vegetation classification. Hernández Stefanoni et al. [35] combined geostatistical modelling and RS data to improve tropical species richness mapping. In another study, Adhikari et al. [36] investigated the possibility of spatial modelling of *Ilex Khasiana* P., using 16 environmental parameters, while Riemer Sørensen et al. [37] modelled the distribution of 10 commercial palm species using climatic, topographic, and spectral data, as well as a combination of them.

In the past two decades, parametric algorithms such as multiple linear and non-linear regressions, have been popular methods for estimating forest characteristics using non-spectral and spectral data [38,39]. Recently, non-parametric algorithms have been used for the prediction and estimation of forest attributes, because of some advantages they have over parametric algorithms, i.e., flexibility and the ability to describe non-linear dependencies, the fact that they are free from the assumption of any given probability distribution, and the fact that the observations are assumed to be independent of each other [40–42]. So far, non-parametric algorithms such as the generalized linear model (GLM), artificial neural networks (ANN), random forest (RF), support vector machine (SVM), and the k -nearest neighbor (k -NN), are used to predict the biological properties of forests [43]. The non-parametric machine learning techniques, have demonstrated superior performances over classic regression analysis for estimating forest attributes. The non-parametric models are more efficient than parametric models due to the possibility of using more than one independent variable simultaneously and also because there is no need for the data to be normally distributed. Among the non-parametric algorithms; the RF, SVM and k -NN are three of many machine learning algorithms that demonstrated good performance in forest attribute estimations [2]. Thus, in the present study, these three commonly used data mining classifier algorithms are used for tree species prediction.

1.1. k -Nearest Neighbor

The k -NN method is one of the simplest and most popular data-mining algorithms used for classification and regression. k -NN is widely used for the estimation of forest description using various topographic and remote-sensing data [44–49]. In k -NN implementations, three factors should be determined including the number of k , the type of distance measured and weights for nearest neighbors.

1.2. Support Vector Machine Classification

This algorithm is suitable for both classification and regression techniques based on statistical learning theory [50]. Generally, SVMs focus on the boundary between classes and map the input space created by independent variables using a non-linear transformation according to a kernel function. Linear, polynomial radial basis function (RBF) and sigmoid are the most commonly used kernel types. The RBF is the most popular kernel, which is used in SVMs [51,52]. According to our literature review, SVM has been used for forest classification [53–55]. In their work, Sheeren et al. [10] found SVM, among various non-parametric techniques, to be the best classifier with very close results to other classifiers among them (k -NN and RF).

1.3. Random Forest

RF is a new algorithm to the field of data mining and is designed to produce accurate predictions that do not overfit the data [56,57]. RF can also be used for regression-type problems (to predict a continuous dependent variable) and classification problems (to predict a categorical dependent variable). Implementation of RF depends on the regularization of decision tree and stopping parameters. The decision tree model parameters include the maximum number of trees that must be grown in the forest and the number of variables (k predictor or independent variables in each node for predicting depend values) that are randomly selected in each node [58]. Alternatively, choosing a small number of predictor variables may downgrade prediction performance, because this can exclude variables that may account for most of the variability and trends in the data [58]. The stopping parameters or control parameters are used to stop running the algorithm when satisfactory results have been achieved [2]. In some studies, such as Shataee et al. [2] and Garzón et al. [59] the RF has been used for the prediction of forest attributes.

The main aim of this study was to investigate the possibility of using high resolution satellite images derived from QuickBird with non-parametric algorithms, to estimate the distribution of tree species and also to investigate the effect of the environmental factors upon tree group species. Our literature review showed that no study has been carried out on the application of machine-learning algorithms (k -NN, SVM or RF algorithms) using QuickBird data for the prediction of distribution of tree species groups in the Hyrcanian forest. Therefore, for the purposes of this study, three of the most commonly used non-parametric machine-learning methods were applied for the first time to classify the spatial distribution of forest tree species as grouped (forest stand types) by climate, topography, soil data and a combination of them with the spectral data of QuickBird, in the Dr. Bahramnia forestry plan site in the Golestan province, in northeastern Iran. Overall, we wanted to study how helpful the combination between spectral and non-spectral data can be, in increasing the accuracy of the species distribution modelling (SDM).

2. Materials and Methods

2.1. Study Area

The study area is located south-west of Gorgan city in Golestan province in Iran ($36^{\circ}43' N$ to $36^{\circ}46' N$ and $54^{\circ}21' E$ to $54^{\circ}24' E$). The projection which was used in both Figures 1 and 2 was UTM WGS 1984 zone 40 N. The total area is 1714 hectares, where the green markers indicate the sample plots as seen in Figure 1. The elevation of the study area ranges from 220 to 1012 m, the slope range is between 0 and 80% and the soil type is characterized as brown and grey-brown (Figure 2). The average precipitation is 649 mm. In regards to aspect, 45% of the total area is facing to the west, 42% to the north, 10% to the east and 3% to the south. Regarding slope, 81% of the study area was between 0% and 30%, 18% between 31% and 60% and 1% above 61%. Information for different species is presented in Table 1.

Table 1. Statistical table displays the species distribution based on inventory data.

Species Name	Minimum of Cross Section Area (cm ²)	Maximum of Cross Section Area (cm ²)	Average of Cross Section Area (cm ²)	Basal Area (cm ²)
<i>Fagus orientalis</i> L.	78.50	18,859.62	2514.16	3,539,937.87
<i>Carpinus betulus</i> L.	78.50	20,096	1568.18	5,087,192.50
<i>Quercus castanaefolia</i> L.	78.50	9498.5	689.99	234,597.25
<i>Alnus miller</i> L.	78.50	18,859.62	795.95	506,226.87
<i>Acer velutinum</i> L.	176.60	1256	330.35	11,892.75
<i>Zelkova carpinifolia</i> Pall.	78.50	5671.62	330.20	200,430.12
<i>Diospyros lotus</i> L.	78.50	5671.62	330.20	200,430.12
<i>Parrotia persica</i> D.C.	78.50	13,266.50	703.41	3,014,812.12
Other species	176.62	3316.62	488.11	83,955.75

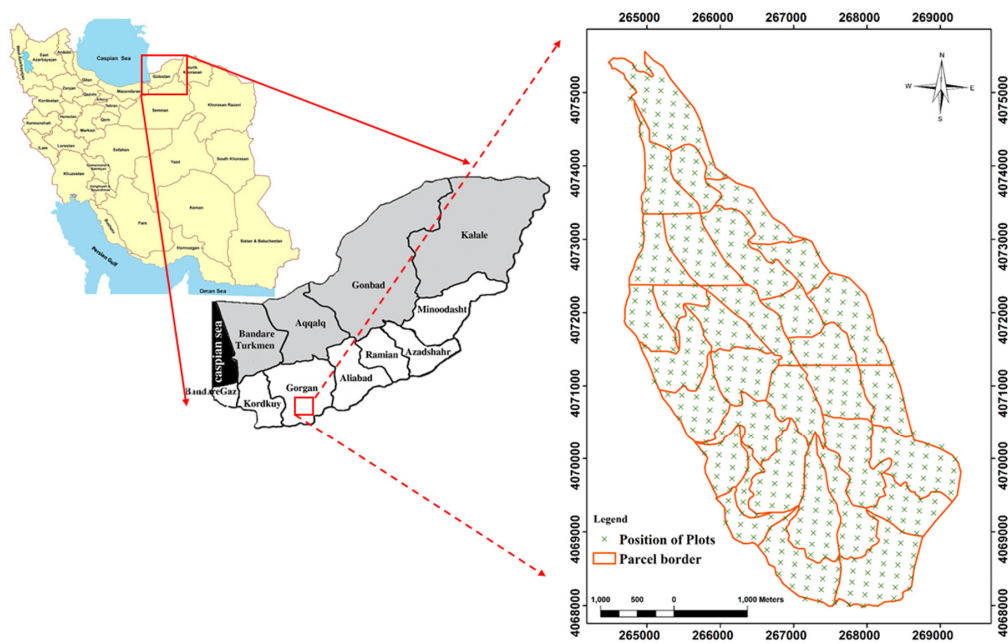


Figure 1. Location of the research area in Dr. Bahramnia’s forestry district, Golestan province in northern Iran.

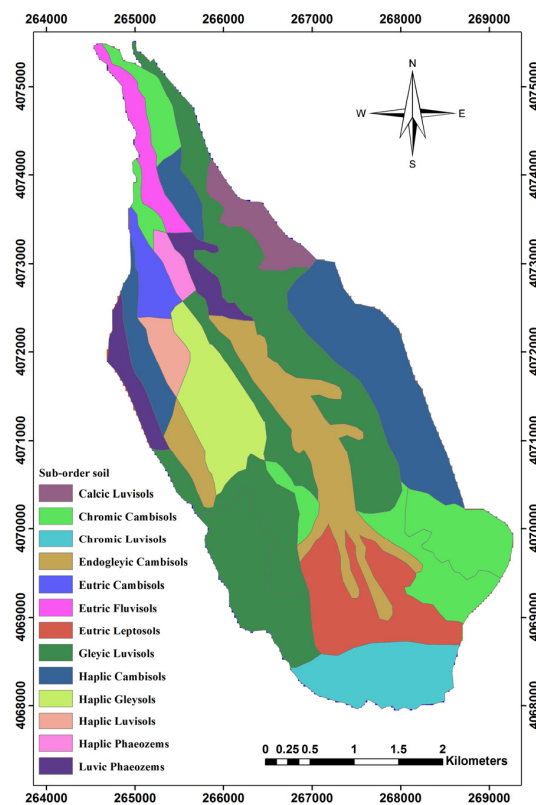


Figure 2. Classification of Sub-order types of soil within the study area.

2.2. Field Survey

To determine the dominant trees, we used the information of measured trees from 518 permanent sample plots with a radius of 17.5 m, plotted in a systematic network of 150m × 200 m grids.

The geographic position of plot centers and forest attributes such as diameter, height, crown diameter, name of species and tree health status were recorded on inventory forms. We included all trees with diameters greater than 10 cm for measuring the basal area. Based on the basal area factor (BAF) for every species in each plot, we were able to extract and specify the dominant tree species. Thus, any species with a basal area frequency above 50% in a plot, was selected as the dominant species. In addition, plots that did not contain a particular species group (basal area frequency of all species was lower than 50%) were excluded from the analysis, because they can cause uncertainty in tree species identification.

2.3. Environmental Data

We then constructed a DEM (cell size = 30 m) for the study area by using the topographic map (1/25,000 scale with 10 m contour interval). To evaluate the DEM quality, we first applied the hillshade tool in order to create a shaded relief from the DEM, by considering the illumination source angle and shadows to be able to visually identify large errors (noise and sudden change in values). Then, we used the contour tool in ArcGIS 10.1 software (ESRI, Redlands, LA, USA), to recreate the contours from the DEM (15 m contour interval) and finally compared the original contours with new contours which came from the interpolation.

Using a variety of software such as ArcGIS V.9.3 (ESRI, Redlands, LA, USA, 2008), and terrain analysis system (TAS V.1.0), (University of Western Ontario, Ontario, Canada, 2003), we used the DEM to construct primary and secondary topographic characteristic maps (Tables 2 and 3) and climatic maps such as average annual precipitation (Equation (1)), average annual temperature (Equation (2)) and average annual evaporation (Equation (3)). These equations were derived based on information provided by metrological stations during two decades [60]:

$$Y = \frac{282X^2 - 285,000X + 18.10^7}{X^2 - 1000X + 45.10^4} \quad (1)$$

where Y is the average annual precipitation (mm) and X is representing the elevation.

$$T = -0.006X + 17.75 \quad (2)$$

where T is the average annual temperature ($^{\circ}C$) and X is representing the elevation.

$$ETP = 651 - 0.092X \quad (3)$$

where ETP is the average annual evaporation (mm) and X is representing the elevation.

Table 2. Primary topographic attributes that were computed by terrain analysis from digital elevation model (DEM) data in this study. Sources: [61–63].

Characteristic	Definition	Importance
Altitude	elevation	vegetation, climate conditions, solar energy
Slope	gradient	flow rate, precipitation, vegetation, flow velocity, soil conditions
Aspect	slope azimuth	evapotranspiration, species distribution (fauna and flora), solar energy
Specific catchment area	used to estimate saturation excess overland flow	runoff volume and rate, soil characteristics, water viscosity, geomorphological conditions
Profile curvature	slope profile curvature	runoff acceleration, erosion/deposition percentage, geomorphological conditions
Tangential curvature	plan curvature multiplied by slope	an alternative measure of local flow conditions and divergence
Plan curvature	contour curvature	soil and water content, soil characteristics

Table 3. Secondary topographic attributes that were computed by terrain analysis from DEM data in this study. Sources: [61–63].

Characteristics	Definition	Importance
Stream power indices (SPI)	$SPI = A_s \tan \beta_R$ where: A_s —specific catchment area, β_R —local slope angle.	It is a measure of erosive power of flowing water, predicts tangential concavity and net deposition in areas of profile concavity and net erosion in areas of profile convexity.
	$LS = (M + 1) \left(\frac{A_s}{22.13} \right)^m \left(\frac{\sin \beta}{0.0896} \right)^n$ where: LS —length-slope factor, $M = 0$, A_s —specific catchment area ($m^2 \cdot m^{-1}$), $m = 0.4$, β ($^\circ$)—slope gradient, $n = 1.3$.	It is the Revised Universal Soil Loss Equation in certain circumstances, predicts locations of net erosion and net deposition areas.
Topographic wetness index	$\ln \frac{a}{\tan b}$ where: a —local upslope area draining through a certain point per unit contour length, b —local slope in radians.	For steady-state flow conditions, it describes the spatial distribution of the saturation zone for runoff generation, soil transition, slope gradient.
Radiation indices	$R_{ne} = (1 - \alpha) R^\downarrow + \sigma (\epsilon_a T_a^4 - \epsilon_s T_s^4)$ where: R_{ne} —estimated net radiation ($W \cdot m^{-2}$), α —albedo (dimensionless), R^\downarrow —incoming short wave solar radiation ($W \cdot m^{-2}$), σ —Stefan-Boltzmann constant ($5.67 \times 10^{-8} W \cdot m^{-2} \cdot K^{-4}$), ϵ_a —atmospheric emissivity (dimensionless), determined according to equation $\epsilon_a = \phi(e_a / T_a)1/7$ ϕ —empirical coefficient, e_a —air vapor pressure (kPa), T_a —air temperature ($^\circ K$), ϵ_s —surface emissivity (dimensionless), T_s —surface temperature ($^\circ K$).	The three main terms account for direct-beam, diffuse, and reflected irradiance. A variety of methods are used by different authors to calculate these individual components. The methods vary tremendously in terms of sophistication, input data, and accuracy.

The zonal statistic algorithm was used for extracting the sub-order soil factor (Table 4) from [64] and other layers such as topographic and climatic factors, by using a buffer layer around the center of plots (17.5 m radius).

Table 4. Characteristics of Sub-order types of soil.

Sub-Order Soil	Area (ha)	Area (%)
Calcic Luvisols	48.07	2.82
Chromic Cambisols	206.93	12.12
Chromic Luvisols	106.32	6.23
Endogleyic Cambisols	188.35	11.03
Eutric Cambisols	34.18	2
Eutric Fluvisols	43.99	2.58
Eutric Leptosols	137.37	8.05
Gleyic Luvisols	476.32	27.90
Haplic Cambisols	252.54	14.79
Haplic Gleysols	108.50	6.35
Haplic Luvisols	26.48	1.55
Haplic Phaeozems	17.20	1.01
Luvic Phaeozems	61.10	3.58

2.4. QuickBird Data

QuickBird was a high-resolution, commercial earth observation satellite, owned by Digital Globe launched in 2001 and decayed in January 2015. QuickBird used Ball Aerospace’s Global Imaging System 2000. The satellite collected panchromatic (black and white) imagery at 61 cm resolution and multispectral imagery resolution of 2.44 m at 450 km, and up to 1.63 m at 300 km, respectively as shown in Figure 3. In this study, a window of QuickBird image from 8 October 2008 has been used in

four multi-spectral bands with 2.4 m spatial resolution and one panchromatic band with 60 cm spatial resolution. The quantization radiometric level of images was 11 bits [65].

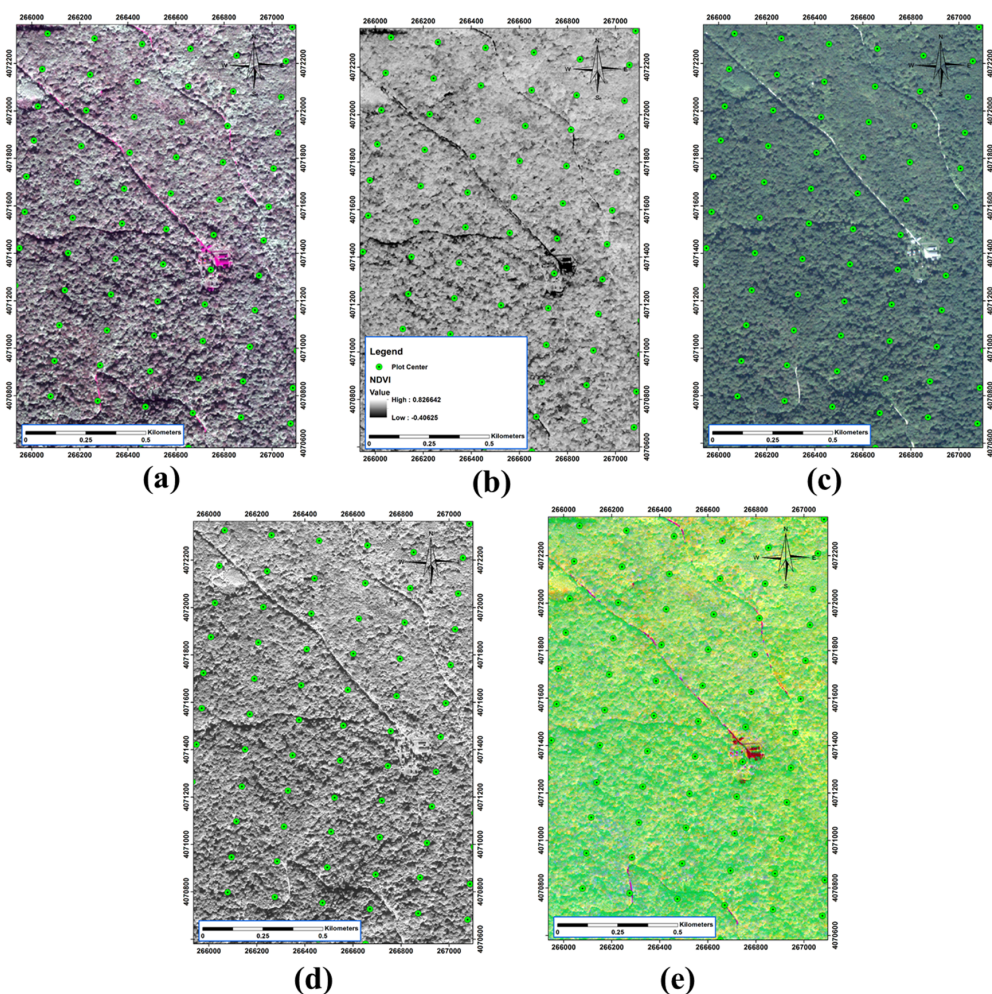


Figure 3. Illustration of some of the satellite outputs. (a) tasseled cap; (b) normalized difference vegetation index (NDVI); (c) multispectral image; (d) panchromatic image; (e) principal component analysis (PCA) band 1.

2.5. Pre-Processing and Processing of Spectral Data

In the first step, geometric correction was done using DEM with 10 m accuracy in order to eliminate the effect of displacement caused by the topography, using the quadratic equation and nearest neighbor sampling. The geo-referencing mean square error was less than one meter for multi-spectral bands and less than 40 cm for the panchromatic band. Processing of satellite images such as principal component analysis (PCA), tasseled cap transformation (brightness, greenness and moistness) (Table 5), normalized difference vegetation index (NDVI) and texture analysis was performed in order to enhance the spectral differences between tree species.

In this study, 13 texture analysis variables were used, such as mean, variance, entropy, contrast, heterogeneity, homogeneity, angular second moment, correlation, gray-level difference vector (GLDV) Angular second moment, GLDV entropy, GLDV mean, GLDV contrast and inverse difference with kernel window size 12×12 for RGB and infrared (IR). While 50×50 kernel size was used for the panchromatic band. Overall, 75 RS data layers were included in this study with other non-spectral layers.

Table 5. Conversion coefficient of tasseled cap for QuickBird images.

Coefficient	Component			
	A1	A2	A3	A4
Brightness	0.319	0.542	0.490	0.604
Greenness	−0.121	−0.331	−0.517	−0.780
Moisture(wetness)	0.652	0.375	−0.639	0.163

2.6. Extracting Data

The whole structure of layer preparation in this work is shown in Figure 4. To extract data, the zonal statistic algorithm in ArcGIS 10.1 software was used to extract the mean of non-spectral and spectral values for each plot, by using a buffer layer around the center of plots with a radius of 17.5 m.

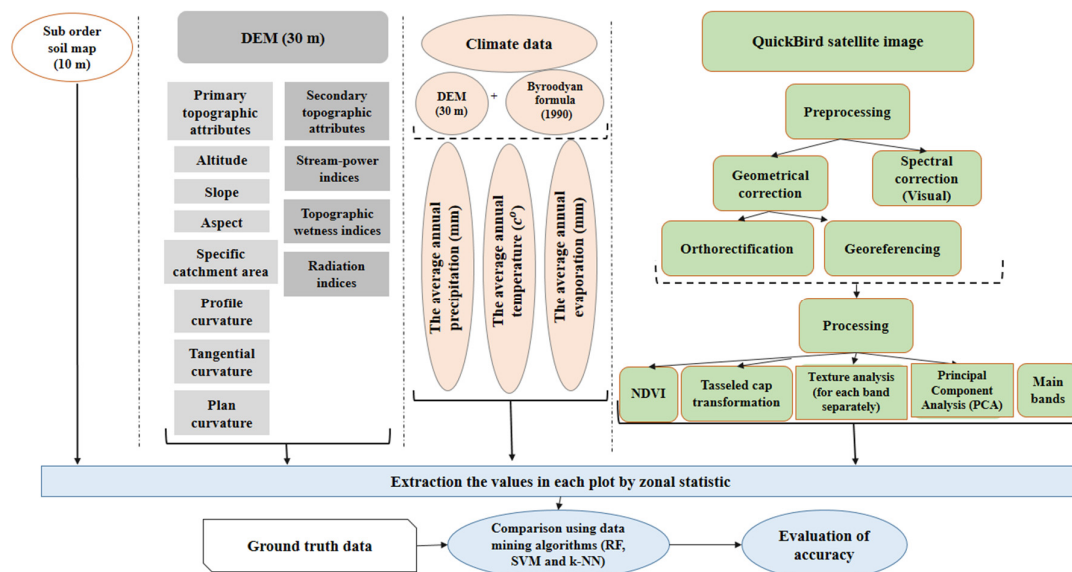


Figure 4. Data processing workflow illustrates the structure and order of layer preparation and different categories of layers which are used in our study.

Spatial distribution of forest species group modelling was done by grouping variables of topographic, climate, soil, and spectral RS data as a single group or combined using RF, *k*-NN and SVM algorithms, on 80% of the plots. Algorithm performances were evaluated by using 20% of unused plots in the modelling process by overall accuracy, user, and producer accuracy indicators.

2.7. Randomly Stratified Sample Splitting Method

In all statistical analyses that need to have prototype samples, i.e., training, test and validation data sets, the sample should be uniform and representative of all the data [66]. Therefore, the plots were first stratified based on their frequency distributions in different internal categories and then randomly divided into two training and validation sets. In other words, they had a sufficient distribution range in terms of their frequency in whole classes. This sample splitting method can be called a randomly stratified sample splitting [2] (Table 6). To achieve this, the values of plots based on basal area were stratified into categories that are currently being used for classification or mapping of discrete variables in forest management, and the frequency of plots was computed for each category. The training and validation samples were randomly selected as 80% for training plots and 20% for validation of the modelling in each category.

Table 6. Number of train and test plots in each class.

Species groups	Training Plots (80%)	Test Plots (20%)	Total
<i>Fagus orientalis</i> L.	91	24	115
<i>Carpinus betulus</i> L.	130	34	164
<i>Quercus castanaefolia</i> L.	7	2	9
<i>Alnus miller</i> L.	15	4	19
<i>Acer veutinum</i> L.	18	4	22
<i>Zelkova carpinifolia</i> Pall	1	0	1
<i>Parrotia persica</i> D.C.	59	15	74
Total	321	83	404

2.8. Implementation of Machine Learning Methods

k-Nearest Neighbor. In *k*-NN implementations, the number of *k*-NNs, the type of distance measure and the weighting for nearest neighbors are three important parameters. Determination of *k* is very important in terms of calculation time and producing unbiased results.

According to McRobert [67] a 1–20 range of *k* is the best option in terms of using the *k*-NN algorithm. He also stated that a higher *k* value may lead to lower noise effect during calculation, however, this requires taking a larger number of samples. It is important to mention that in the case of a higher *k* value, the pixel-level results will average towards the mean [68], leading to a higher bias and less precision. In most of the studies, the optimal *k* was reported to be between 5 and 10 [69–71]. However, in some other studies, such as [48] with *k* between 1 and 35 and [67] with *k* between 1 and 50, *k* values decreased the bias of the modelling process.

A smaller *k* often leads to a higher variance and less stable results [72]. Therefore, the optimal *k* is dependent on the data and goals of the estimation [73]. In addition, for an efficient comparison of distance measures in the *k*-NN algorithm, four distance measuring methods are available in Statistica software (StatSoft. Inc., Tulsa, OK, USA), including Euclidean, squared Euclidean, city block (Manhattan) and Chebychev. In this study, after a primary test on the results of different distance measurement methods and *k* values (between 5 and 50), weighted squared Euclidean distance, (*k* = 50) was used for modelling with the *k*-NN algorithm and its results were compared with each other.

Support Vector Machine Regression. The prerequisite for SVR to achieve better results, is the appropriate determination of the parameters that play key roles in achieving higher accuracy and better performance [74]. The specified grid search using *v*-fold cross-validation [52] is the most commonly used method to identify suitable parameters, i.e., epsilon (ϵ) and capacity (*C*) with fixed gamma that would produce high-accuracy results. A brief description of the proposed methods is summarized in [52]. In this study, RBF kernel was examined in a fixed number of gamma that are calculated based on 1/number of independent variables [75]. For selecting the best parameters, 10-fold cross-validation with 1000 iterations were used for minimizing the error function [76]. Chang and Lin et al. [75] used a specified grid search method to determine the best capacity and epsilon rates. The specified grid search included a range of capacity from 1 to 40, which is equal to the range of input variables [77] and epsilon values from 0.1 to 0.5.

Random Forest. For a high quality of RF classification performance, the decision tree model and stopping parameters should be regularized. For determination of the optimal tree number, 2000 initial trees, were used to produce a graph, showing the average squared error rates against each number of trees for training and test samples. RF is a powerful analytical tool for the exploration of data and verification of the optimal number of trees. By interpreting the graph, the optimal number of trees is found based on a tree number that produces a stable error. Then, the RF implementation was again repeated by using this optimal number of trees and other fixed parameters. In addition, default rates of stopping and splitting parameters were used to stop the process of growing trees, when stopping conditions were reached. The stopping parameters for all estimations included a minimum of one child node and a maximum of 100 nodes to stop growing the trees in 10 iterations for calculating the mean error and a 5% decrease in the training error.

3. Results and Discussion

Table 6 contains the numbers of training and test samples and indicates that the *Fagus orientalis* L., *Carpinus betulus* L., and *Parrotia persica* are the most frequent and dominant species. After modelling with each of the group variables individually as well as combined, the overall producer and user accuracies of the classes were presented in Table 7. The RF algorithm had a slightly higher performance (higher overall accuracy compared to the SVM and *k*-NN algorithm in almost all group variables). The best overall accuracy was obtained using a combination of topography, soil and climate factors (63.85%) by the RF algorithm, as shown in Table 7. Also, the results showed that modelling based on basal area frequencies for the species of *Carpinus betulus* L., *Fagus orientalis* L. and *Parrotia persica*, had the highest accuracy in classification compared to other species shown in Table 6. This may be due to greater frequency of training and test samples for these species groups.

3.1. Topographic Variables

The results (Table 7) showed that use of primary and secondary topographic variables derived from DEM could have a better result for two dominant group species—*Fagus orientalis* L. and *Carpinus betulus* L.—compared to any other species. Generally, the use of primary and secondary topographic variables can produce better results for modelling the distribution of a group or community, than for the distribution of individual tree species [78]. In other words, the spatial distribution of species communities are more dependent on the environmental conditions compared to the spatial presence of single trees or species. For example, one tree species such as *Fagus orientalis* L., can be presented in an extended geographical area, but as a community, it can only be present in limited areas. Topographical factors play a significant role in many environmental processes since they can create different microclimate conditions that could affect the habitats of species. In general, topography is one of the most important factors affecting species composition and distribution [79].

Elevation changes can cause ecological changes such as the air pressure, change of the amount of ultraviolet light, type and amount of rainfall, fluctuations in relative humidity and absolute humidity. Hence, these factors could also affect the composition and distribution of tree species. Based on the results, elevation between topographic variables had the highest effect on the spatial distribution of tree species. According to the RF results, elevation was introduced as the most important factor in the process of modelling. Our results corroborated previous studies [22,26,36,80,81] when we tried to examine the relationship between elevation and the spatial occurrence of species. However, our results contradicted the results of Wheatley et al. [32].

Also, primary topographic features such as slope and aspect had lower correlations with the spatial distribution of tree species and are in accordance with the results of Wheatley et al. [32], but opposed the results of other studies [23,24,27,80]. Similar biological characteristics of dominant species, as well as human interference, for example silvicultural treatments, management and utilization of forests, can lead to an ineffective relationship between slope-aspect parameters and the spatial distribution of tree species. Moreover, in our study, 80% of the area was in range of 0% to 30% slope; this homogeneity can reduce the effectiveness of the slope factor. In addition, about 90% of the total area has north and west aspects, due to the wide area in each direction, dominant species groups (*Fagus orientalis* L., *Parrotia persica*, and *Carpinus betulus* L.) are distributed in both directions and can result in the reduction of the coefficient of the aspect layer in modelling.

Table 7. The classification accuracy assessment of tree and shrub species groups by using Random forest (RF), support vector machine (SVM) and *k*-nearest neighbor (*k*-NN) algorithms.

Variables	Algorithm	Overall Accuracy (%)	Producer Accuracy (%)				User Accuracy (%)			
			Fagus Orientalis	Carpinus Betulus	Parrotia Persica	Other Species	Fagus Orientalis	Carpinus Betulus	Parrotia Persica	Other Species
Topography	RF	62.67	91.67	88.24	0	0	75	52.94	0	0
	SVM	61.44	79.16	91.17	0	0	79.16	54.38	0	0
	<i>k</i> -NN	61.44	79.16	91.17	0	0	79.16	54.38	0	0
soil	RF	53.01	58.33	88.24	0	0	66.67	48.38	0	0
	SVM	53.01	58.33	88.24	0	0	66.67	48.38	0	0
	<i>k</i> -NN	50.6	58.34	82.35	0	0	66.67	57.14	0	0
climate	RF	50.6	75	67.65	13.34	0	64.51	54	100	0
	SVM	56.62	79.17	82.35	0	0	65.51	50.9	0	0
	<i>k</i> -NN	54.42	79.16	79.41	0	0	67.85	55.1	0	0
Topography and climate	RF	62.65	87.5	91.18	0	0	80.76	54.38	0	0
	SVM	57.831	79.16	88.23	0	0	73.07	51.72	0	0
	<i>k</i> -NN	60.24	87.5	85.29	0	0	75	52.73	0	0
Topography and soil	RF	63.85	91.67	91.18	0	0	80	54.38	0	0
	SVM	61.44	87.5	88.23	0	0	77.78	51.72	0	0
	<i>k</i> -NN	61.44	79.16	91.76	6.67	0	79.16	54.38	50	0
Climate and soil	RF	53.01	75	82.35	0	0	69.23	52.73	0	0
	SVM	56.62	75	73.75	0	0	58.62	48.14	0	0
	<i>k</i> -NN	55.42	75	82.35	0	0	69.23	49.12	0	0
Topography, soil and climate	RF	63.85	87.5	75	0	0	85.29	52.73	0	0
	SVM	61.44	91.67	91.18	0	0	64.51	54	0	0
	<i>k</i> -NN	60.24	75	82.35	0	0	69.23	49.12	0	0
Spectral data	RF	54.21	87.5	70.58	0	0	60	50	0	0
	SVM	54.21	87.5	70.58	0	0	60	50	0	0
	<i>k</i> -NN	54.21	83.33	64.7	25	0	69.23	49.12	25	0
Topography and spectral	RF	60.24	87.5	85.3	0	0	75	52.72	0	0
	SVM	60.24	87.5	85.29	0	0	72.41	53.7	0	0
	<i>k</i> -NN	60.24	87.5	85.29	0	0	72.41	53.7	0	0
Topography, climate and spectral	RF	63.85	91.67	85.3	0	0	75.86	53.7	0	0
	SVM	59.03	87.5	82.35	0	0	72.41	51.85	0	0
	<i>k</i> -NN	59.03	87.5	82.35	0	0	72.41	51.85	0	0
Spectral and non-spectral data	RF	57.83	95.83	79.41	0	0	69.69	54	0	0
	SVM	61.44	87.5	76.47	6.67	0	63.63	53.06	100	0
	<i>k</i> -NN	60.24	95.83	79.41	0	0	69.69	54	0	0

Following the elevation parameter, solar radiation was the most important and influential independent variable in predicting the spatial distribution of species groups. The importance of solar radiation in the growing season on species spatial distribution was agreeable with the results of [22,24,26], but opposed the results of Wheatley et al. [32]. Wheatley et al. [22] proposed that using ground variables for mapping regions with homogenous elevation is the main reason for this contradiction. Spatial prediction of tree species, for the species which had limited ecological range (i.e., *Fagus orientalis* L.) were more correspondent in reality, compared to species which had wider distribution [82]. However, distribution of plants cannot be completely restricted by the topographic characteristics. This is because some species are resistant to various conditions, and they have wider ecological domains and higher ecological tolerance.

3.2. Climate

The results which predicted the spatial distribution of tree species using the climatic variables group showed that all three variables—temperature, evapo-transpiration and rainfall—were significant parameters in modelling the distribution of tree species, and agreed with the results of [25,33]. Also, our results agreed with many other studies with regards to the importance of temperature [23,25] and rainfall [23,28,37], in modelling the distribution of tree group species.

3.3. Sub-Order Soil

In general, Soil type is one of the most important factors in the development of plants. The physical, chemical, and biological properties of soil affected the distribution and establishment of plants. Previous studies [25,83,84] emphasized the importance of soil characteristics in species distribution.

3.4. Spectral Data

Using spectral data in modelling had lower accuracy (54.21%) in all algorithms compared to non-spectral data (Table 7). The result showed that applying spectral data as auxiliary data cannot improve the accuracies. The results of modelling with QuickBird spectral data proved that—with main bands and the texture analysis, such as mean and greenness—they were the best satellite data in modelling according to the importance factors obtained from RF analysis results [33,85]. Mean texture analysis of all bands had the highest importance coefficient in the distribution modelling of tree species groups, which agreed with [63], demonstrating the power of data obtained from the texture analysis of high-resolution images in presentation of the forest qualitative and quantitative characteristics.

Overall, according to Table 7, forest species group modelling, based on the frequency of the basal area, showed that topography variables had the best results (overall accuracy of 62.67%) compared to soil or climate variables. However, a combination of topography and soil variables improved the overall accuracy to 63.85%, but the combination of three variables, namely topography, soil and climate, could not improve the overall accuracy (60.24%). Finally, in this study, the synthesis of spectral data with the data from topography, climate and soil did not improve the results (overall accuracy of 61.44% using the SVM).

4. Conclusions

In remote sensing studies, it is common that a combination of spectral information of different objects can cause errors in sampling; in other words, uncertainty in samples. In our study, we experienced a similar problem because of the uncertainty of a few samples, in which the percentage of dominant tree species was close to 50%. To solve this problem, we tested our models by using plots in which the frequency of dominant trees was more than 75%, but still it could not significantly increase the accuracy of the models. Besides, we had to eliminate some tree species whose frequency was ranging between 50% and 75% per plot. Consequently, we decided to ignore the uncertainty of samples and the errors that they can cause.

In most studies, spectral data serve as the basis and non-spectral data are used as the auxiliary. However, the combined use of spectral and non-spectral data in this study showed much higher accuracies than modelling with only spectral data. Our results correspond to results of previous studies [8,33,34,36,37,86]. Integrating auxiliary data with data obtained from RS increases the modelling accuracy and contravenes with the results of [32].

In dense forests such as those found in the Hyrcanian region in the north-east part of Iran, spectral data can only be representative of the reflectances of dominant tree species in the upper part of the forest canopy and cannot provide spectral information regarding species, which may be encountered lower in the canopy. Therefore, the integration of non-spectral data with spectral data has caused an increase in classification accuracy of *Fagus orientalis* L., which is the dominant species in the canopy (over 90%). Consequently, inadequate accuracy of the exact location of the measured samples could be one of the reasons for reducing the effect of spectral data. High-resolution satellite images are sensitive to low errors in the spatial coordinates recorded by GPS devices.

Regarding the comparison analysis between the algorithms that we used, the classification accuracy of tree species by the RF algorithm could achieve higher results in terms of accuracy than the SVM and *k*-NN algorithms. Also, the RF algorithm was shown to provide generally higher accuracy and validity of the user and producer accuracy indicators than the SVM and *k*-NN algorithms. The capability of the RF algorithm in determining the weight coefficients of the independent variables without pruning the tree structure, increased the classification accuracy of the algorithm. On the contrary, the *k*-NN and SVM algorithms are not able to recognize the importance of each variable and they consider the same weight for all the independent variables in the modelling process. Moreover, RF results had the highest user accuracy compared to other algorithms. This means that understanding of the RF algorithm in change of tree species in each region is more reliable.

According to Yanoviak et al. [87], the RF algorithm compared to ANN and the decision tree methods, had higher modelling accuracy for the distribution of pine. In addition, according to Kernes et al. [88] tree models had a better understanding of the relationship and the boundaries, than logistic regression models for predicting the shrub cover spatial shifting. Also, according to [89], RFs are often used in very large geographical areas and when the number of samples in classes is unbalanced, the RF algorithm can be used with an acceptable level of accuracy for classification in such instances and it can be one of the factors for superiority over the SVM and *k*-NN. Naidoo et al. [90] studied the possibility of modelling savanna tree species in the Kruger national park in South Africa, using integrated hyperspectral, light detection and ranging (LiDAR) data and the RF algorithm and their results showed that the RF model produced 87% accuracy.

In general, data mining algorithms could not provide the desired accuracy in the mapping distribution of tree species. Various factors, such as lithology, geology, and human activities affect the distribution of tree species, however, these effects have not been considered in this study. Furthermore, recorded data plots with low accuracy (error in sampling) may reduce the classification accuracy of the distribution of tree species [36].

In conclusion, our results showed that topography, soil, and climate variables, influenced the distribution of tree species. However, topographic variables were the most important factors affecting the distribution of tree species. Combining spectral data with auxiliary data did not improve classifications results.

Acknowledgments: This research was supported by the projects of the Internal Grant Agency (IGA) of Faculty of Forestry and Wood Sciences, Czech University of Life Sciences (CULS) in Prague (No. B07/15) and (No. A14/16) and the research project of Ministry of Agriculture of Czech Republic (Grant No. QJ1520187). The authors would also like to acknowledge substantial help from Adam Dziadula of the Czech University of Life Sciences (CULS), for his contribution to the language review of the article.

Author Contributions: Azadeh Abdollahnejad designed the experiment and analyzed the data. Dimitrios Panagiotidis wrote the manuscript, helped on data interpretation and collaborated on the literature research, discussion and conclusion. Shaban Shataee helped on satellite data acquisition and ground data collection. Shaban shataee and Peter Surový supervised the project and the manuscript and granted the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Food and Agriculture Organization. Forests and the Forestry Sector. Available online: <http://www.fao.org/forestry/site/23747/en/Iran> (accessed on 21 December 2009).
2. Shataee, S.; Kalbi, S.; Fallah, A.; Pelz, D. Forest attribute imputation using machine-learning methods and ASTER data: Comparison of k -NN, SVR and RF regression algorithms. *Int. J. Remote Sens.* **2012**, *19*, 6254–6280. [[CrossRef](#)]
3. Guyot, V.; Castagneyrol, B.; Vialatte, A.; Deconchat, M.; Selvi, F.; Bussotti, F.; Jactel, H. Tree diversity limits the impact of an invasive forest pest. *PLoS ONE* **2015**, *10*, e0136469. [[CrossRef](#)] [[PubMed](#)]
4. Austin, M.P. Spatial prediction of species distribution: On interface between ecological theory and statistical modelling. *Ecol. Model.* **2002**, *157*, 101–118. [[CrossRef](#)]
5. Ferrier, S.; Watson, G.; Pearce, J.; Drielsma, M. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiv. Conserv.* **2002**, *11*, 2275–2307. [[CrossRef](#)]
6. Boyd, D.S.; Danson, F.M. Satellite remote sensing of forest resources: Three decades of research development. *Prog. Phys. Geogr.* **2005**, *29*, 1–26. [[CrossRef](#)]
7. Salajanu, D.M.; Jacobs, D. Predicting Spatial Distribution of Privet (*ligustrum* spp.) In South Carolina from MODIS and forest Inventory Plot data. In Proceedings of the 17th Willam T. Memorial Remote Sensing Symposium, Denver, CO, USA, 1–20 November 2008.
8. Shataee, S.H. Course Seeks Maps of Forest Types Using Satellite Data (Case Study Kheyroud Kenar Noushahr). Ph.D. Thesis, Faculty of Forestry, Tehran, Iran, 2003.
9. Mohammadi, J.; Shataee, S.H.; Babanezhad, M. Estimation of forest stand volume, tree density and biodiversity using Landsat ETM + Data, comparison of linear and regression tree analyses. *Procedia Environ. Sci.* **2011**, *7*, 299–304. [[CrossRef](#)]
10. Sheeren, D.; Fauvel, M.; Josipovic, V.; Lopes, M.; Planque, C.; Willm, J.; Dejoux, J.F. Tree species classification in temperate forests using Formosat-2 satellite image time series. *Remote Sens.* **2016**, *8*, 734. [[CrossRef](#)]
11. Féret, J.B.; Asner, G.P. Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Trans. Geosci. Remote Sens.* **2012**, *51*, 73–84. [[CrossRef](#)]
12. Ghiyamat, A.; Shafri, H.Z.; Mahdiraji, G.A.; Shariff, A.R.M.; Mansor, S. Hyperspectral discrimination of tree species with different classifications using single- and multiple-endmember. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 177–191. [[CrossRef](#)]
13. George, R.; Padaliab, H.; Kushwahab, S.P. Forest tree species discrimination in western Himalaya using EO-1 Hyperion. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 140–149. [[CrossRef](#)]
14. Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [[CrossRef](#)]
15. Engler, R.; Waser, L.T.; Zimmermann, N.E.; Schaub, M.; Berdos, S.; Ginzler, C.; Psomas, A. Combining ensemble modeling and remote sensing for mapping individual tree species at high spatial resolution. *For. Ecol. Manag.* **2013**, *310*, 64–73. [[CrossRef](#)]
16. Ghosh, A.; Fassnacht, F.E.; Joshia, P.K.; Koch, B. A framework for mapping tree species combining hyperspectral and LiDAR data: Role of selected classifiers and sensor across three spatial scales. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *26*, 49–63. [[CrossRef](#)]
17. Holmgren, J.; Persson, A.; Soderman, U. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *Int. J. Remote Sens.* **2008**, *29*, 1537–1552. [[CrossRef](#)]
18. Guisan, A.; Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecol. Model.* **2000**, *135*, 147–186. [[CrossRef](#)]
19. Kint, V.; Van Meirvenne, M.; Nachtergale, L.; Gendens, G.; Lust, N. Spatial methods for quantifying forest stand structures development. A comparison between nearest neighbor indices and variogram analysis. *For. Sci.* **2003**, *49*, 36–49.
20. Hunter, J.T. Factors affecting range size differences for plant species on rock outcrops in eastern Australia. *Divers. Distrib.* **2003**, *9*, 211–220. [[CrossRef](#)]

21. Wilson, M.F.G.; O'Connell, B.; Brown Connell, B.; Brown, C.; Guinan, J.C.; Grehan, A.G. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Mar. Geod.* **2007**, *30*, 3–35. [[CrossRef](#)]
22. Ejtehadi, H.; Sepehry, A.; Horvath, F. Separability of forest vegetation type using environmental variables including elevation, slope, aspect and direct incoming solar radiation: A GIS application. In Proceedings of the 5th Conference on Geographic Information System, National Cartographic Center (NCC), Tehran, Iran, 10 May 1998.
23. Garzona, M.; Blazekb, B.; Netelerb, M.; Sanchez de Diosa, R.; Sainz Olleroa, H.; Furlanellob, C. Predicting habitat suitability with machine learning models: The potential area of *Pinus sylvestris* L. in the Iberian Peninsula. *Ecol. Model.* **2006**, *197*, 383–393. [[CrossRef](#)]
24. Ghanbari, F.; Shataee, S.H.; Mohseni, A.; Habashi, H. Using a logistic regression model to predict the spatial characteristics of topography and forest type (case study of a forest series Shastklath Gorgan). *J. Res. For. Poplar Res.* **2011**, *19*, 27–41.
25. Mellert, K.H.; Fensterer, V.; Küchenhoff, H.; Reger, B.; Kölling, C.; Klemmt, H.J.; Ewald, J. Hypothesis-driven species distribution models for tree species in the Bavarian Alps. *J. Veg. Sci.* **2011**, *22*, 635–646. [[CrossRef](#)]
26. Pfeffer, K.; Pebesma, E.J.; Burrough, P.A. Mapping alpine vegetation using vegetation observation and topographic Attributes. *Lands. Ecol.* **2003**, *18*, 759–776. [[CrossRef](#)]
27. Navroudi, B.H.; Namiranian, M.; Mohanjer, M.M.; Azizi, P. Effect of aspect and altitude above sea level on the amount of volume of natural Rashsthanhay. *Iran. J. Nat. Resour.* **2000**, *53*, 201–215. (In Persian)
28. Mehdinya, T.; Ejtehadi, H.; Sepehri, A. Physiographic variables and the correlation between rainfall and vegetation communities present in the watershed of the Babol, Mazandaran province using geographic information systems. *J. Agric. Sci. Nat. Resour.* **2006**, *13*, 107–199. (In Persian)
29. Wong, C.S.C.; Li, X.; Thornton, I. Urban environmental geochemistry of trace metals. *Environ. Pollut.* **2006**, *142*, 1–16. [[CrossRef](#)] [[PubMed](#)]
30. Shataee, S.H. Improved classification of forest types by combining spectral data and help establish a method to determine the probability of occurrence of classes of models. *Nat. Mapp. Agency* **2004**, *83*, 1–6.
31. Mohammadi, J.; Shataee, S.H.; Yaghmaee, F.; Mahiney, A. Modelling forest stand volume and tree density using Landsat ETM+ data. *Int. J. Remote Sens.* **2010**, *7*, 2959–2975. [[CrossRef](#)]
32. Wheatley, J.M.; Wilson, J.P.; Redmond, R.L.; Ma, Z.; Dibenedetto, J. Automated land cover mapping using Landsat thematic mapper images and topographic attributes. In *Terrain Analysis, Principles and Applications*; Wilson, J.P., Gallant, J.C., Eds.; John Wiley and Sons: Hoboken, NJ, USA, 2000.
33. Saatchi, S.; Buermann, W.; Steege, H.; Mori, S.; Smith, B. Modelling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sens. Environ.* **2008**, *112*, 2000–2017. [[CrossRef](#)]
34. Wang, X.P.; Tan, Z.Y.; Fang, J.Y. Climatic Control on Forests and Tree Species Distribution in the Forest Region of Northeast China. *J. Integrative Plant Biol.* **2006**, *48*, 778–789. [[CrossRef](#)]
35. Hernández-Stefanoni, J.L.; Gallardo-Cruz, J.A.; Meave, J.A.; Dupuy, J.M. Combining geostatistical models and remotely sensed data to improve tropical tree richness mapping. *Ecol. Indic.* **2011**, *11*, 1046–1056. [[CrossRef](#)]
36. Adhikari, D.; Barik, S.K.; Upadhaya, K. Habitat distribution modelling for reintroduction of *Ilex khasiana* Purk, a critically endangered tree species of northeastern India. *Ecol. Eng.* **2012**, *40*, 37–43. [[CrossRef](#)]
37. Riemer-Sørensen, S.; Parkinson, D.; Davis, T.M. Simultaneous constraints on the number and mass of rekativistic species. *Astrophys. J.* **2013**, *30*, 763.
38. Hall, R.J.; Skakun, R.S.; Arsenault, E.J.; Case, B.S. Modelling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. *For. Ecol. Manag.* **2006**, *225*, 378–390. [[CrossRef](#)]
39. Gebreslasie, M.T.; Ahmed, F.B.; Aardt, J. Estimating plot-level forest structural attributes using high spectral resolution ASTER satellite data in even-aged Eucalyptus plantations in southern KwaZulu-Natal, South Africa. *South. For.* **2008**, *70*, 1–10.
40. Hyvonen, P. The Updating of Forest Resource Data for Management Planning for Privately Owned Forests in Finland. Ph.D. Thesis, Faculty of Forestry, Joensuu, FL, USA, 2007.
41. McRobert, R.; Tomppo, E.; Finley, A.; Heillinen, J. Estimating aerial means and variances of forest attributes using the *k*-nearest neighbor's technique and satellite imagery. *Remote Sens. Environ.* **2007**, *111*, 466–480. [[CrossRef](#)]

42. Sironen, S.; Kangas, A.; Maltamo, M. Comparison of different non-parametric growth imputation methods in the presence of correlated observations. *Forestry* **2010**, *83*, 39–51. [[CrossRef](#)]
43. Franklin, J.; Mc Cullough, P.; Gray, C. Terrain variables used for predictive mapping of vegetation communities in southern California. In *Terrain Analysis, Principles and Applications*; Wilson, J.P.L., Gallant, J.C., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2000.
44. Franco-Lopez, H.; Ek, A.R.; Bauer, M.E. Estimation and mapping of forest stand density, volume, and cover type using the *k*-nearest neighbor's method. *Remote Sens. Environ.* **2001**, *77*, 251–274.
45. Katila, M.; Tomppo, E. Selecting estimation parameters for the Finnish multisource national forest inventory. *Remote Sens. Environ.* **2001**, *76*, 16–32. [[CrossRef](#)]
46. Ohmann, J.L.; Gregory, M.J. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A. *Can. J. For. Res.* **2002**, *32*, 725–741. [[CrossRef](#)]
47. Makela, H.; Pekkarinen, A. Estimation of forest stands volumes by Landsat TM imagery and stand-level field-inventory data. *For. Ecol. Manag.* **2004**, *196*, 245–255. [[CrossRef](#)]
48. Finley, A.O.; McRobert, R.E.; Ek, A.R. Applying an efficient *k*-nearest neighbor search to forest attribute imputation. *For. Sci.* **2006**, *52*, 130–135.
49. Tatjana, K.; Suppan, F.; Schneider, W. The impact of relative radiometric calibration on the accuracy of *k*-NN predictions of forest attributes. *Remote Sens. Environ.* **2007**, *110*, 431–437.
50. Walton, J.T. Sub pixel urban land cover estimation: Comparing cubist, random forests, and support vector regression. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 1213–1222. [[CrossRef](#)]
51. Cortez, P.; Morais, A. A data mining approach to predict forest fires using meteorological data. In Proceedings of the EPIA 2007—Portuguese Conference on Artificial Intelligence, Guimarães, Portugal, 3–7 December 2007.
52. Durbha, S.S.; King, R.L.; Younan, N.H. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sens. Environ.* **2007**, *107*, 348–361. [[CrossRef](#)]
53. Zhang, R.; Ma, J. An improved SVM method P-SVM for classification of remotely sensed data. *Int. J. Remote Sens.* **2008**, *29*, 6029–6036. [[CrossRef](#)]
54. Shafri, H.Z.M.; Ramle, F.S.H. A comparison of support vector machine and decision tree classifications using satellite data of Langkawi islands. *Inf. Technol. J.* **2009**, *8*, 64–70. [[CrossRef](#)]
55. Ostapowicz, K.; Lakes, T.; Kozak, J. Modelling of land cover change using support vector machine. In Proceedings of the 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 10–14 May 2010.
56. Breiman, L. Random Forests. *Mach. Learn* **2001**, *45*, 5–32. [[CrossRef](#)]
57. Breiman, L. Using Models to Infer Mechanisms. IMS Wald Lecture 2. Available online: <http://www.stat.berkeley.edu/~breiman/wald2002--2.pdf> (accessed on 19 September 2002).
58. Statistica. Electronic Text Book, Stat Soft Inc. Available online: <http://www.Statsoft.com> (accessed on 20 November 2010).
59. Garzón, M.B.; Dios, R.S.; Ollero, H.S. Effects of climate change on the distribution of Iberian tree species. *Appl. Veg. Sci.* **2008**, *11*, 169–178. [[CrossRef](#)]
60. Byroodyan, M. Weather and Climatology (Ghare Sou River watershed studies). *Agric. Publ.* **1990**, *13*, 300. (In Persian)
61. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [[CrossRef](#)]
62. Moore, I.D.; Gessler, P.E.; Nielsen, G.A.; Petersen, G.A. Terrain attributes: Estimation methods and scale effects. In *Modelling Change in Environmental Systems*; Jakeman, A.J., Beck, M.B., McAleer, M., Eds.; Wiley: London, UK, 1993; pp. 189–214.
63. Wilson, J.P.; Gallant, J.C. *Terrain Analysis: Principles and Applications*; John Wiley and Sons: Hoboken, NJ, USA, 2000.
64. Kardgar, N. Accuracy Assessment of Soil Maps in Dr. Bahramnia Forestry Plan Bahramnia (Section A). Master's Thesis, Faculty of Forestry, Tehran, Iran, 2012. (In Persian)
65. DigitalGlobe. QuickBird Satellite Imagery Products. Product Guide. Available online: http://www.glc.umd.edu/library/guide/QuickBird_Product_Guide.pdf (accessed on 1 May 2006).

66. Keller, F. Evaluation, Connectionist and Statistical Language Processing. Lecture. Available online: <http://www.coli.unisaarland.de/~crocker/Teaching/Connectionist/lecture114up.pdf> (accessed on 7 July 2010).
67. McRobert, R.E. Diagnostic tools for nearest neighbor techniques when used with satellite imagery. *Remote Sens. Environ.* **2009**, *113*, 489–499. [[CrossRef](#)]
68. Nilsson, M. Estimation of Forest Variables Using Satellite Image Data and Airborne LiDAR. Ph.D. Thesis, Faculty of Forest Science, Umeå, Sweden, 1997.
69. Reese, H.; Nilsson, M.; Sandstrom, P.; Olsson, H. Applications using estimates of forest parameters derived from satellite and forest inventory data. *Comput. Electr. Agric.* **2002**, *37*, 37–55. [[CrossRef](#)]
70. Gu, H.Y.; Dai, L.M.; WU, G.; Xu, D.; Wang, S.Z.; Wang, H. Estimation of forest volumes by integrating Landsat TM imagery and forest inventory data. *Sci. Chin. Ser. E Technol. Sci.* **2006**, *49*, 54–62. [[CrossRef](#)]
71. Kutzer, C. Potential of the *k*-NN Method for Estimation and Monitoring Off-Reserve Forest Resources in Ghana. Ph.D. Thesis, Faculty of Forest and Environmental Sciences, Albert-Ludwigs Universität, Freiburg im Breisgau, Germany, 2008.
72. Kozma, L. *k*-Nearest Neighbor Algorithm (*k*-NN). Helsinki University of Technology Special Course in Computer and Information Science. Available online: <http://www.lkozma.net/knn2.pdf> (accessed on 20 February 2008).
73. Larose, D.T.; Larose, C.D. *k*-Nearest neighbor algorithm. In *Discovering Knowledge in Data: An Introduction to Data Mining*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; pp. 1–316.
74. Wang, Y.; Wang, J.X.; Du, W.; Wang, C.C.; Liang, Y.C.; Zhou, C.G.; Huang, L. Immune particle swarm optimization for support vector regression on forest fire prediction. *Int. Symp. Neural Netw. Adv. Neural Netw.* **2009**, *5552*, 382–390.
75. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines (Taipei: Department of Computer Science, National Taiwan University). Available online: <http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf> (accessed on 4 March 2013).
76. Schölkopf, B.; Smola, A.; Muller, K.R. Nonlinear component analysis as a kernel eigen value problem. *Neural Comput.* **1998**, *10*, 1299–1319. [[CrossRef](#)]
77. Matterna, D.; Haykin, S. Support vector machines for dynamic reconstruction of a chaotic system. In *Workshop in Advances in Kernel Methods: Support Vector Learning*; Schölkopf, B., Burges, C.J.C., Smola, A.J., Eds.; MIT Press: Cambridge, MA, USA, 1999.
78. Zimmermann, N.E.; Kienast, F. Predictive mapping of alpine grassland in Switzerland: Species versus community approach. *J. Veg. Sci.* **1999**, *10*, 469–482. [[CrossRef](#)]
79. Mohammadi, J. *Pedometry (Geographic Information Systems)*, 1st ed.; Pelk Publication: Tehran, Iran, 2006; pp. 1–637.
80. Guoyu, L. Topography related spatial distribution of dominant tree species in a tropical seasonal rain forest in China. *For. Ecol. Manag.* **2011**, *262*, 1507–1513.
81. Urli, M.; Delzon, S.; Eyermann, A.; Couallier, V.; García-Valdés, R.; Zavala, M.A.; Porté, A.J. Inferring shifts in tree species distribution using asymmetric distribution curves: A case study in the Iberian mountains. *INRA Sci. Impact* **2013**, *25*, 147–159. [[CrossRef](#)]
82. Chahouki, M.A.Z.; Ahvazi, L.K.; Azarnivand, H. Comparison of three modelling approaches for predicting plant species distribution in mountainous scrub vegetation (Semnan rangelands, Iran). *Pol. J. Ecol.* **2012**, *60*, 277–289.
83. McKenney, D.W.; Pedlar, J.H. Spatial models of site index based on climate and soil properties for two boreal tree species in Ontario, Canada. *For. Ecol. Manag.* **2003**, *175*, 497–507. [[CrossRef](#)]
84. Stephenson, C.M.; Mackenzie, M.L.; Edwards, C.; Travis, J.M.J. Modelling establishment probabilities of an exotic plant, *Rhododendron ponticum*, invading a heterogeneous woodland landscape using logistic regression with spatial autocorrelation. *Ecol. Model.* **2006**, *193*, 747–758. [[CrossRef](#)]
85. Shataee, S.H. *Course Booklet Remote Sensing*; Faculty of Forestry, Gorgan University of Agricultural Sciences and Natural Resources: Gorgan, Iran, 2005.
86. Hernández, P.A.; Franke, I.; Herzog, S.K.; Pacheco, V.; Paniagua, L.; Quintana, H.L.; Soto, A.; Swenson, J.J.; Tovar, C.; Valqui, T.H.; et al. Predicting species distributions in poorly-studied landscapes. *Biodivers. Conserv.* **2008**, *17*, 1353–1366. [[CrossRef](#)]

87. Yanoviak, S.P.; Gora, E.M.; Fredley, J.; Bitzer, P.M.; Muzika, R.M.; Carson, W.P. Direct effects of lightning in temperate forests: A review and preliminary survey in a hemlock–hardwood forest of the northern United States. *Can. J. For. Res.* **2015**, *45*, 1258–1268. [[CrossRef](#)]
88. Kernes, B.K.; Ohmann, J.L. Evaluation and prediction of shrub cover in coastal Oregon forests (USA). *Catena* **2004**, *55*, 341–365.
89. Freeman, E.A.; Moisen, G.G.; Frescino, T.S. Evaluating effectiveness of down-sampling for stratified designs and unbalanced prevalence in Random Forest models of tree species distributions in Nevada. *Ecol. Model.* **2012**, *233*, 1–10. [[CrossRef](#)]
90. Naidoo, L.; Mathieu, R.; Main, R.; Kleynhans, W.; Wessels, K.; Asner, P.G.; Leblon, B. The assessment of data mining algorithms for modelling Savannah Woody cover using multi-frequency (X-, C- and L-band) synthetic aperture radar (SAR) datasets. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 14 July 2014.



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Paper IV





Determining tree height and crown diameter from high-resolution UAV imagery

Dimitrios Panagiotidis, Azadeh Abdollahnejad, Peter Surový and Vasco Chiteculo

Department of Forest Management, Czech University of Life Sciences, Prague, Czech Republic

ABSTRACT

Advances in computer vision and the parallel development of unmanned aerial vehicles (UAVs) allow for the extensive use of UAV in forest inventory and in indirect measurements of tree features. We used UAV-sensed high-resolution imagery through photogrammetry and Structure from Motion (SfM) to estimate tree heights and crown diameters. We reconstructed 3D structures from 2D image sequences for two study areas (25 × 25 m). Species composition for Plot 1 included Norway spruce (*Picea abies* L.) together with European larch (*Larix decidua* Mill.) and Scots pine (*Pinus sylvestris* L.), whereas Plot 2 was mainly Norway spruce and Scots pine together with scattered individuals of European larch and Silver birch (*Betula pendula* Roth.). The involved workflow used canopy height models (CHMs) for the extraction of height, the smoothing of raster images for the determination of the local maxima, and Inverse Watershed Segmentation (IWS) for the estimation of the crown diameters with the help of a geographical information system (GIS). Finally, we validated the accuracies of the two methods by comparing the UAV results with ground measurements. The results showed higher agreement between field and remote-sensed data for heights than for crown diameters based on RMSE%, which were in the range 11.42–12.62 for height and 14.29–18.56 for crown diameter. Overall, the accuracy of the results was acceptable and showed that the methods were feasible for detecting tree heights and crown diameter.

ARTICLE HISTORY

Received 8 July 2016

Accepted 5 November 2016

1. Introduction

The determination of tree height is important, mainly because of biological and commercial reasons. It is a significant indicator, reflecting a site's productive capacity of the species concerned, when it is growing on a particular site. Site productive capacity means the total stand biomass produced by a stand on a particular site, up to any particular stage of its development, when the stand has been using fully the resources necessary for tree growth, which are available from the site. Furthermore, it is height, rather than biomass, that best reflects the site productive capacity (Shao, Zhao, and Shugart 1996).

CONTACT Dimitrios Panagiotidis  panagiotidis@fd.czu.cz  Department of Forest Management, Czech University of Life Sciences, Prague, Czech Republic

© 2016 Informa UK Limited, trading as Taylor & Francis Group

Similarly, tree crown size has been shown to be a multipurpose ecological indicator. It determines, among others, carbon sequestration, shading, risk of wind-breakage, and tree growth. The dependence of crown size on resource supply, species, and tree age complicates an accurate evaluation of a tree's space requirement, and its size-dependent functions and services in forested areas. Many remote-sensing applications involve the estimation of individual tree canopy area (Kalliovirta and Tokola 2005), as an intermediate stage in differentiating the signals reflected from the forest canopy and the forest floor, e.g. estimation of timber volume (Bolduc et al. 1999; Maltamo et al. 2004). Forest canopy height is an important structural parameter in several forest inventories; very-high-resolution digital models have been developed to identify and quantify individual tree crowns separately using remote technology. A digital elevation model (DEM) or a digital terrain model (DTM) is used for bare ground modelling without considering the features on the ground surface (Peckham and Gyozo 2007). On the other hand, digital surface models (DSMs) consider the features on the ground surface, e.g. trees and buildings. Methods used to generate DSMs usually focus on photogrammetric methods. However, this technique requires a well-trained workforce, high-quality digital cameras, and precise technology that will ensure the integrity of the data for accurate results. Several examples of the retrieval of canopy height using multi-angle/multi-view passive imagery from airborne platforms exist (Waser et al. 2008). Emerging techniques of computer vision in parallel with the recent advances of unmanned aerial systems (UASs) enable the extraction of reliable 2D and 3D imagery information from the collection of multi-angle imagery analysis taken with standard (non-metric) cameras and derived by Structure from Motion (SfM) algorithms (Küng et al. 2011; James and Robson 2012). These developments have made it possible to construct dense point clouds based on feature matches (alignment) within a series of overlapping photographs, even if internal and external camera orientation parameters are not available (Turner, Lucieer, and Watson 2012). New methods of photograph reconstruction enable the generation of very-high-resolution DSMs and orthophotograph mosaics with high spatial resolution (Turner, Lucieer, and Watson 2012; Fritz, Kattenborn, and Koch 2013; Dandois and Ellis 2013; Gini et al. 2014; Sona et al. 2014), and they have been used for environmental monitoring (Diaz Varela et al. 2014) and quantifying tree heights and crown dimensions (Zarco-Tejada et al. 2014). Several studies have identified the potential of combining UAV image flights and the SfM processing chain for terrain modelling (Remondino et al. 2011; Dandois and Ellis 2013), in terms of runtime and accuracy, as well as visual appearance of the resulting reconstructions.

A canopy height model (CHM) represents the difference between the top canopy surface and the underlying ground topography, i.e. subtraction of DSM from DTM. Popescu (2007) proposed that a CHM could be described as a 3D surface that contains all the necessary information about vegetation height above ground surface. The CHM can be effectively derived from digital image raw data through the appropriate filtering/classification of photograph point clouds to distinguish ground and canopy points; this separation can be achieved using various software, such as LasTools and Agisoft PhotoScan ©, and the latter is equipped with a tool for automatic and manual ground point classification (Agisoft). CHMs are raster images that allow us to delineate individual tree crowns and detect treetops. Individual tree identification (ITD) based on the CHM can be performed with a range of algorithms and procedures, including image smoothing, local maxima finding, and template matching (Persson, Holmgren, and Soderman

2002; Popescu, Wynne, and Nelson 2002; Koch, Heyder, and Weinacker 2006; Chen et al. 2006; Falkowski et al. 2006). By identifying the individual tree positions based on smoothing and local maxima filtering, tree height can be directly measured from CHMs using regression models (Pyysalo and Hyyppä 2002; Popescu, Wynne, and Nelson 2003).

Watershed segmentation is an important member of the boundary detection-based segmentation family (Carleer, Debeir, and Wolff 2005). Additionally, it is the most popular technique used for segmenting a CHM, because it is intuitive to treat each concave tree crown in the inverted CHM as a catchment basin. The surface of the raster is segmented into the equivalent of individual drainage basins by identifying local maximum and nearest minima values. Based on this approach, Inverse Watershed Segmentation (IWS) can delineate distinct tree entities with height values and crown diameters (Edson and Wing 2011). The method assumes that local maxima present in the CHM are treetops; however, caution should be exercised when applying local maxima in structurally complex forest structures. For example, multiple local maxima are often identified within a single broadleaf crown because of the irregularity of crowns and, in part, from random errors associated with the creation of the CHM. A possible solution is the use of smoothing filters when constructing the CHM; however, caution is required because strong filters may eliminate smaller trees. Popescu, Wynne, and Nelson (2003) suggested an improved version of local maxima filtering using a locally variable window size relative to tree height by referring to a predefined height-crown equation that accounts for tree crown development related to tree growth. However, other studies have reported problems correctly identifying tree crowns using a tree-level approach, particularly in closed-canopy and denser forests. According to Wang, Gong, and Biging (2004), the assumption that treetops are located around the vicinity of the centre of a crown can be met only when the view area is within $\pm 15^\circ$ of nadir. This assumption may not be applicable to trees located outside of this range because treetops lean away from the nadir point. Moreover, for trees viewed from outside of the near-nadir direction, the silhouettes detected from the edge-detection methods are inconsistent with the real tree crown boundaries, whereby trees are often misidentified or even missing in some cases.

The number and quality of algorithms are increasing rapidly. Sophisticated algorithms, such as IWS, are designed to consider increasingly more aspects, in either the preprocessing or post-processing of images, in an attempt to minimize oversight errors and provide a friendlier processing environment for the analyst. Delineation of individual trees is a complex process and sometimes an acceptable algorithm does not exist, even when using systems such as human vision, which is currently considered the most advanced image-processing tool. Bortolot (2006) investigated the practical advantages of using CHMs and proposed the use of CHM for identifying clusters of trees, which would correspond to a group of tree crowns for easier post-processing.

The key contribution of this article was to test the performance of photogrammetric techniques for the estimation of tree heights and crown diameters, based on UAV (DJI S800) SfM point clouds. We assessed the accuracy of height and crown diameter measurements derived from point clouds compared with ground measurements.

2. Materials and methods

2.1. Study area and experimental design

Two research plots from the Czech Republic (see [Figure 1](#)) were included in this work; Plot 1 is located southeast of Prague (N50.01577, E14.75998) and Plot 2 is located south of the city (N49.81452, E14.700590). Both plots were located on flat terrain and were 625 m² (25 X 25 m) in size. Plot 1 had 48 trees with predominantly Norway spruce (*Picea abies* L.), with a few European larch (*Larix decidua* Mill.) and Scots pine (*Pinus sylvestris* L.) individuals scattered throughout. Plot 2 had 39 trees with mainly Norway spruce and Scots pine trees, with scattered individuals of European larch and Silver birch (*Betula pendula* Roth.). We measured all trees on each plot to estimate tree heights and crown diameters.

2.2. Analysis overview

An illustration of the methodology is summarized in [Figure 2](#). Images were acquired using a Sony NEX-5 R digital camera, embedded on a UAV platform DJI S800 to create 3D reconstruction models for both study areas. The images were processed and analysed in Agisoft PhotoScan© to derive the DSM and DTM for both plots, using the SfM approach. The reconstructed mesh surface was derived from the point cloud by using

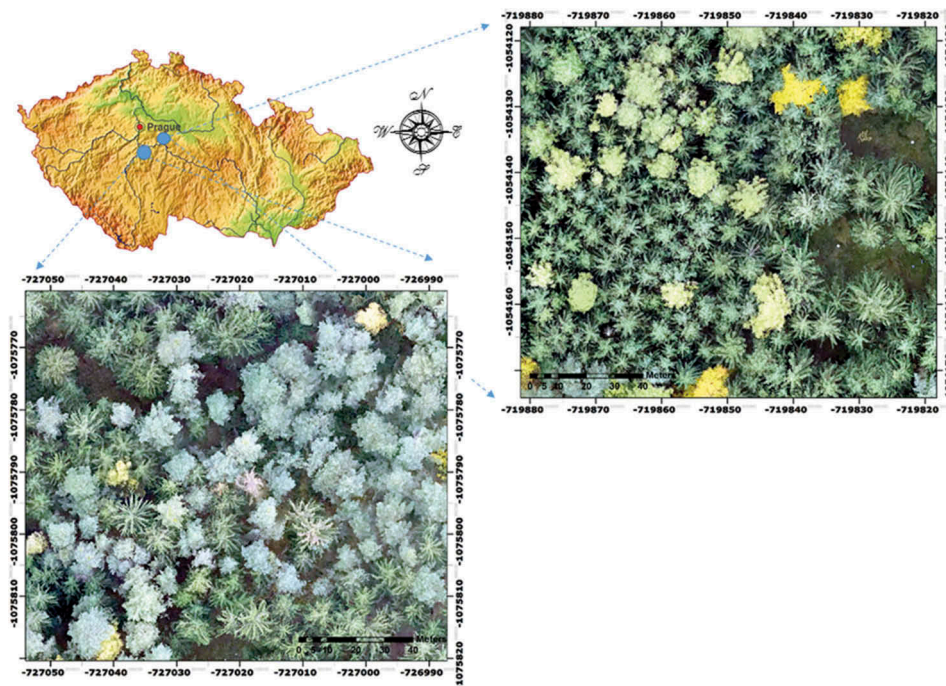


Figure 1. Plot locations in the Czech Republic. The top right photograph displays Plot 1 and the bottom left photograph displays Plot 2. Image sources: images provided as captured and post-processed from an RGB high-resolution embedded UAV camera using ESRI Inc. Map of the Czech Republic is from <http://www.mapcruzin.com/free-czech-republicmaps> (assessed on January 2016).

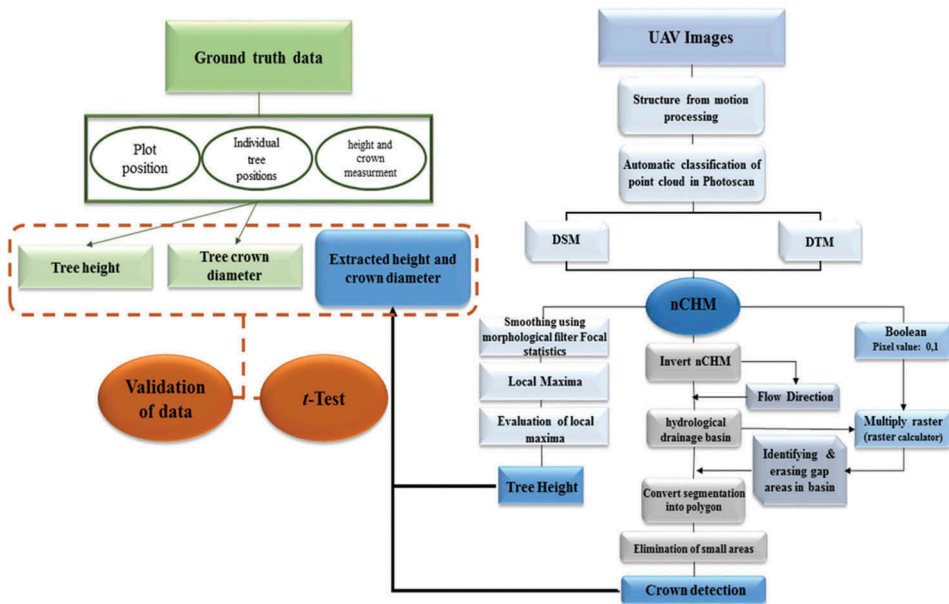


Figure 2. A flow chart to illustrate the analysis and outputs; the main objectives are in the shaded nodes.

the automatic classification of points procedure, based on three parameters: a) max angle, b) max distance, and c) cell size. To increase the accuracy of the resultant DSM and DTM models, we placed four ground control points (GCPs) in the corners of each square plot and we measured them with geodetic Real Time Kinematic (RTK) GPS. ArcGIS 10.3.1 by ESRI was used to create CHMs, from the difference in elevation between DSM and DTM (i.e. $CHM = DSM - DTM$). To estimate the tree heights, we used local maxima filtering, and for the crown diameters we used IWS in both study areas. Finally, statistical paired *t*-tests analyses were performed on the estimated and measured variables for both plots in order to evaluate the performance of the remote-sensing approach based on UAV SfM point clouds.

2.3. Field measurements

Tree heights were recorded using a TruPulse 360 /B laser range finder. To minimize height evaluation errors, the horizontal distance from the objective tree was at least equal to the tree height to be measured. We also measured the width of horizontal tree crown projections (east–west and north–south) for each tree using the TruPulse 360 /B laser range finder. We measured the positions of plot trees using the azimuth–distance approach from the plot centre, which was determined using RTK GPS Trimble Juno 3D handheld equipment; azimuth was estimated using a handheld compass and tree distance from plot centre was estimated using a Vertex III and Transponder T3 (Haglöf Sweden, 2002). We also created an index of the tree species for all measured trees. All field measurements for both plots were collected in the same month as the aerial images were taken.

2.4. Photograph acquisition and 3D model reconstruction

The DJI S800 UAV platform carried an RGB camera (see Figure 3). The camera was a Sony NEX-5R, with a 16 MP resolution and 4912×3264 pixels. The camera was mounted with a Voigtlander Super Wide Heliar lens, with a focal length of 15 mm. A drone needed approximately 7 min to complete a flight based on the predefined parameters (i.e. number of waypoints and time of flight), which we set using the DJI ground station software. A point cloud of images comprising the flight path lines covered the entire study area (see Figure 4). We performed all six flights, three flights per plot, in three different elevations using the same parameters, i.e. flight path pattern and altitudes.

The SfM image reconstruction process resulted in 673 of the original 718 images aligned in Plot 1 and 457 of the original 480 images aligned in Plot 2. We extracted both



Figure 3. This is a scaled image that illustrates the actual size of the multirotor UAV (DJI S800) that was used for photograph acquisition (source: Přemysl Janata).

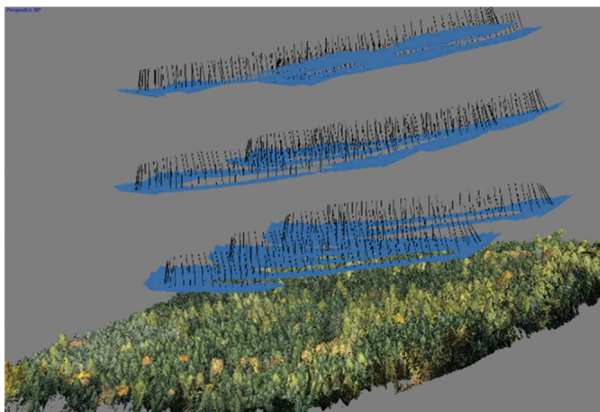


Figure 4. Flight performance for image acquisition at three different elevations (Plot 1, autopilot mode).

DSM and DTM (0.05 x 0.05 m cell size) as constructed from the total amount of aligned images for each plot using Agisoft PhotoScan®. We initially reconstructed the DTM mesh using automatic classification, based on classified sparse point clouds of the ground (as opposed to the construction of the DSM, which was based on the complete dense point cloud). The automatic classification was based on three parameters: i) maximum angle (deg), ii) max distance (m), and iii) cell size (m). In addition to the extraction of DSM and DTM, we located four reference points for each plot. GCPs distributed within the study areas (mainly at larger gap areas), measured with geodetic RTK GPS with centimetre accuracy, whereas the coordinate system was set to S-JTSK/Krovak East North (a local coordinate system that is mainly used in the Czech Republic).

The imagery and the synchronized on-board GPS position for each single image were used as input for image positions. During the alignment process, when the algorithm is trying to find and match points between overlapping images and refine the image positions for each photograph separately, we set the accuracy to high.

2.5. Calculation of tree heights

The setting parameters for local maxima were determined using the morphological filter-focal statistics tool in ArcGIS. This morphological filter is preferred when we need to identify the highest pixel value on the treetop from the CHM; therefore, it can be used to eliminate the possibility of having multiple local maxima within a single crown area. For individual tree position detection from the UAV platform, we used the adaptive filtering method based on CHM height values. In this method, an image-smoothing step was applied to the CHMs using low-pass filters to reduce the noise effect and regulate the values of the smoothing window as a stepwise function of the heights of the CHM (Dralle and Rudemo 1996; Pitkänen et al. 2004). Tree height was derived from the CHM based on the evaluation of local maxima, which were considered to be treetops (Figure 5(a)). The CHM was calculated by subtracting the DSM from the DTM using ArcGIS 10.3.1 by ESRI ©. This decision is usually based on the particular geometry of the forest structure and the spatial pattern of the targeted trees. Prior to any further calculations, and depending on the structure of tree crowns and resolution, values

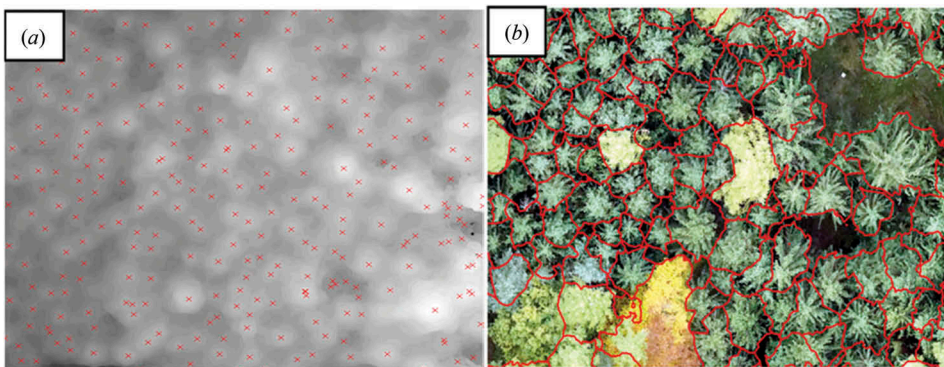


Figure 5. Example to illustrate (a) local maxima seeds identified as tree tops, and (b) tree crowns derived from the process of the IWS algorithm.

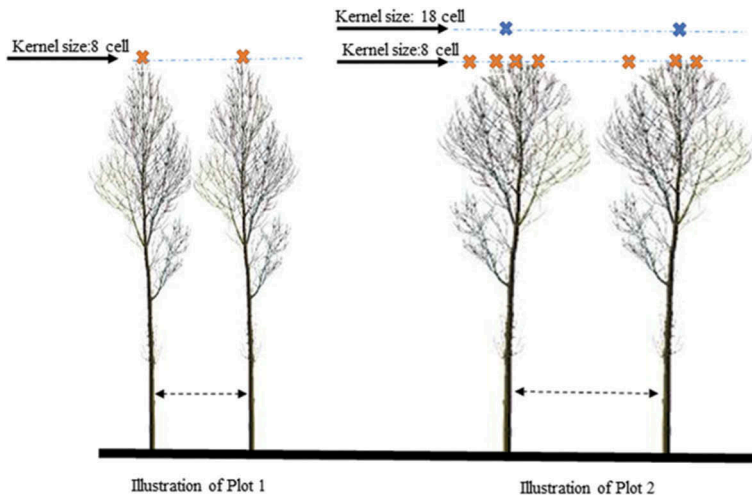


Figure 6. Example to illustrate the potential problems of identifying tree heights using different kernel radius sizes in forests with different tree shapes and structure for Plots 1 and 2.

were approximated to the nearest integer number of pixels. Among the several processing types we tested, in different variances of radius by using circular-shaped areas, we obtain the best results by setting a radius of eight-cell units within a kernel range $rm \in \{1, 2, 3, 4, 6, 8, \dots, 18\}$ for Plot 1 and 18-cell radius for Plot 2, respectively. We hypothesized that different kernel radii were necessary for plots because the canopy structure of the trees of plot 2 was more circular; when we applied the same eight-cell kernel radius, similar to that in Plot 1, multiple pixels were identified as a treetop; thus, we increased the filtering value to identify a single treetop. Plot 2 trees were older and taller compared with those in Plot 1, which produced more oval/spherical shapes, unlike Plot 1 trees, which had more conical crown shapes (see Figure 6).

Focal statistics is a filtering tool that we used to identify the treetop location, based on the maximum pixel value in each kernel. Then we extracted the common pixel values, derived from the process of the focal statistics tool together with the CHM, to determine the positions of the highest pixel values (tree positions). For each input cell location, focal statistics calculate a statistic of the values in the form of weights within a specified neighbourhood around it; it returns Mean, Standard Deviation, or Sum values as needed. To match the pixel values from CHM and the focal statistic result, we used equation (1):

$$\text{Con}('CHM' == 'focal statistics result', 1). \quad (1)$$

Conditional tool (Con) performs a conditional if/else evaluation on each of the input cells of an input raster and returns back the binary value of 0 (for non-data) or 1 (for data value). The return value was the value when the CHM value equalled the focal statistics output (raster).

When we tried to match the efficiency of tree positions obtained from the UAV platform with the field measurements, we detected a slight spatial inaccuracy (tilted positions of trees resulted from the UAV) caused by the lack of precision of the on-board

GPS. To solve this problem, we used the azimuth and horizontal distances, measured in the field to identify the X, Y tree positions. For each plot, we georeferenced four tree point positions, in order to match the tree positions acquired from the UAV with those of the terrestrial data.

2.6. Calculation of crown diameters

For individual tree canopy delineation, we used the method of IWS, as proposed by Edson and Wing (2011). We used the IWS method in ArcGIS 10.3.1 (ESRI) mainly because of its ability to delineate distinct tree entities, especially in relatively closed forest canopy structures. To implement the IWS, we inverted the CHM raster surface, whereby treetops become 'ponds' and tree branches/crowns become watersheds. Using the inverted CHM, we then created the flow direction layer to create individual hydrologic drainage basins (Wannasiri et al. 2013), which identified the crown delineations in the form of a raster layer. However, this layer contained large gaps located between tree crowns. To identify the tree crowns for delineation, we used the Boolean layer of the CHM to reclassify the new output into two different categories. For Plot 1, the first category identified all of the pixels that had values greater than 17 m (pixel values = 1) and the second category identified all pixel values less than 17 m, including gap areas (values = 0). We determined this threshold value by testing a range of different height values (14–19 m); this range was based on the shortest trees in the plot and the ground field height measurements. Plot 2 had taller trees; thus, we had to increase the threshold value to 20 m. To identify all the gap areas in the plots, we multiplied the result of the Boolean layers with the basin layer.

Finally, we converted the inverted CHM to polygons, to provide a layer for measuring the crown diameters, as shown in Figure 5(b). Shapefiles were exported as polygons to delineate tree crowns; we used the QGIS V 2.12.0 (©Free-Open Source Software Foundation, Inc., Boston, MA, USA) package to measure the crown diameters. We established the centre for each polygon and we converted the polygons to lines and then line to points to assign the points to the periphery of each polygon; we then used the distance tool to find the diameter of the crown. To preserve the integrity of our data, all GIS measurements of the crown diameters were performed in the same manner as the field data collection when applicable, e.g. orientation of diameter measurements (north–south and east–west).

2.7. Statistics and validation of data

To compare the estimates of tree height and crown diameter from UAV flights with field measurements, we used a tree-level approach. All statistical analyses were conducted in MATLAB (MathWorks®), Inc. and Excel (Microsoft® office). We used linear regression analysis to model the relationship between the estimated and measured variables and we calculated the R-squared as a metric for accuracy. Additionally, box-and-whisker plots were used to illustrate the variance for the measured and estimated variables. To determine whether the estimated heights and crown diameters were significantly different from the measured values, we compared the mean of differences using confidence intervals and paired *t*-tests. We also calculated the residuals as the difference

between individual tree parameters estimated from the CHM and the field measurements. As a better illustration for the distribution of error around the mean of data (how close predictions are to the eventual outcomes), we also computed the mean absolute error (MAE) and then we evaluated the residuals for each plot for both measured and estimated variables. The MAE is given by Equation (2) as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| = \frac{1}{n} \sum_{i=1}^n |e_i|. \quad (2)$$

As the name suggests, the MAE is an average of the absolute errors, $|e_i| = |f_i - y_i|$ where f_i is the prediction and y_i is the true value.

3. Results

3.1. Tree height and crown estimation

Plot 1 had 49 trees and Plot 2 had 39 trees; the height measures (terrestrial data) and crown diameter estimations (UAV data) of each plot are summarized in Table 1. In Plot 1, the median and mean height values indicate UAV data underestimated tree heights (see Table 1) and overestimated crown diameters; in Plot 2, both tree heights and crown diameters were overestimated.

Linear regression (see Figure 7) exhibited a strong relationship between the estimated and measured tree heights in Plot 1 ($R^2 = 0.75$) and Plot 2 ($R^2 = 0.72$). Similarly, the estimated crown diameters in Plot 1 ($R^2 = 0.63$) and Plot 2 ($R^2 = 0.85$) also showed good fit with the estimated values (see Figure 10). In general, taking a better look (see Figure 8; Figure 11) for both tree variables and both study areas and in accordance with the above results, we can point out that we have wider ranges of variation of mean values between the estimated values compared with the terrestrial measurements in case of height. On the contrary, exception is the means of crown variable for both plots (see Table 3).

To improve the regression model, we also calculated the residuals. More specifically, in Plot 1, the MAE of height was 2.62 m, and the height accuracy was lower in Plot 2 with MAE of height equal to 2.88 m (see Figure 9). For crown diameter, the results were similar to the higher accuracy in Plot 1 (MAE = 0.73 m) compared with Plot 2 (MAE = 0.80 m) (see Figure 12). The estimated and measured tree heights and crown diameters in Plot 1 had a mean difference of 1.55 m and -0.42 m, respectively, whereas in Plot 2 the mean differences were -2.35 m and -0.70 m, respectively (see Figure 13; see Table 5).

Table 1. Statistics of the measured and estimated height variables for the two study areas at the tree level.

	Plot 1		Plot 2	
	Measured tree height (m)	Estimated height (UAV) (m)	Measured tree height (m)	Estimated height (UAV) (m)
H min	13.3	16.2	17.9	20.4
H Q1	21.3	20.8	24.5	27.5
H median	24.2	22.5	26.7	29.9
H Q3	28.1	24.1	30.1	31.2
H max	34.4	30.1	32.7	34.8
H mean	24.2	22.3	27.0	29.4

CD: crown diameter (min, max, and median values); Q: quartiles (Q1:25% (lower) – 50% (median) – Q3: 25% (upper)).

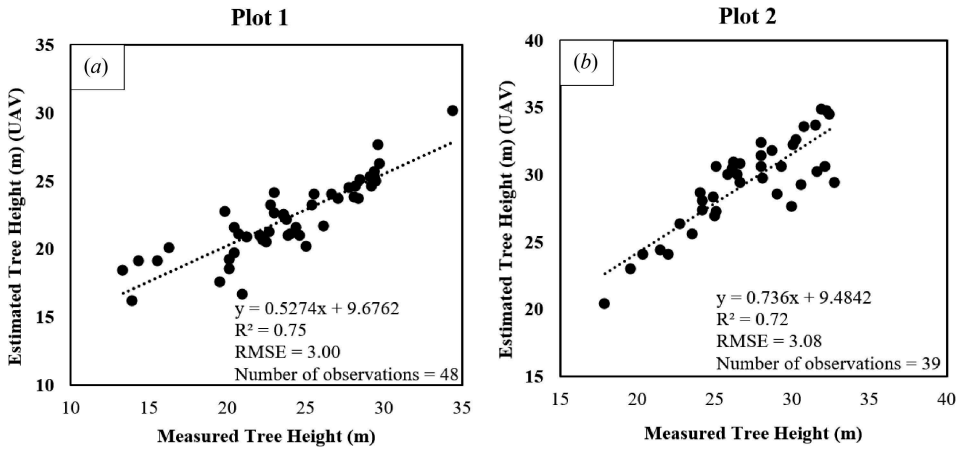


Figure 7. Linear regression models of the estimated (UAV) and measured (ground truth) heights for Plot 1 (a) and Plot 2 (b).

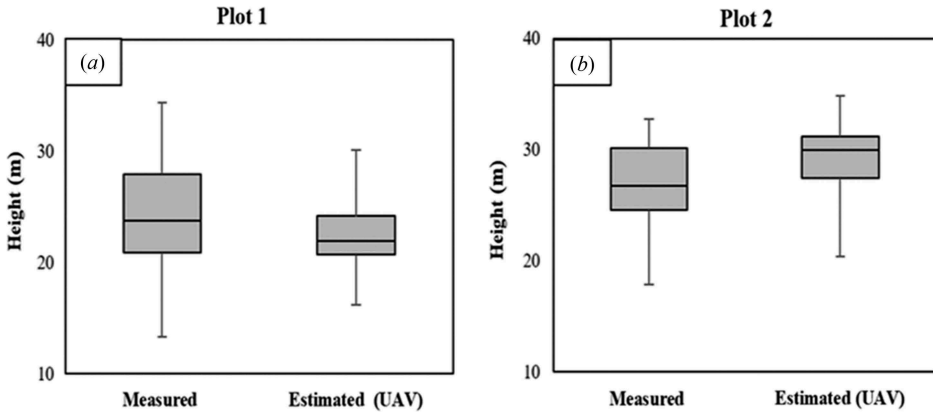


Figure 8. Box-and-whisker plots of the measured and estimated height variables for both study areas based on 48 trees in Plot 1 (a) and 39 trees in Plot 2 (b).

The root mean square error (RMSE) (m) for height (see Table 2) in Plot 1 was 3.00 m and it was 3.08 m in Plot 2, whereas the crown projection RMSE (m) in Plot 1, 0.82 m, was obviously lower than the RMSE (m) of 1.04 in plot 2 (see Table 4). However, because the plots had different means, we wanted to normalize the RMSE by dividing the RMSE by the mean of the plots to calculate the RMSE% for comparing the accuracy between the plots.

RMSE% for height (see Table 2) and crown diameter (see Table 4) evaluation varied by plot, ranging from 11.42 to 18.56. Particularly, in Plot 1 the estimation of height and crown diameter RMSE% was similar, 12.62 and 14.29, respectively, but for Plot 2 the RMSE% were quite different for height (11.42) and crown diameter (18.56) estimates. In both plots, the estimated and measured values for height and crown diameter were similar, as supported by the R^2 values of the linear regression models.

According to the t -test result, the absolute value of the t -statistics for the estimated heights and crown projections in both plots exceeded the critical values (see Tables 2

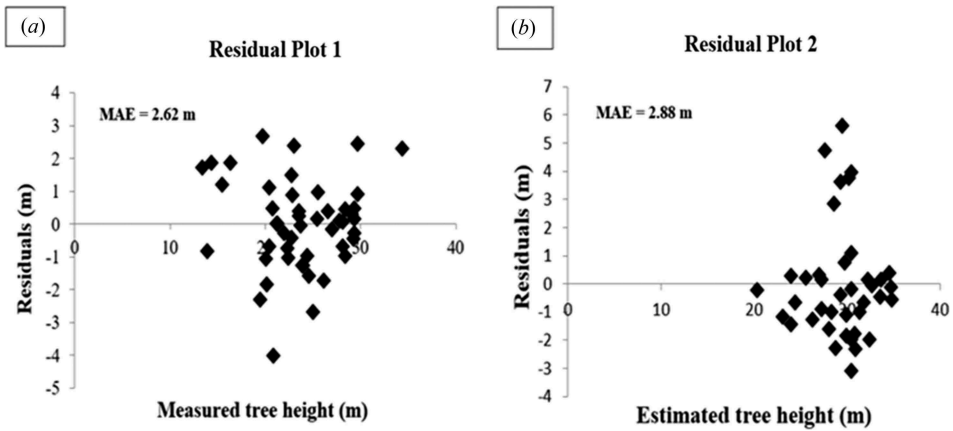


Figure 9. Residual plots displaying the distribution of errors around the mean estimated height for Plot 1 (a) and Plot 2 (b).

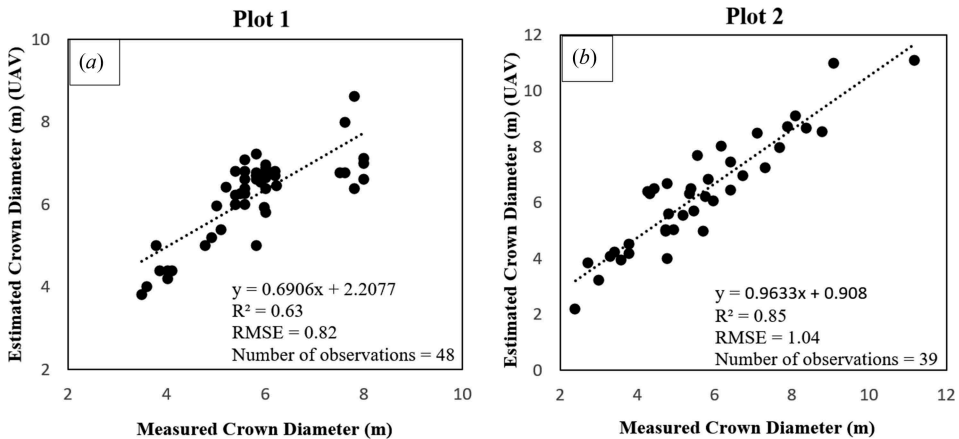


Figure 10. Linear regression models of the estimated (UAV) and measured (ground truth) crown diameters for Plot 1 (a) and Plot 2 (b).

and 4); based on this result, we rejected the null hypothesis and concluded that the two plot means were different at a 0.05 significance level.

4. Discussion

Low-altitude UAV imagery may be used to systematically observe forest canopy height (Baltsavias et al. 2008; Dandois and Ellis 2010). Structural forest attributes are commonly extracted from the CHM by means of regression models to predict tree features for forest inventories (e.g. descriptive statistic of the CHM on a particular area) (Næsset, 2002; Næsset et al. 2004). The practical outcomes of CHMs, in combination with terrestrial measurements, are the prediction of forest attributes of interest, e.g. crown diameter and height estimation.

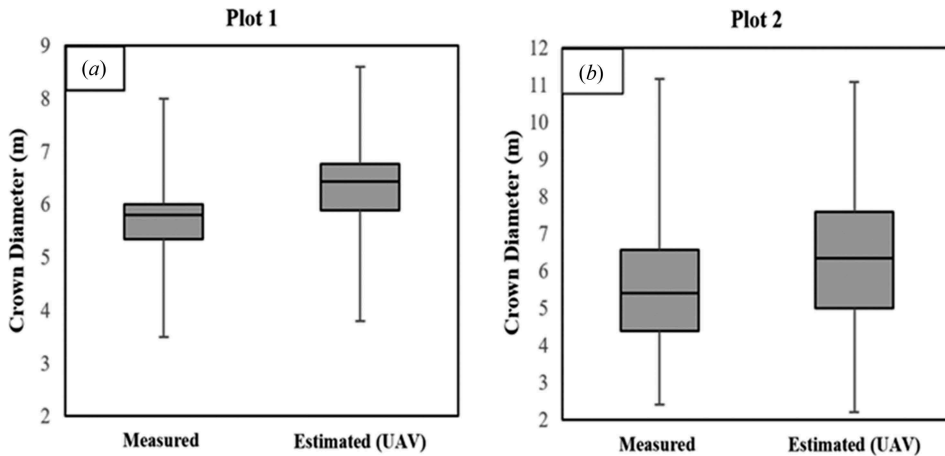


Figure 11. Box-and-whisker plots of both measured and estimated crown diameters for both study areas based on 48 trees in Plot 1 (a) and 39 trees in Plot 2 (b).

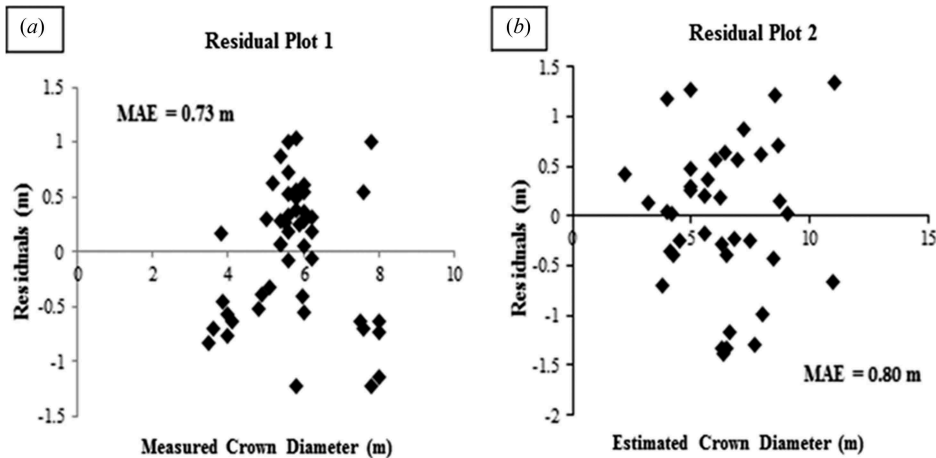


Figure 12. Residual plots displaying the distribution of errors around the mean estimated crown diameter for Plot 1 (a) and Plot 2 (b).

Based on the *t*-test results, we rejected the null hypothesis because the means for both plots were different at a 0.05 significance level. Although the differences between means were significant according to paired *t*-tests, the error of estimates from the UAV was relatively small and therefore considered acceptable. However, low accuracy, especially crown diameter, was observed in both plots (RMSE % = 14.29), but with lower accuracy in Plot 2 (RMSE% = 18.56) (see Table 4). This can be partially explained by the combination of pixels from gap areas with the pixels of ‘real’ crown area, which can cause overestimation of crown diameter (Diaz-Varela et al. 2015).

Absolute RMSE (m) showed a higher agreement for crown diameter than height for both Plot 1 (0.82 m) and Plot 2 (1.04 m). Estimates of height and crown diameter had an RMSE%

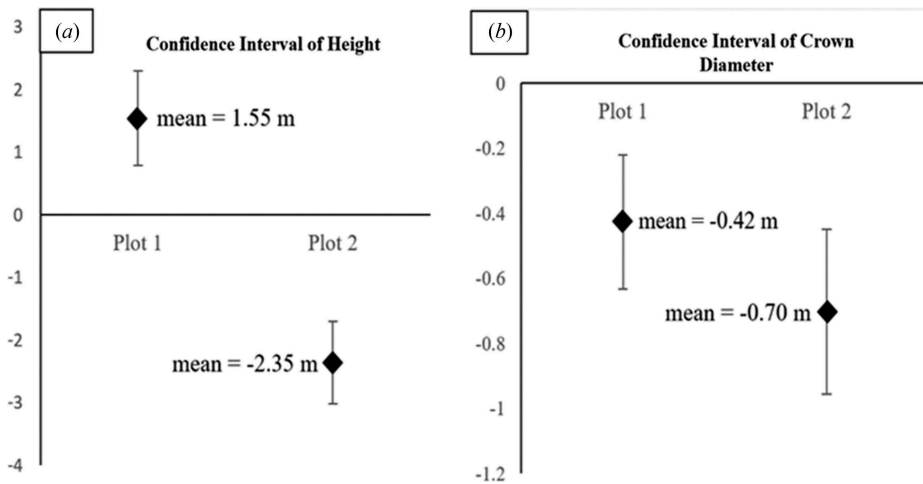


Figure 13. The mean of differences for the estimated–measured tree heights and crown diameters for both plots.

Table 2. Statistical summary of height estimates for Plots 1 and 2.

	RMSE (m)	RMSE %	Bias	Bias %	MAE (m)	P(T ≤ t) two-tail t-test	t Critical two-tail t-test	t Stat	p-Value	df
Plot 1	3.00	12.62	-1.55	-6.99	2.62	0.15	2.01	4.14	0.00	47
Plot 2	3.08	11.42	2.35	8.02	2.88	0.00	2.02	7.27	0.00	38

RMSE: root mean square error; RMSE%: root mean square error percentage; MAE: mean absolute error; bias; bias%: two-tailed *t*-test; *t*-statistic; df: degrees of freedom.

Table 3. Statistics of the measured and estimated crown diameter variables for the two study sites at the tree level.

	Plot 1		Plot 2	
	Measured CD (m)	Estimated CD (UAV) (m)	Measured CD (m)	Estimated CD (UAV) (m)
CD min	3.5	3.8	2.4	2.2
CD Q1	5.4	5.9	4.4	5.0
CD median	5.8	6.4	5.4	6.3
CD Q3	6.0	6.8	6.6	7.6
CD max	8.0	8.6	11.2	11.1
CD mean	5.7	6.2	5.6	6.3

CD: crown diameter (min, max, and median values); Q: quartiles (Q1:25% (lower) – 50% (median) – Q3: 25% (upper)).

Table 4. Statistical summary of crown diameter estimates for Plots 1 and 2.

	RMSE (m)	RMSE %	Bias	Bias%	MAE (m)	P(T ≤ t) two-tail t-test	t Critical two-tail		t Stat	p-Value	df
							t-test	t-test			
Plot 1	0.82	14.29	0.43	6.92	0.73	0.00	2.01	4.17	0.00	47	
Plot 2	1.04	18.56	0.70	11.11	0.80	0.00	2.02	5.61	0.00	38	

RMSE: root mean square error; RMSE%: root mean square error percentage; MAE: mean absolute error; bias; bias%, two-tailed *t*-test; *t*-statistic; df: degrees of freedom.

lower than 20 for most results. Zarco-Tejada et al. (2014) compared the height data retrieved from high-resolution images and ground measurements; similarly, they determined that the RMSE% ranged between 10 and 13, depending on the study area. In Plot 2, where the

Table 5. Statistical summary table displays the confidence interval of mean value of differences, standard deviation, the size of the sample plots (total number of observations), the confidence interval as an estimated range of values, and the two-sided confidence limits for both plots.

	Number of Plots	Mean (m)	Standard Dev. (m)	Sample Size	Confidence Interval (m)	Confidence Range (m)	
						Upper Bound	Lower Bound
Height	1	1.55	2.59	48	2.04	0.75	0.75
	2	-2.35	2.02	39	2.02	0.65	0.65
Crown Diameter	1	-0.42	0.71	48	2.04	0.21	0.21
	2	-0.70	0.78	39	2.02	0.25	0.25

discontinuous canopy yielded a higher percentage of gap areas, RMSE% for crown diameter was 18.56, which is still acceptable when compared with errors that may occur with field measurements (Diaz-Varela et al. 2015; Paulus et al. 2014; Paproki et al. 2012).

Height and crown diameter estimates for Plot 1 were closer to the measured variables (see Table 5; see Figure 13). In addition, the confidence intervals indicated that the mean of the estimated crown diameters for both plots was significantly more similar than the estimated height variables, which were significantly different (underestimation of heights for Plot 1 and overestimation of height for Plot 2). These findings may be attributable to the differences in heights of trees (mean Plot 1 = 24.2 m, mean Plot 2 = 27 m) and similarity of crown diameters (mean Plot 1 = 5.7 m, mean Plot 2 = 5.6 m) in the two plots.

Although the proposed approach caused some systematic errors related to image smoothing, it can result in low and acceptable errors, which can be useful for monitoring the variation of forest attributes relevant to forest management (Diaz-Varela et al. 2015).

5. Conclusions

In this study, we tested the performance of tree detection algorithms based on UAV SfM point clouds, and we assessed the accuracy of tree height and diameter measurements derived from both point clouds. We reconstructed 3D forest canopy models based on SfM techniques to evaluate tree height and crown diameter. The proposed workflow consists of UAV image acquisition data from an RGB camera, processing of the point cloud, construction of the CHM from DSM and DEM, and GIS processing-analyses and watershed algorithm application for the calculation of tree variables. Many studies have previously treated estimates of tree parameters, such as tree crown delineation and treetop detection, as two separate procedures (Wang, Gong, and Biging 2004). In fact, tree crown boundaries and treetop detection should be considered as continuous processes; the derivation of one aspect can assist in the solution of the other. However, these assumptions have some limitations and can only be applied to forest structures with more homogeneous crown structures, especially with large crowns, because otherwise several local maxima can be found for a single crown surface. In addition, the location of the treetop is the basis for extracting tree height and delineating the crown projection area. We found that the choice of filters for smoothing the CHM has a significant influence on the detection of treetops. Therefore, the decision on how to determine the treetop location is critical and it is largely based on the knowledge and experience we have about the structure of the forest. It is also worth noting the fact that estimations from photos tend to show a slightly higher dispersion of data,

which could be associated with image artefacts related to data acquisition, errors during processing, or human factors during ground measurements. In fact, the nadir view of the tree crown, taken from remote sensors in the height estimation, could account for tree irregularities, which may go unidentified by individuals taking ground measurements (Moorthy et al. 2011).

To potentially improve the results of the 3D image reconstruction model and ensure the integrity of the results based on CHM, we used four ground reference points, measured with RTK GPS. In this study, instead of using a more sophisticated approach for treetop estimates, we simply tested and used threshold values applied on morphological filters to remove non-treetop maxima. Concerning the estimation of crown diameters, large gaps between the trees within our study area produced an inaccurate delineation of solitary trees using IWS. To correct the model and leave only the tree crowns for delineation, we excluded the areas with elevation lower than 17 m in the layer of CHM for Plot 1, and 20 m in the same layer for Plot 2 (because the shortest trees found in Plot 1 and Plot 2 were 17 m and 20 m, respectively); this produced the best fitted delineation for our trees. Pixels with high elevations may be the result of bushes or leaf present in the gap areas, a factor that can result in the misidentification of tree crowns.

As a general conclusion regarding the performance of remote sensing versus field measurements, we can say that the comparison between reference field measurements and remote-sensing estimation of crown parameters confirmed the performance of the workflow applied as a quick and effective alternative to characterize forest tree crown diameters. Finally, high accuracy of height and crown diameter estimates was produced for both plots; based on the RMSE% values comparing remote-sensing and field techniques, crown diameter appeared to be of lower accuracy for both plots, but the results were considered satisfactory.

Acknowledgements

This research was funded by the Ministry of Finance of the Czech Republic, grant of Iceland, Liechtenstein and Norway 'Frameworks and possibilities of forest adaptation measures and strategies connected with climate change' No. EHP-CZ02- OV-1-019-2014, and the Internal Grant Agency (IGA) of the Faculty of Forestry and Wood Sciences at the Czech University of Life Sciences in Prague (No. B07/15). The authors would also like to acknowledge Dr Přemysl Janata of the Mendel University of Brno for his support and participation in the field campaign and measurements.

Disclosure statement

The authors declare no conflict of interest.

Funding

This work was supported by the Ministry of Finance of the Czech Republic, grant of Iceland, Liechtenstein and Norway; Frameworks and possibilities of forest adaptation measures and strategies connected with climate change [EHP-CZ02- OV-1-019-2014]; Internal Grant Agency; Forestry and Wood Sciences at the Czech University of Life Sciences Prague [B07/15].

Author contributions

Dimitrios Panagiotidis, Azadeh Abdollahnejad, and Vasco Chiteculo collected the ground measurements. Dimitrios Panagiotidis and Azadeh Abdollahnejad wrote the manuscript, analysed the images, and made the data analyses and interpretations. Vasco Chiteculo elaborated and adjusted the figures and Peter Surový supervised the manuscript.

References

- Baltsavias, E., A. Gruen, H. Eisenbeiss, L. Zhang, and L. T. Waser. 2008. "High-Quality Image Matching and Automated Generation of 3D Tree Models." *International Journal of Remote Sensing* 29 (5): 1243–1259. doi:10.1080/01431160701736513.
- Bodluc, P., K. Lowell, and G. Edwards. 1999. "Automated Estimation of Localized Forest Volumes from Large-Scale Aerial Photographs and Ancillary Cartographic Information in a Boreal Forest." *International Journal of Remote Sensing* 20 (18): 3611–3624. doi:10.1080/014311699211237.
- Bortolot, Z. 2006. "Using Tree Clusters to Derive Forest Properties from Small Footprint Lidar Data." *Photogrammetric Engineering & Remote Sensing* 72: 1389–1397. doi:10.14358/PERS.72.12.1389.
- Carleer, A. P., O. Debeir, and E. Wolff. 2005. "Assessment of Very High Spatial Resolution Satellite Image Segmentations." *Photogrammetric Engineering & Remote Sensing* 71 (11): 1285–1294. doi:10.14358/PERS.71.11.1285.
- Chen, Q., D. D. Baldocchi, P. Gong, and M. Kelly. 2006. "Isolating Individual Trees in a Savanna Woodland Using Small Footprint Lidar Data." *Photogrammetric Engineering & Remote Sensing* 72: 923–932. doi:10.14358/PERS.72.8.923.
- Dandois, J. P., and E. C. Ellis. 2010. "Remote Sensing of Vegetation Structure Using Computer Vision." *Remote Sensing* 2 (4): 1157–1176. doi:10.3390/rs2041157.
- Dandois, J. P., and E. C. Ellis. 2013. "High Spatial Resolution Three-Dimensional Mapping of Vegetation Spectral Dynamics Using Computer Vision." *Remote Sensing of Environment* 136 (136): 259–276. doi:10.1016/j.rse.2013.04.005.
- Diaz-Varela, R., R. De La Rosa, L. Leon, P. J. Zarco-Tejada, and J. Pablo. 2015. "High-Resolution Airborne UAV Imagery to Assess Olive Tree Crown Parameters Using 3D Photo Reconstruction: Application in Breeding Trials." *Remote Sensing* 7: 4213–4232. doi:10.3390/rs70404213.
- Diaz-Varela, R. A., P. J. Zarco-Tejada, V. Angileri, and P. Loudjani. 2014. "Automatic Identification of Agricultural Terraces through Object-Oriented Analysis of Very High Resolution Dsms and Multispectral Imagery Obtained from an Unmanned Aerial Vehicle." *Journal of Environmental Management* 134 (134): 117–126. doi:10.1016/j.jenvman.2014.01.006.
- Dralle, K., and M. Rudemo. 1996. "Stem Number Estimation by Kernel Smoothing of Aerial Photos." *Canadian Journal of Forest Research-Revue* 26 (7): 1228–1236. Canadienne De Recherche Forestiere. doi:10.1139/x26-137
- Edson, C., and M. G. Wing. 2011. "Airborne Light Detection and Ranging (Lidar) for Individual Tree Stem Location, Height, and Biomass Measurements." *Remote Sensing* 3: 2494–2528. doi:10.3390/rs3112494.
- Falkowski, M. J., A. M. S. Smith, A. T. Hudak, P. E. Gessler, L. A. Vierling, and N. L. Crookston. 2006. "Automated Estimation of Individual Conifer Tree Height and Crown Diameter via Two-Dimensional Spatial Wavelet Analysis of Lidar Data." *Canadian Journal of Remote Sensing* 32: 153–161. doi:10.5589/m06-005.
- Fritz, A., T. Kattenborn, and B. Koch. 2013. "UAV-Based Photogrammetric Point Clouds—Tree Stem Mapping in Open Stands in Comparison to Terrestrial Laser Scanner Point Clouds." *International Archives of Photogrammetry, Remote Sensing & Spatial Information Sciences* XL-1/W2: 141–146. doi:10.5194/isprsarchives-XL-1-W2-141-2013.
- Gini, R., D. Passoni, L. Pinto, and G. Sona. 2014. "Use of Unmanned Aerial Systems for Multispectral Survey and Tree Classification: A Test in A Park Area of Northern Italy." *European Journal of Remote Sensing* 2014 (47): 251–269. doi:10.5721/EuJRS20144716.

- HAGLÖF SWEDEN, 2002, Vertex III and Transponder T3 Manual (English). Accessed 22 October 2003. <http://www.forestrytools.com.au/VertexIII%20users%20guide.pdf>
- James, M. R., and S. Robson. 2012. "Straight Forward Reconstruction of 3D Surfaces and Topography with a Camera: Accuracy and Geoscience Application." *Journal of Geophysical Research* 2012: 117.
- Kalliovirta, J., and T. Tokola. 2005. "Functions for Estimating Stem Diameter and Tree Age Using Tree Height, Crown Width and Existing Stand Database Information." *Silva Fennica* 39 (2): 227–248. doi:10.14214/sf.386.
- Koch, B., U. Heyder, and H. Weinacker. 2006. "Detection of Individual Tree Crowns in Airborne Lidar Data." *Photogrammetric Engineering & Remote Sensing* 72: 357–363. doi:10.14358/PERS.72.4.357.
- Küng, O., C. Strecha, A. Beyeler, J. C. Zufferey, D. Floreano, P. Fua, and F. Gervais. 2011. "The Accuracy of Automatic Photogrammetric Techniques on Ultra-Light UAV Imagery." *International Archives of Photogrammetry, Remote Sensing & Spatial Information Sciences* XXXVIII-1: 125–130.
- Maltamo, M., K. Eerikainen, J. Pitkaken, J. Hyypä, and M. Vehmas. 2004. "Estimation of Timber Volume and Stem Density Based on Scanning Laser Altimetry and Expected Tree Size Distribution Functions." *Remote Sensing of Environment* 90: 319–330. doi:10.1016/j.rse.2004.01.006.
- Moorthy, I., J. R. Miller, J. A. J. Berni, P. Zarco-Tejada, B. Hu, and J. Chen. 2011. "Field Characterization of Olive (*Olea Europaea* L.) Tree Crown Architecture Using Terrestrial Laser Scanning Data." *Agricultural & Forest Meteorology* 151 (2): 204–214. doi:10.1016/j.agrformet.2010.10.005.
- Naesset, E. 2002. "Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data." *Remote Sensing of Environment* 80 (1): 88–99. doi:10.1016/S0034-4257(01)00290-5.
- Paproki, A., X. Sirault, S. Berry, R. Furbank, and J. Fripp. 2012. "A Novel Mesh Processing Based Technique for 3D Plant Analysis." *BMC Plant Biology* 12 (1): 63. doi:10.1186/1471-2229-12-63.
- Paulus, S., J. Behmann, A. K. Mahlein, L. Plumer, and H. Kuhlmann. 2014. "Low-Cost 3D Systems: Suitable Tools for Plant Phenotyping." *Sensors* 14 (2): 3001–3018. doi:10.3390/s140203001.
- Peckham, R. J., and J. Gyozo. 2007. *Development and Applications in a Policy Support Environment Series: Lecture Notes in Geoinformation and Cartography*. Germany: University of Heidelberg.
- Persson, Å., J. Holmgren, and U. Soderman. 2002. "Detecting and Measuring Individual Trees Using Airborne Laser Scanning." *Photogrammetric Engineering and Remote Sensing* 68 (9): 925–932.
- Pitkänen, J., M. Maltamo, J. Hyypä, and X. Yu. 2004. Adaptive Methods for Individual Tree Detection on Airborne Laser Based Canopy Height Model. In: M. Theis, B. Koch, H. Spiecker, and H. Weinacker, eds. *Proceedings of ISPRS Working Group VIII/2: "Laser-scanners for forest and landscape assessment"*. Germany: University of Freiburg; pp. 187–191.
- Popescu, S. C. 2007. "Estimating Biomass of Individual Pine Trees Using Airborne Lidar." *Biomass and Bioenergy* 31: 646–655. doi:10.1016/j.biombioe.2007.06.022.
- Popescu, S. C., R. H. Wynne, and R. F. Nelson. 2002. "Estimating Plot-Level Tree Heights with Lidar: Local Filtering with a Canopy Height Based Variable Window Size." *Computers and Electronics in Agriculture* 37: 71–95. doi:10.1016/S0168-1699(02)00121-7.
- Popescu, S. C., R. H. Wynne, and R. F. Nelson. 2003. "Measuring Individual Tree Crown Diameter with Lidar and Assessing Its Influence on Estimating Forest Volume and Biomass." *Canadian Journal of Remote Sensing* 29: 564–577. doi:10.5589/m03-027.
- Pyysalo, U., and H. Hyypä. 2002. Reconstructing Tree Crowns from Laser Scanner Data for Feature Extraction. In *ISPRS Commission III, Symposium*, 13 September 2002. Graz, Austria. pp. 218–221.
- Remondino, F., L. Barazzetti, F. Nex, M. Scaioni, and D. Sarazzi. 2011. "UAV Photogrammetry for Mapping and 3D Modeling—Current Status and Future Perspectives. *International Archives of the Photogrammetry.*" *Remote Sensing and Spatial Information Sciences* 38: 25–31.
- Shao, G., S. Zhao, and H. H. Shugart. 1996. *Forest Dynamics Modeling – Preliminary Explanations of Optimizing Management of Korean Pine Forests*, 159. Beijing: Chinese Forestry Publishing House.

- Sona, G., L. Pinto, D. Pagliari, D. Passoni, and R. Gini. 2014. "Experimental Analysis of Different Software Packages for Orientation and Digital Surface Modelling from UAV Images." *Earth Science Informatics* 7: 97–107. doi:[10.1007/s12145-013-0142-2](https://doi.org/10.1007/s12145-013-0142-2).
- Turner, D., A. Lucieer, and C. Watson. 2012. "An Automated Technique for Generating Georectified Mosaics from Ultra-High Resolution Unmanned Aerial Vehicle (UAV) Imagery, Based on Structure from Motion (Sfm) Point Clouds." *Remote Sensing* 4 (12): 1392–1410. doi:[10.3390/rs4051392](https://doi.org/10.3390/rs4051392).
- Wang, L., P. Gong, and G. S. Biging. 2004. "Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery." *Photogrammetric Engineering & Remote Sensing* 70: 351–357. doi:[10.14358/PERS.70.3.351](https://doi.org/10.14358/PERS.70.3.351).
- Wannasiri, W., M. Nagai, K. Honda, P. Santitamont, and P. Miphokasap. 2013. "Extraction of Mangrove Biophysical Parameters Using Airborne LIDAR." *Remote Sensing* 5 (4): 1787–1808. doi:[10.3390/rs5041787](https://doi.org/10.3390/rs5041787).
- Waser, L. T., E. Baltsavias, K. Ecker, H. Eisenbeiss, C. Ginzler, M. Kuchler, P. Thee, and L. Zhang. 2008. "High-Resolution Digital Surface Models (Dsms) for Modelling Fractional Shrub/Tree Cover in a Mire Environment." *International Journal of Remote Sensing* 29 (5): 1261–1276. doi:[10.1080/01431160701736422](https://doi.org/10.1080/01431160701736422).
- Zarco-Tejada, P. J., R. Diaz-Varela, V. Angileri, and P. Loudjani. 2014. "Tree Height Quantification Using Very High Resolution Imagery Acquired from an Unmanned Aerial Vehicle (UAV) and Automatic 3D Photo-Reconstruction Methods." *European Journal of Agronomy* 55 (55): 89–99. doi:[10.1016/j.eja.2014.01.004](https://doi.org/10.1016/j.eja.2014.01.004).





Paper V



Article

Estimation and Extrapolation of Tree Parameters Using Spectral Correlation between UAV and Pléiades Data

Azadeh Abdollahnejad * , Dimitrios Panagiotidis and Peter Surový 

Department of Forest Management, Faculty of Forestry & Wood Sciences, Czech University of Life Sciences (CULS), 165 00 Prague, Czech Republic; panagiotidis@fld.czu.cz (D.P.); surový@fld.czu.cz (P.S.)

* Correspondence: abdollahnejad@fld.czu.cz; Tel.: +420-774-844-679

Received: 20 December 2017; Accepted: 8 February 2018; Published: 11 February 2018

Abstract: The latest technological advances in space-borne imagery have significantly enhanced the acquisition of high-quality data. With the availability of very high-resolution satellites, such as Pléiades, it is now possible to estimate tree parameters at the individual level with high fidelity. Despite innovative advantages on high-precision satellites, data acquisition is not yet available to the public at a reasonable cost. Unmanned aerial vehicles (UAVs) have the practical advantage of data acquisition at a higher spatial resolution than that of satellites. This study is divided into two main parts: (1) we describe the estimation of basic tree attributes, such as tree height, crown diameter, diameter at breast height (DBH), and stem volume derived from UAV data based on structure from motion (SfM) algorithms; and (2) we consider the extrapolation of the UAV data to a larger area, using correlation between satellite and UAV observations as an economically viable approach. Results have shown that UAVs can be used to predict tree characteristics with high accuracy (i.e., crown projection, stem volume, cross-sectional area (CSA), and height). We observed a significant relation between extracted data from UAV and ground data with $R^2 = 0.71$ for stem volume, $R^2 = 0.87$ for height, and $R^2 = 0.60$ for CSA. In addition, our results showed a high linear relation between spectral data from the UAV and the satellite ($R^2 = 0.94$). Overall, the accuracy of the results between UAV and Pléiades was reasonable and showed that the used methods are feasible for extrapolation of extracted data from UAV to larger areas.

Keywords: downscaling; Pléiades imagery; unmanned aerial vehicle; stem volume estimation; remote sensing

1. Introduction

Technological advances in unmanned aerial vehicles (UAVs) have made it feasible to obtain high-resolution imagery and three-dimensional (3D) data for assessing tree attributes and forest monitoring. Methods of data acquisition with remotely-sensed aerial or satellite data at high spatial resolution have partially replaced conventional methods of field measurement for forest inventory purposes [1–4]. Repeated observation from modern satellites, as well as improvements in UAV technology, have contributed significantly to our understanding of the dynamics of complex ecosystems, particularly forests. Accurate quantification of tree basic parameters, such as height, crown diameter, and diameter at breast height (DBH), is essential for decision-making and planning. Modern techniques of remote sensing can provide accurate estimations of tree height and crown area characteristics at the individual level using a series of algorithms [5]. The individual tree identification (IDS) algorithm allows for the estimation of crown diameter [6], and the smoothing of the canopy height model (CHM) using local maxima techniques [7] can provide estimates of individual tree heights.

Volume and above ground biomass (AGB) of small spatial extent areas can be derived with the help of forest inventory data. However, field measurements are typically unbiased, time-consuming, and expensive. For spatially larger areas, modelling of volume and biomass requires the use of remote sensing information for practical purposes. Additionally, remote sensing techniques are able to improve the value of inventoried data with detailed coverage at affordable costs [8,9]. The strength of remote sensing approaches based on very high resolution (VHR) images using aerial imagery or downscaling methods (i.e., calibration of satellite images based on UAVs) is that they allow for the construction of high-quality 3D digital surface models (DSMs) that can be used to estimate several forest tree attributes, such as height, DBH, and crown diameter [3].

Landsat satellites have been used to predict volume and AGB, mainly because of their long-term data record and the favorable compromise between aerial coverage, spectral sensitivity, and spatial resolution [10]. Newer satellites (e.g., GeoEye-1, IKONOS, WorldView-2, and Pléiades) have significantly increased the potential for data acquisition at higher spatial resolutions (i.e., 2 m multispectral). However, satellites are still unable to provide the desired spatial resolution needed for forestry applications; UAVs, on the other hand, can provide higher resolution data and are frequently used to capture aerial images for planning purposes [11].

Only a limited number of studies have used optical VHR sensors for image matching and estimating forest parameters. In previous studies, the IKONOS, Cartosat-1, and Worldview-2 satellites were used for estimation of structural (e.g., height, DBH, stem volume) and textural (i.e., composition, AGB) metrics with acceptable accuracy [12–15].

However, a review of past studies found that few researchers have used the Pléiades satellite for forestry purposes [16] and no studies used it for AGB estimation. While previous studies using Pléiades emphasized combining spectral derivatives and textural metrics to image matched height metrics, these studies utilized combinations of either height and textural metrics or height metrics and spectral derivatives to improve the remote sensing estimation of forest parameters.

The main objectives of this study were to: (i) evaluate UAV performance to estimate key individual tree parameters, including height, crown diameter, cross-sectional area (CSA), and stem volume; and (ii) to extrapolate the extracted UAV data to larger areas, based on spectral correlation between UAV and satellite, as a practical method for detailed information collection across a large area that would not be economically feasible with ground-based assessments.

2. Materials and Methods

2.1. Study Area

The Doksy territory lies on the shores of Lake Mácha in Northern Bohemia in the Czech Republic (Figure 1). The lake is largely surrounded by dense forests covering an area of 300 km². Geologically, the area is characterized by sandstone pseudokarst in the late stages of development, and the soils are either sandy or peaty, with shallow, peaty basins over rocky sandstone hummocks and sporadic volcanic hills. We selected three evenly-aged (managed stands) 40 × 40 m experimental plots and sampled all trees in each plot to estimate the tree heights, DBH, crown diameters, and stem volume. All plots were located northeast of the city of Doksy (−718000, −991250 NW to −717000, −991950 SE in the local coordinate system S-JTSK/Krovak East North, Figure 1). The research was carried out in a 140-year-old *Pinus sylvestris* L. (Scots pine) monoculture natural stand established on sandy soils (68%). The vegetative period tends to be rather warm and dry. The mean annual air temperature is 7.3 °C and the average annual maximum temperature is 31.5 °C. The mean annual precipitation is 635 mm, with only 354 mm during the growing season.

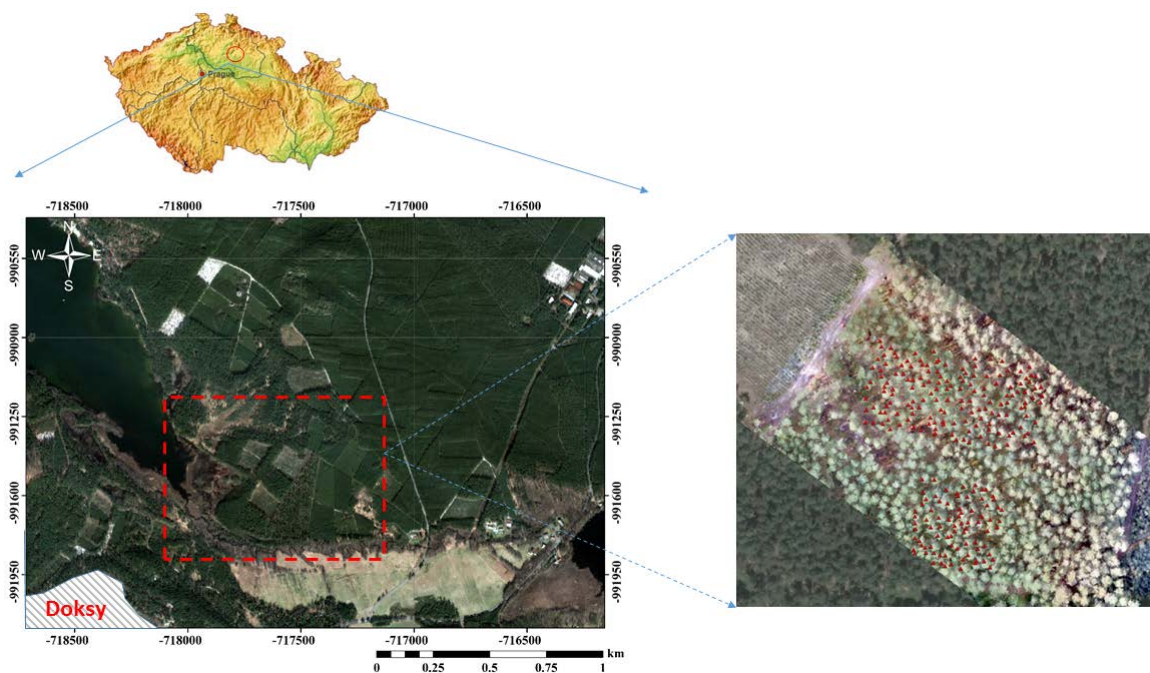


Figure 1. Location of the study area in the northern part of the Czech Republic in the local coordinate system S-JTSK/Krovak East North. Source: Google Earth (elaborated work in ArcGIS for higher map detail).

2.2. Field Survey

For the acquisition of field measurements, we used field-map technology. Field-map is a software and hardware product which combines flexible real-time geographic information system (GIS) software field-map with electronic equipment of high accuracy for mapping and dendrometric measurement. For this study, the dendrometric device which we used was the TruPulse 360/B laser range finder by Inc. Centennial, Colorado, USA. For the purpose of this study, we used the TruPulse 360/B to measure (i) the positions of individual trees by distance based on the field triangulation approach, (ii) tree heights based on the functionality of the digital relascope, and (iii) the mean of the horizontal widths of the tree crown radius in four orientations (east, west, north, and south). The coordinate system was set to S-JTSK/Krovak East North (a local coordinate system that is mainly used in the vicinity of the Czech Republic). The whole data collection process is rather ergonomic since mapped data can be seen simultaneously in the monitor of the computer in the form of layers (either point, lines, or polygons). Each layer can have a number of attributes (i.e., tree positions, heights, etc.) which are stored in a fully relational database.

In addition, the DBH of trees (1.3 m above ground) was measured by a Häglof digital caliper (Häglof Sweden AB, Långsele, Sweden) in two perpendicular directions. Diameters were then transferred via bluetooth to the field-map computer and stored in the relative database. We considered the mean of measured diameters at breast height for computing the CSA. Finally, all the data was extracted by connecting the field-mapper with a universal serial bus (USB) to the personal computer (PC) for further processing. All field measurements were taken in November, 2015. In total, we sampled 223 live trees from all three plots.

2.3. Aerial and Satellite Imagery

The UAV data were acquired in March 2016. The platform model used was an octocopter SteadiDrone EI8HT, ready to fly (RTF) embedded with an RGB (red, green, blue) high-resolution camera. The camera was a Sony Alpha 6000 with an adjusted focal length of 25 mm. The octocopter

needed approximately 7 min to complete a flight for each plot based on the predefined parameters (e.g., the number of waypoints) and the flight mode was set to semi-automatic. The flight path lines covered the entire study area and produced a set of images of the area. The octocopter was guided by a DJI (Dà-Jiāng Innovations Science and Technology Company, Shen Zhen, China) ground station, which is a global positioning system (GPS) flight planning and waypoint-based autopilot software. We performed three flights in total, one flight per plot, at a height of approximately 70 m above the ground with 80% frontal overlap and 70% side overlap.

In order to improve the accuracy of the 3D model, we also set up four ground control points (GCPs) randomly distributed within each plot; these points were measured using the Leica real-time kinematic (RTK) system, model RX1250XC with centimeter accuracy. Due to the low image quality, four of the 596 original images were excluded from the alignment process. During the alignment process, we set the accuracy to high for optimization of the final 3D model. We used Agisoft Photoscan® software (V 1.2.6, St. Petersburg, Russia) to construct the digital terrain model (DTM) and DSM from the 3D model with a cell size of 0.01×0.01 m. The reconstructed mesh of the 3D model was based on automatic classification on certain point classes through the triangulated irregular network (TIN) method. Due to the relatively open canopy in large parts of the study area, small bushes were often abundant in the understory, and, therefore, we classified them as ground points. For setting the parameters for the automatic classification, due to the presence of small bushes near the trees, we decreased the maximum angle from 15 (default value) to 11, the maximum distance from 4 to 1.5 m, while the cell size remained the same. All of the processing was conducted by one computer operator using an Intel® Core™ i7-6700K with a base clock of 4 GHz and 32 GB random access memory (RAM) running with the Windows 10 Professional Edition 64-bit operating system.

We used the Pléiades 1A satellite (launched 16 December 2012) to acquire the space-borne image. The image was taken 27 March 2016, and it had 20 bits/pixel dynamic range of acquisition. For this study, we used one frame with a total area of 25 km^2 ($5 \times 5 \text{ km}$). The image consisted of four multispectral bands: RGB, infrared (IR), and one panchromatic (PAN), as can be seen in Table 1. We used six GCPs for georeferencing the satellite image. For point acquisition, we used GPS RTK Leica model RX1250XC with a maximum error of two centimeters. The RTK correction was carried out by using a base/rover set, which sends and receives fast-rate over-the-air RTK data corrections, using the PDLGFU15 radio module.

Table 1. Pleiades-1A satellite sensor characteristics [17].

Imagery Products	Panchromatic: 50-cm resolution, black and white
	2-m multispectral (RGB—red, green, blue) Bundle: 50-cm black and white and 2-m multispectral
Spectral Bands	Panchromatic: 480–830 nm Blue: 430–550 nm Green: 490–610 nm Red: 600–720 nm Near Infrared: 750–950 nm
Image Location Accuracy	With ground control points: 1 m Without ground control points: 3 m (CE90)

Also, dark object correction was used to derive atmospheric optical information for radiometric normalization using the minimum digital number (DN) value of satellite images = water.

Normalized values of each band (RGB) from the UAV and satellite were then used to compare the spectral data (DN) between the UAV and Pléiades bands (separately) at the individual tree level using Equation (1):

$$Z_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

where Z_i describes the normalized data between 0 and 1, x_i describes the spectral data (for both the UAV and satellite), and x_{\max} and x_{\min} are the maximum and minimum value for each band, respectively. In addition, using the nearest neighbor method, we resampled all three multi-spectral bands from the satellite from 2 m to 1 cm to assign more weight to pixels that cover more crown area (Figure 2). For extracting the spectral data (UAV and satellite), we used zonal statistics in ArcGIS desktop V.10.4.1 (ESRI Inc., Redlands, CA, USA), with the crown area for each individual tree as the zonal layer. We considered the same weight for averaging the DN values of pixels within the zonal layer.

To evaluate the greenness of the detected trees from the UAV and eliminate the dry trees and gap areas, the normalized difference vegetation index (NDVI) was used as a detector index (Equation (2)). This index is usually used to determine the visible spectral response by defining the ratio of greenness per individual tree applied to satellite data [18]:

$$\text{NDVI} = \frac{\text{NIR} - R}{\text{NIR} + R} \quad (2)$$

where *NIR* stands for near-infrared and *R* refers to the red band.

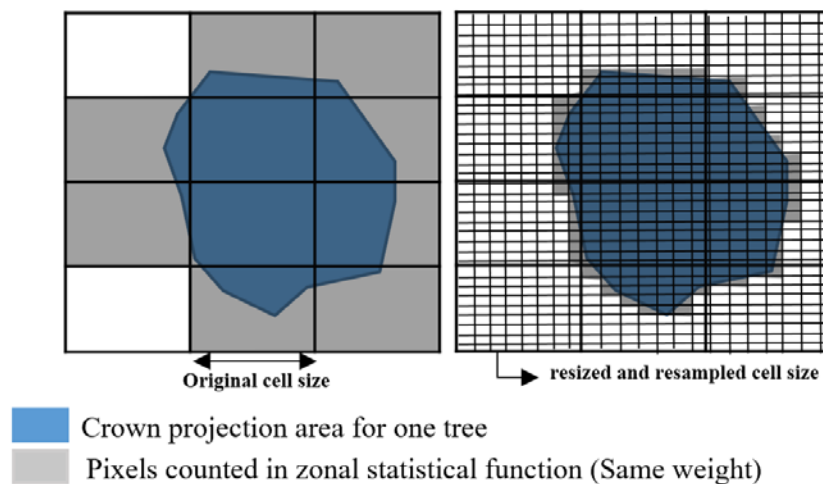


Figure 2. Example of resampled pixels, from 2 m to 1 cm.

2.4. Estimation of Height

To extract the height from the UAV, we computed the CHM, which was derived from the subtraction of the DSM from the DTM. The local maxima algorithm was then used to estimate the height; this algorithm enhances the maximum value within a specified kernel size. As a first step, we used the focal statistics tool in ArcGIS to identify the highest pixel value using the CHM as the input data layer. We performed a low-pass filter to reduce the noise effect and regulate the values of the smoothing window [19]. Among the several processing types we tested in different variances of radius using circular-shaped areas, the best results were at a kernel size with a radius of 1 m based on the average crown diameter derived from the ground measurements (Figure 3). For matching the pixel values, we used the conditional if/else statement on each of the input cells of CHM and focal statistics results, by entering the following command “Con (“CHM” = “focal statistics result”, 1)” using the ArcGIS V. 10.4.1 (ESRI, Redlands, CA, United States) raster calculator.

This conditional tool performs an if/else statement on each input cell and it returns a binary layer with a value of zero assigned as no data and a value of one for data. Finally, the return value was the value when the CHM value equaled the focal statistics output.



Figure 3. The applied method of local maxima seeds for the derivation of tree heights (red crosses show the detected treetops by the local maxima approach).

2.5. Estimation of Crown Projection

For the extraction of the crown area, we used the method of inverse watershed segmentation (IWS), as proposed by Panagiotidis et al. [20]. To implement the IWS, we multiplied the CHM model by -1 so that treetops would appear as ponds and crowns as watersheds. Then, we created a flow direction layer for creating hydrological drainage basins [21]. We applied the ArcGIS reclassify tool on the CHM to create a Boolean layer comprised of two categories; we used a threshold value of 15 m as the classification delimiter. Essentially, that means that pixels with values above 15 m were assigned with a value of 1, and those that were lower than 15 m were assigned a value of 0. Afterwards, the CHM was converted to polygons; the center of each polygon was identified, the polygons were converted to lines, and the lines to points. This allowed us to assign points to the periphery of each polygon. Finally, we used the ArcGIS “zonal geometry as table” tool to define individual tree crown diameters.

2.6. Estimation of Cross-Sectional Area and Stem Volume

To estimate the CSA and volume of individual trees, the coefficient of determination (R^2) between the crown diameter and DBH was calculated. The formula of this coefficient was then used to measure the CSA, by replacing the measured crown size (ground) with the crown size extracted from the UAV. Once we had defined the height and CSA, we were able to estimate the stem volume from the UAV. Additionally, for modelling the stem volume, we considered the shape of tree stems as cylinders.

2.7. Extrapolation of Tree Attributes

We were able to calculate calibrated equations that can be used to determine tree attributes such as CSA and stem volume directly using the spectral information from Pléiades for each tree based on the following equations: linear relation between the UAV spectral data and tree parameters such as CSA and volume (Equation (3)) and linear relation between the UAV and Pléiades spectral data (Equation (4)). In details:

$$CSA = AX_1 + B \quad (3)$$

$$X_1 = CX_2 + D \quad (4)$$

where CSA is cross-section area, X_1 is UAV spectral data, and X_2 is satellite spectral data.

By replacing X_1 in Equation (3) with the Equation number (4), we will be able to measure the CSA directly by satellite spectral data using the below equation:

$$CSA = A(CX_2 + D) + B \quad (5)$$

2.8. Statistical Evaluation and Validation of Data

All statistical analyses were conducted in IBM SPSS V.24 (64-bit 2016) and Excel (Microsoft® Office). The linear regression was used to study correlation between the ground data and tree parameters predicted by the UAV.

Pearson correlation coefficient was computed to analyze the relationships between the spectral values of the UAV RGB bands and tree parameters derived from ground inventory and UAV in two different probability values ($p < 0.05$ and $p < 0.01$). In this study, due to the lack of an IR band in the UAV approach, we computed the vegetation index (VI) [22], green-red vegetation index (GRVI) [22], and visible atmospherically-resistant index, green (VARI g) [23] using the following equations:

$$VI = \frac{\text{Green}}{\text{Red}} \quad (6)$$

$$GRVI = \frac{(\text{Green} - \text{Red})}{(\text{Green} + \text{Red})} \quad (7)$$

$$VARI\ g = \frac{(\text{Green} - \text{red})}{\text{Green} + \text{Red} - \text{Blue}} \quad (8)$$

3. Results

For the sake of simplicity, we divided the results into three parts. In the first part, we presented an overview of tree characteristics in our study area. In the second part, we mainly focused on the potential of the UAV platform deployed with an RGB camera to act as an accurate, alternative field measurement technique. In the third part, we tried to extrapolate the extracted data from the UAV to a larger forested area based on spectral correlation between Pléiades and the UAV.

3.1. General Evaluation of the Study Area

The evaluation of tree attributes showed similarities between the sample plots. However, there was a difference between the variability of sample plots, whereas plot 1 had a lower amount of variability compared with the other two plots (Table 2).

Table 2. Descriptive statistics of the three plots based on ground survey.

Sample	Index	Crown Projection (m ²)	DBH (m)	CSA (m ²)	Height (m)	Volume (m ³)
Plot 1 N = 74	Mean	15.24	0.28	0.06	21.23	1.30
	Variability	30.61	0.15	0.06	7.00	1.30
	Std.	5.20	0.03	0.01	1.33	0.30
Plot 2 N = 72	Mean	14.71	0.28	0.06	23.03	1.43
	Variability	41.54	0.17	0.08	13.80	2.10
	Std.	7.75	0.04	0.02	2.24	0.46
Plot 3 N = 77	Mean	14.94	0.28	0.06	24.49	1.56
	Variability	28.67	0.17	0.08	8.20	2.13
	Std.	6.45	0.03	0.02	2.00	0.45
Total N = 223	Mean	14.97	0.28	0.06	22.96	1.43
	Variability	41.54	0.18	0.08	13.80	2.21
	Std.	6.51	0.03	0.01	2.32	0.42

Std. = standard deviation; DBH = diameter at breast height; CSA = cross-sectional area.

In addition, the NDVI showed that the mean greenness of trees ranged from 0.3 to 0.55 (Figure 4); this range is associated with shrubs-grasslands and temperate forest land-cover classes [24].

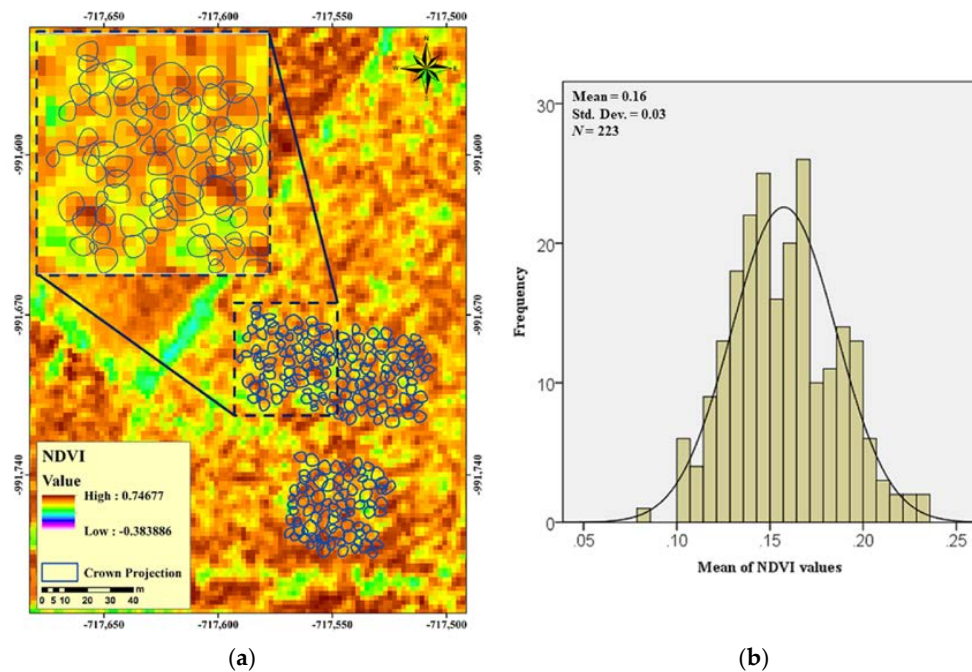


Figure 4. (a) The NDVI (normalized difference vegetation index) calculated for the study area and (b) the histogram of NDVI derived from Pléiades 1A satellite imagery.

3.2. UAV Performance

Linear regression (Figure 5) exhibited a strong relationship in the significant level of $\alpha = 0.05$ ($R^2 = 0.78$) between the cross-section area and crown projection derived from ground data with a root mean square error percent (RMSE%) = 11.35 (Table 3; ID = 1). For the crown projections, a strong relationship ($R^2 = 0.78$; Figure 6a) between the ground and UAV data was observed with an RMSE% = 20.96 (Table 3; ID = 2). Based on these results, we were able to estimate cross-sectional areas from the UAV using adjusted R and RMSE% with an $R^2 = 0.60$ (Figure 6b) and RMSE% = 15.24 (Table 3; ID = 3). Additionally, the results of the height estimation showed strong correlation between the ground and extracted data from the UAV with an $R^2 = 0.87$ (Figure 6c) and RMSE% = 3.73 (Table 3; ID = 4). Finally, we calculated the stem volume based on data from the UAV and the ground. The comparison of the stem volume between the ground and UAV data showed significant correlation with $R^2 = 0.71$ (Figure 6d) and RMSE% = 15.88 (Table 3; ID = 5).

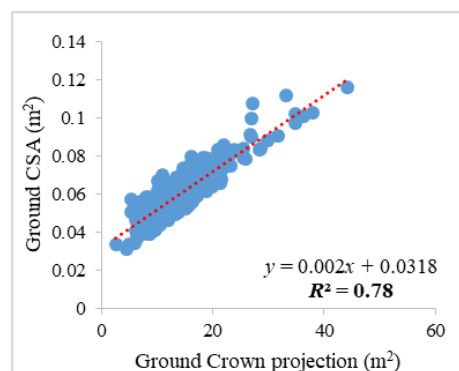


Figure 5. The correlation between crown projection and cross-sectional areas.

Table 3. Statistical summary table, where ID represents the simplicity of identification of each tested regression; N: number of observations; RMSE: root mean square error; and df: degrees of freedom. ID represents the simplicity of identification of each tested regression.

ID	N	RMSE	RMSE%	Bias	Bias %	df	p-Value
1	223	0.0069	11.35	0.0018	2.96	222	0.00
2	223	3.14	20.96	-	-	222	0.00
3	223	0.01	15.24	-	-	222	0.00
4	223	0.86	3.73	-	-	222	0.00
5	223	0.23	15.88	-	-	222	0.00

ID 1 indicates the relation between the ground CSA versus ground crown projection; ID 2 indicates the relation between the crown projection derived from UAV and ground data; ID 3 indicates the relation between the CSA derived from the UAV and ground data; ID 4 indicates the relation between the estimated height from the UAV and ground data; and ID 5 indicates the relation between the stem volume derived from the UAV and ground data.

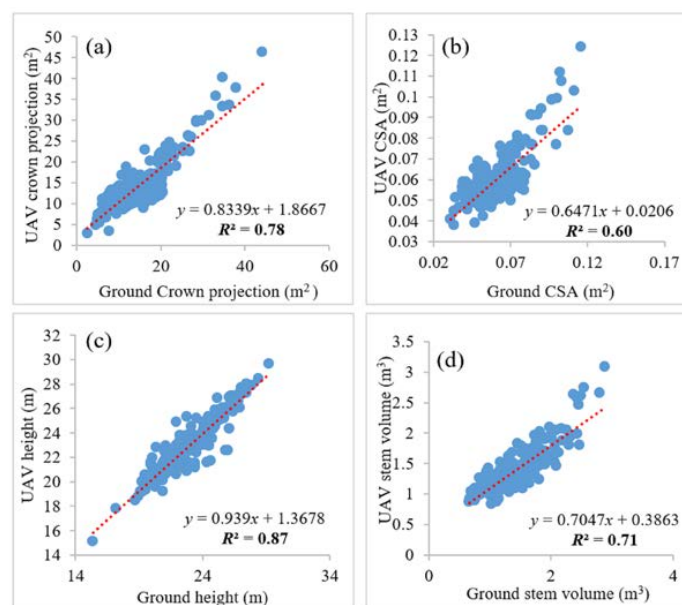


Figure 6. The correlation of (a) crown projection between the ground and UAV (unmanned aerial vehicle) data in m²; (b) cross-sectional areas (CSA) between the ground and UAV data in m²; (c) height between the ground and UAV data in m; and (d) stem volume between the ground and UAV data in m³.

Moreover, the results of the correlation coefficient indicated that the most important descriptive statistic index for all models was the sum of the spectral data as can be seen in Table 4. Finally, our results showed that vegetation indicators that were computed using RGB bands had a significant correlation at the level of $\alpha = 0.01$ with the individual tree stem volume, as can be seen in Table 5.

Table 4. Pearson correlation between the UAV spectral descriptive data and tree attributes at the level of individual trees.

Tree Parameter	Min	Max	Mean	Std.	Sum	Median
CSA (Ground)	−0.208 **	0.133 *	0.224 **	0.160 *	0.864 **	0.229 **
Stem volume (Ground)	−0.200 **	0.131	0.239 **	0.180 **	0.795 **	0.255 **
CSA (UAV)	−0.140 *	0.0641	0.055	0.012	0.821 **	0.049
Stem volume (UAV)	−0.158 *	0.072	0.107	0.05	0.795 **	0.115

** Correlation is significant at the 0.01 level (two-tailed). * Correlation is significant at the 0.05 level (two-tailed).

Table 5. The Pearson correlation coefficient between the different vegetation indicators derived from UAV and tree attributes at the level of individual trees. VI: vegetation index; GRVI: green-red vegetation index; and VARI g: vegetation atmospherically resilient index, green.

Vegetation Index	CSA (Ground)	Stem Volume (Ground)	CSA (UAV)	Stem Volume (UAV)
VI	0.103	0.208 **	0.110	0.218 **
GRVI	−0.106	−0.210 **	−0.113	−0.221 **
VARI g	−0.106	−0.211 **	−0.112	−0.221 **

** Correlation is significant at the 0.01 level (two-tailed).

The selection process for the best independent variable for further analysis was based on the largest positive or negative correlations with the dependent variables. Based on our results, we chose the sum of pixel values as an independent variable to calculate the regression between the spectral data and tree attributes (Figure 7). Our results showed there was high correlation between the UAV main bands and tree attributes based on R^2 (Figure 7).

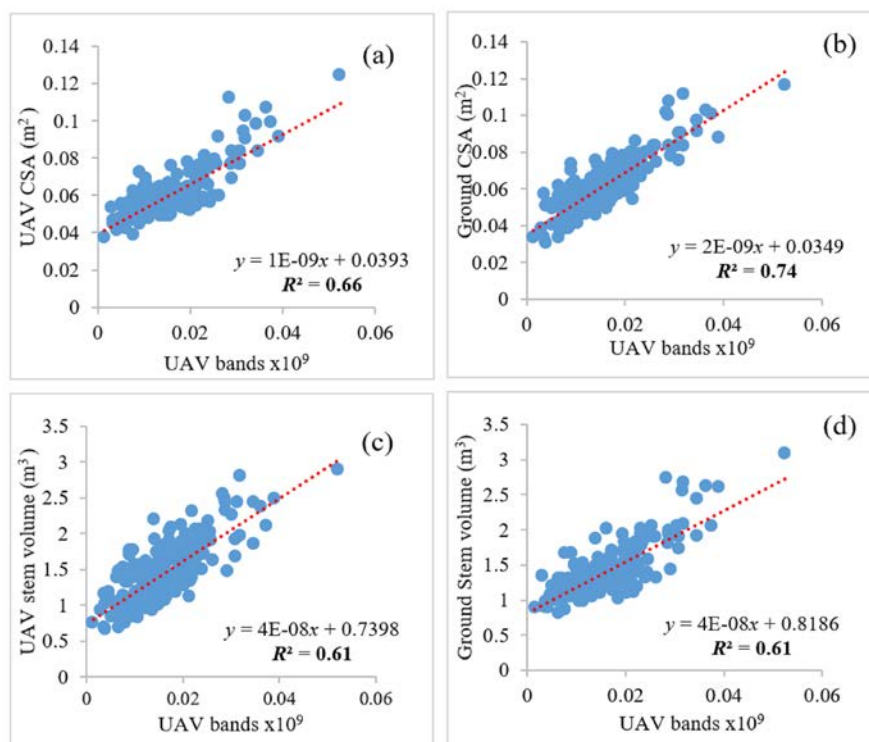


Figure 7. The relationship between spectral data of the UAV bands ($\times 10^9$) and tree characteristics: (a) UAV cross-section area in m^2 ; (b) ground cross-section area in m^2 ; (c) UAV stem volume in m^3 ; and (d) ground stem volume in m^3 .

3.3. UAV-Pléiades Extrapolation

Overall, our findings indicated that there is a strong relationship between the UAV and Pléiades spectral data with an $R^2 = 0.94$ (Figure 8). Based on the regression model equations (Figures 7 and 8), we were able to calculate calibrated formulas that could determine tree attributes, such as CSA and stem volume (Table 6), directly by using the spectral information from Pléiades for each tree. Formulas presented in Table 6 were used to extrapolate the extracted data from the UAV to a larger area based on satellite spectral information.

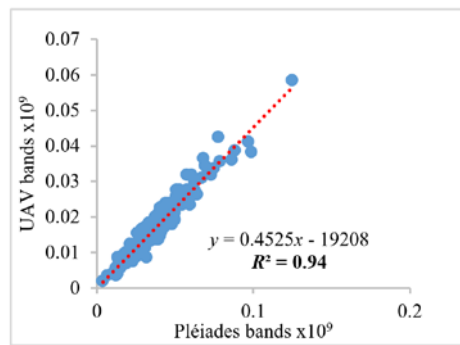


Figure 8. The relationship between spectral data between the UAV and Pléiades bands ($\times 10^9$).

Table 6. Basic errors of CSA and volume based on extrapolated data.

Index	CSA	Volume
RMSE	0.018	0.21
RMSE%	30.56	31.01
Bias	-0.015	-0.39

Our results showed that the computed formulas (9 and 10) could be used for the extrapolation of CSA and volume at the individual tree level with significant accuracy, as can be seen in Table 6. In addition, Figure 9 indicates that there were no significant differences between the mean of CSA and volume derived from UAV and ground data. Also, the same figure and Table 6 show that the extrapolation method has estimated the tree parameters with reasonable accuracy.

$$CSA = 1 \times 10^{-9} (0.4525x - 19208) + 0.0392 \tag{9}$$

$$Volume = 4 \times 10^{-8} (0.4525x - 19208) + 0.7398 \tag{10}$$

where the x is the spectral information derived from satellite imagery.

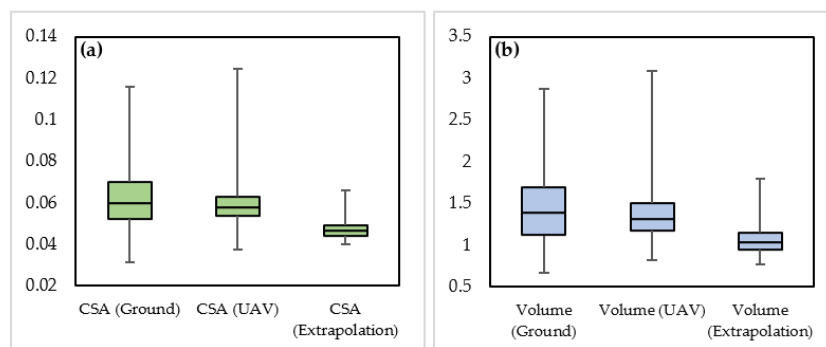


Figure 9. Box-and-whisker plots for comparison of the differences between the means of tree parameters: (a) CSA (m^2) (b) Volume (m^3) derived from different approaches. The medians of the measured values are marked by vertical lines inside the boxes. The ends of the boxes are the interquartile range (upper: Q3 and lower: Q1). The whiskers show the highest and lowest observations.

4. Discussion

Many studies have provided good examples of 3D model reconstructions of VHR DSMs from UAV-based imagery to systematically observe forest attributes, such as tree crown, tree height, and DBH [25,26]. Structural forest attributes are commonly extracted from the CHM using regression

models to predict tree characteristics for forest inventory purposes [27]. Detailed CHMs from remote sensing data have recently gained more attention because they can be used to efficiently predict key forest parameters, such as tree height, crown projection, and stem volume. In this study, it was demonstrated that a consumer-grade camera deployed on a UAV platform generated a 3D scene of the entire study area, which was used to quantify single-tree parameters based on automatic and semi-automatic methods that can be used to support detailed construction or update existing forest inventories. Of course, this is just one of the basic reasons why UAVs are gradually replacing conventional field measurements. However, some applications require more accurate results in terms of absolute means, and this low-cost aerial approach would, therefore, require the collection of ground truth data to ensure this.

Based on the linear regression and RMSE results, we concluded that there is generally a high correlation between estimated (UAV) and observed data (ground) (Table 3; Figures 5 and 6). In this particular study, the methodological approach and the algorithms that were used to determine tree heights and crown diameters were influenced by the relative homogeneity of the study area; all trees were of similar age and had similar morphological characteristics (Table 2). Also, as it can be seen in Figure 9, the amount of CSA and volume for 50% of the trees distributed around the median, but the remaining 50% were distributed in a wider range.

In general, estimation of forest parameters in homogeneous forests is preferable for the application of algorithms, such as local maxima and IWS, because homogeneous forests allow for higher precision using a single kernel size method for smoothing the CHM [20,28,29]. It is evident that the UAV can be used to efficiently estimate heights and crown diameters (Figure 6a–c, and Table 3).

Since there was a high correlation between the crown and DBH from the ground measurements (Figure 5), and due to the fact that UAVs can be used to estimate both heights and crown diameters, there is a possibility for indirect measurement of stem diameters (Figure 6b and Table 3; ID 3) and volume (Figure 6d and Table 3; ID 5) [30].

Additionally, we evaluated the relationship between the tree parameters (CSA and stem volume of individual trees) and the spectral information that was extracted from the UAV imagery. Our results showed that there is a significant correlation between the CSA and stem volume and spectral/textural values derived from the UAV. These results are similar to the results of several studies that have emphasized the possibility of the estimation of tree/stand parameters, such as basal area (BA) and volume, using spectral (main bands) and textural (vegetation indicators) values of aerial/satellite images [31–33]. For the vegetation index (VI), a low value for this attribute implies lower density and stem volume, while higher values indicate dense canopy and higher stem volumes. This explains the positive correlation that we found between the VI, CSA, and stem volume in our study. This result is in accordance with the results of Wallner et al. [31].

The inverse relation between the stem volume and VARI g and GRVI indicators (Table 5) can be explained as follows: due to the fact that vegetation absorbs more red light and reflects more green light, a high value in the red band indicates less vegetation occurrence, which can be explained by the negative correlation between (i) VARI g and GRVI, and (ii) CSA and stem volume.

Finally, our results showed that spectral reflection of individual trees from both Pléiades and the UAV had a high correlation of more than 90% (Figure 8). Based on this finding, high-resolution satellite imagery can be used to extrapolate areas that cannot be reached with a UAV. However, the described methodology can be applied only in the case where the areas present similar structural characteristics. Although the extrapolation process in our study caused underestimation of tree parameters compared to that of ground data and UAV (Figure 9), the suggested methodology can be used for practical forestry and can open up new scientific areas of extrapolation methods to inspire other researchers in the forest community. In future work, we plan to use auxiliary data such as environmental data (edaphic, climatic, and topographic) in order to be able to enhance the methodology for homogeneous areas and to calibrate and assess the methodology in the case of diverse forested areas. For areas with higher variability, one suggestion is to divide the whole area into more homogeneous areas, through the

construction of homogenous groups or classes. We also assume that the use of hyperspectral sensors can lead to a better result with higher accuracy.

5. Conclusions

In this study, we proposed a method to test the performance of UAV image-based point clouds to accurately estimate tree attributes. For this purpose, detection algorithms based on high-quality CHM were used. To potentially improve the results of the 3D image reconstruction model and ensure the integrity of the results based on CHM, we used four GCPs, measured with RTK GPS. Many studies have previously treated estimates of tree parameters, such as tree crown delineation and treetop detection, as two separate procedures [28]. We extrapolated the estimated data from the UAV to a larger area based on a significant correlation (R-squared values and Pearson correlation percentages were greater than 0.90) between the spectral data from UAV and Pléiades. This study demonstrated that it is possible to use calibrated (linear regression) formulas (Table 6) to extrapolate data into larger forested areas (downscaling). Precision forestry is focusing on the use of high-resolution data to support site-specific tactical and operational decision-making (e.g., area productivity) over large forest areas. Therefore, the application of this study, as well as other similar studies, should be explored in the content of the European Common Policy to assess the full potential of these methods for covering larger forested areas.

Regarding the performance of remote sensing versus field measurements, based on our empirical data, we can conclude that the positive comparisons between reference ground measurements and remote sensing estimation of tree attributes confirmed the potential of the workflow process that can be applied as a quick and effective alternative technique to characterize forest tree parameters.

Acknowledgments: This research was supported by the project of the Internal Grant Agency (IGA) of the Faculty of Forestry and Wood Sciences, Czech University of Life Sciences (CULS) in Prague (No. A14/16), and by the Ministry of Agriculture of the Czech Republic Project (No. QJ1520037).

Author Contributions: Azadeh Abdollahnejad and Dimitrios Panagiotidis designed the experiment, wrote the manuscript, and led the image processing and evaluation. Azadeh Abdollahnejad performed the statistical analyses. Peter Surový and Dimitrios Panagiotidis acquired the UAV data. Peter Surový supervised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

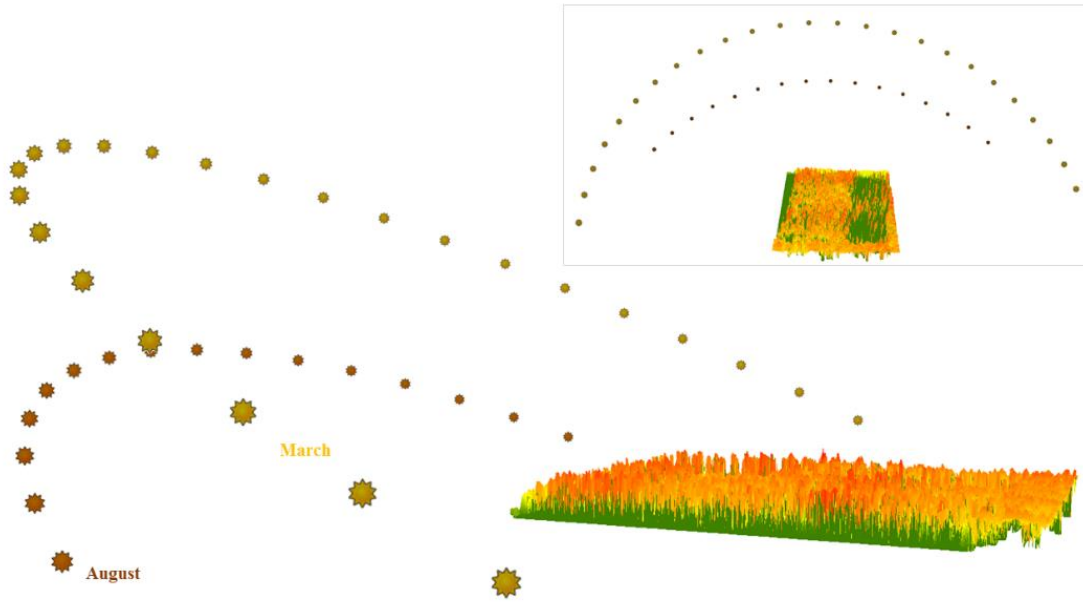
1. Järnstedt, J.; Pekkarinen, A.; Tuominen, S.; Ginzler, C.; Holopainen, M.; Viitala, R. Forest variable estimation using a high-resolution digital surface model. *J. Photogramm. Remote Sens.* **2012**, *74*, 78–84. [[CrossRef](#)]
2. Straub, C.; Stepper, C.; Seitz, R.; Waser, L.T. Potential of UltraCamX stereo images for estimating timber volume and basal area at the plot level in mixed European forests. *Can. J. For. Res.* **2013**, *43*, 731–741. [[CrossRef](#)]
3. White, J.C.; Wulder, M.A.; Vastaranta, M.; Coops, N.C.; Pitt, D.; Woods, M. The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. *Forests* **2013**, *4*, 518–536. [[CrossRef](#)]
4. Stepper, C.; Straub, C.; Pretzsch, H. Using semi-global matching point clouds to estimate growing stock at the plot and stand levels: Application for a broadleaf-dominated forest in central Europe. *Can. J. For. Res.* **2014**, *45*, 111–123. [[CrossRef](#)]
5. Carleer, A.P.; Debeir, O.; Wolff, E. Assessment of Very High Spatial Resolution Satellite Image Segmentations. *Photogramm. Eng. Remote Sens.* **2005**, *71*, 1285–1294. [[CrossRef](#)]
6. Edson, C.; Wing, M.G. Airborne Light Detection and Ranging (LiDAR) for Individual Tree Stem Location, Height, and Biomass Measurements. *Remote Sens.* **2011**, *3*, 2494–2528. [[CrossRef](#)]
7. Popescu, S.C.; Wynne, R.H.; Nelson, R.F. Measuring Individual Tree Crown Diameter with Lidar and Assessing Its Influence on Estimating Forest Volume and Biomass. *Can. J. For. Res.* **2003**, *29*, 564–577. [[CrossRef](#)]

8. Tomppo, E.; Olsson, H.; Ståhl, G.; Nilsson, M.; Hagner, O.; Katila, M. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sens. Environ.* **2008**, *112*, 1982–1999. [[CrossRef](#)]
9. McRoberts, R.E.; Cohen, W.B.; Næsset, E.; Stehman, S.V.; Tomppo, E.O. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scand. J. For. Res.* **2010**, *25*, 340–367. [[CrossRef](#)]
10. Shao, Z.; Zhang, L. Estimating Forest Aboveground Biomass by Combining Optical and SAR Data: A Case Study in Genhe, Inner Mongolia, China. *Sensors* **2016**, *16*, 834. [[CrossRef](#)] [[PubMed](#)]
11. Lehmann, J.R.K.; Nieberding, F.; Prinz, T.; Knoth, C. Analysis of Unmanned Aerial System-Based CIR Images in Forestry—A New Perspective to Monitor Pest Infestation Levels. *Forests* **2015**, *6*, 594–612. [[CrossRef](#)]
12. Kayitakire, F.; Hamel, C.; Defourny, P. Retrieving forest structure variables based on image texture analysis and Ikonos-2 imagery. *Remote Sens. Environ.* **2006**, *102*, 390–401. [[CrossRef](#)]
13. St-Onge, B.; Hu, Y.; Vega, C. Mapping the height and above-ground biomass of a mixed forest using LiDAR and stereo Ikonos images. *Int. J. Remote Sens.* **2008**, *29*, 1277–1294. [[CrossRef](#)]
14. Ozdemir, I.; Karnieli, A. Predicting forest structural parameters using the image texture derived from worldview-2 multispectral imagery in a dryland forest, Israel. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 701–710. [[CrossRef](#)]
15. Shamsoddini, A.; Trinder, J.C.; Turner, R. Pine plantation structure mapping using WorldView-2 multispectral image. *Int. J. Remote Sens.* **2013**, *34*, 3986–4007. [[CrossRef](#)]
16. Immitzer, M.; Stepper, C.; Böck, S.; Straub, C.; Atzberger, C. Forest ecology and management use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. *For. Ecol. Manag.* **2016**, *359*, 232–246. [[CrossRef](#)]
17. Astrium GEO-Information Services, Pléiades Imagery User Guide. Available online: <http://www.cscrs.itu.edu.tr/assets/downloads/PleiadesUserGuide.pdf> (accessed on 1 August 2012).
18. Surový, P.; Ribeiro, N.A.; Pereira, J.S.; Yoshimoto, A. Estimation of Cork Production Using Aerial Imagery. *Rev. Árvore* **2015**, *39*, 853–861. [[CrossRef](#)]
19. Pitkänen, J.; Maltamo, M.; Hyypä, J.; Yu, X. Adaptive Methods for Individual Tree Detection on Airborne Laser Based Canopy Height Model. In *Proceedings of ISPRS Working Group VIII/2: “Laser-Scanners for Forest and Landscape Assessment”*; Theis, M., Koch, B., Spiecker, H., Weinacker, H., Eds.; University of Freiburg: Freiburg, Germany, 2004; pp. 187–191.
20. Panagiotidis, D.; Abdollahnejad, A.; Surový, P.; Chiteculo, V. Determining tree height and crown diameter from high-resolution UAV imagery. *Int. J. Remote Sens.* **2016**, *38*, 1–19. [[CrossRef](#)]
21. Wannasiri, W.; Nagai, M.; Honda, K.; Santitamont, P.; Miphokasap, P. Extraction of Mangrove Biophysical Parameters Using Airborne LiDAR. *Remote Sens.* **2013**, *5*, 1787–1808. [[CrossRef](#)]
22. Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens. Environ.* **1979**, *8*, 127–150. [[CrossRef](#)]
23. Gitelson, A.; Stark, R.; Grits, U.; Rundquist, D.; Kaufman, Y.; Derry, D. Vegetation and Soil Lines in Visible Spectral Space: A Concept and Technique for Remote Estimation of Vegetation Fraction. *Int. J. Remote Sens.* **2002**, *23*, 2537–2562. [[CrossRef](#)]
24. Arulbalaji, P.; Gurugnanam, B. Evaluating the Normalized Difference Vegetation Index Using Landsat Data by Envi in Salem District, Tamilnadu, India. *Int. J. Dev. Res.* **2014**, *4*, 1844–1846.
25. Baltasvias, E.; Gruen, A.; Eisenbeiss, H.; Zhang, L.; Waser, L.T. High-Quality Image Matching and Automated Generation of 3D Tree Models. *Int. J. Remote Sens.* **2008**, *29*, 1243–1259. [[CrossRef](#)]
26. Dandois, J.P.; Ellis, E.C. Remote Sensing of Vegetation Structure Using Computer Vision. *Remote Sens.* **2010**, *2*, 1157–1176. [[CrossRef](#)]
27. Næsset, E. Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data. *Remote Sens. Environ.* **2002**, *80*, 88–99. [[CrossRef](#)]
28. Wang, L.; Gong, P.; Biging, G.S. Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 351–357. [[CrossRef](#)]
29. Jakubowski, M.K.; Li, W.; Guo, Q.; Kelly, M. Delineating Individual Trees from Lidar Data: A Comparison of Vector- and Raster-based Segmentation Approaches. *Remote Sens.* **2013**, *5*, 4163–4186. [[CrossRef](#)]

30. Tuominen, S.; Balazs, A.; Saari, H.; Pölönen, I.; Sarkeala, J.; Viitala, R. Unmanned aerial system imagery and photogrammetric canopy height data in area-based estimation of forest variables. *Silva Fenn.* **2015**, *49*, 1348. [[CrossRef](#)]
31. Wallner, A.; Elatawneh, A.; Schneider, T.; Knoke, T. Estimation of forest structural information using rapid-eye satellite data. *Forestry* **2015**, *88*, 96–107. [[CrossRef](#)]
32. Straub, C.; Weinacker, H.; Koch, B. A comparison of different methods for forest resource estimation using information from airborne laser scanning and CIR orthophotos. *Eur. J. For. Res.* **2010**, *129*, 1069–1080. [[CrossRef](#)]
33. Heiskanen, J. Estimating aboveground tree biomass and leaf area index in a mountain birch forest using ASTER satellite data. *Int. J. Remote Sens.* **2006**, *27*, 1135–1158. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).






Paper VI



Article

UAV Capability to Detect and Interpret Solar Radiation as a Potential Replacement Method to Hemispherical Photography

Azadeh Abdollahnejad ^{1,*} , Dimitrios Panagiotidis ¹ , Peter Surový ¹  and Iva Ulbrichová ²

¹ Department of Forest management, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences (CULS), Kamýcká 129, Prague 165 21, Czech Republic; panagiotidis@fld.czu.cz (D.P.); surov@fld.czu.cz (P.S.)

² Department of Silviculture, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences (CULS), Kamýcká 129, Prague 165 21, Czech Republic; ulbrichova@fld.czu.cz

* Correspondence: abdollahnejad@fld.czu.cz; Tel.: +420-774-844-679

Received: 30 January 2018; Accepted: 6 March 2018; Published: 9 March 2018

Abstract: Solar radiation is one of the most significant environmental factors that regulates the rate of photosynthesis, and consequently, growth. Light intensity in the forest can vary both spatially and temporally, so precise assessment of canopy and potential solar radiation can significantly influence the success of forest management actions, for example, the establishment of natural regeneration. In this case study, we investigated the possibilities and perspectives of close-range photogrammetric approaches for modeling the amount of potential direct and diffuse solar radiation during the growing seasons (spring–summer), by comparing the performance of low-cost Unmanned Aerial Vehicle (UAV) RGB imagery vs. Hemispherical Photography (HP). Characterization of the solar environment based on hemispherical photography has already been widely used in botany and ecology for a few decades, while the UAV method is relatively new. Also, we compared the importance of several components of potential solar irradiation and their impact on the regeneration of *Pinus sylvestris* L. For this purpose, a circular fisheye objective was used to obtain hemispherical images to assess sky openness and direct/diffuse photosynthetically active flux density under canopy average for the growing season. Concerning the UAV, a Canopy Height Model (CHM) was constructed based on Structure from Motion (SfM) algorithms using Photoscan professional. Different layers such as potential direct and diffuse radiation, direct duration, etc., were extracted from CHM using ArcGIS 10.3.1 (Esri: California, CA, USA). A zonal statistics tool was used in order to extract the digital data in tree positions and, subsequently, the correlation between potential solar radiation layers and the number of seedlings was evaluated. The results of this study showed that there is a high relation between the two used approaches (HP and UAV) with $R^2 = 0.74$. Finally, potential diffuse solar radiation derived from both methods had the highest significant relation (-8.06% bias) and highest impact in the modeling of pine regeneration.

Keywords: fisheye camera; unmanned aerial vehicle; canopy height model; solar radiation; crown openness

1. Introduction

Incoming global solar radiation is a key factor which influences energy and water balance and thus is fundamental to most biophysical and physical processes [1,2]. Identifying factors that influence variation in light availability within forested ecosystems represents an important component in our understanding of the complex determinants of tree seedling regeneration, growth, and increment. Also, solar radiation is one of the most influential independent variables in predicting the spatial distribution

of species groups [3,4]. Within forest stands, variation in vegetation composition, structure, crown dimensions, and foliage distribution create spatial variation in light transmittance, affecting growth dynamics. Characterization of the solar radiation regime and forest canopy structural architecture has undergone considerable evolution since [5,6] first reported using Hemispherical Photography (HP).

Hemispherical or fisheye photography is a field-based Remote Sensing (RS) technique to characterize tree canopies from analysis of wide angle (usually between 100 to 180°) photographs [6]. HP photographs are typically acquired looking upward from beneath the forest canopy, where geometric analysis of images is used to characterize the potential light irradiation conditions below the canopy [7]. Taking advantage of fisheye lens polar projection, Coombe and Evans [8], who pioneered the method to characterize light conditions, super-imposed sun path diagrams on HP to predict periods of direct sunlight and the fraction of light that was transmitted to the ground using manual analysis of the photographs. Soon, other studies followed by introducing more specific applications to forestry [9–11]. Since then, a number of studies have been used to estimate particular attributes from HP based on a computerized analysis of HP [7,11,12]. For instance, Chazdon and Field [13] were the first to describe a computerized method to estimate canopy openings, diffuse and direct sunlight, as well as Photosynthetic Photon Flux Density (PPFD), which could be used as key descriptors of the light environment.

Nowadays, with the technological advances in both hardware and software, digital photographs can be analyzed in an optimal manner, using a variety of options and innovative algorithms, to determine more accurately the geometry and position of canopy openings, the path of the sun at different time periods, and subsequently, to indirectly estimate light parameters under forest canopies. Therefore, HP can be used to assess local light environments beneath forest canopies and infer the properties of those canopies.

However, it must be understood that there are limitations associated with the acquisition of HP for further analysis. First, a number of assumptions must be made. One of the most significant is the assumption of foliage effect, meaning that the total amount of Leaf Area Index (LAI) is responsible for the incoming penetration of light through openings, which will eventually determine the conditions for tree regeneration [14]. In digital photographs, canopy areas are assigned to either white (clear sky) or black (referring to the blocked solar radiation by leaves). This may introduce significant errors in darker areas where a significant proportion of the total solar radiation arrives via partial transmission or reflection through a complete canopy layer. Another problem is associated with the acquisition of HP over large areas (i.e., our case study) in regard to exposure time [15]. This may introduce substantial errors during the image analysis, due to the differences of the recorded amount of solar radiation between plots. In general, measuring the light intensity by ground methods is very demanding and labor-costly for practical purposes.

On the other hand, primary platforms such as satellites and piloted aircrafts are used to obtain remote images of the Earth's surface, but these instruments frequently do not deliver adequate spatial and temporal resolutions [16]. The use of Unmanned Aerial Vehicles (UAVs) as a remote sensing platform has recently gained increasing attention, but their applications in forestry are still at an experimental stage. The opportunity to observe the world from the sky using modern approaches (i.e., UAVs) offers the possibility to study forest canopy architecture from an unusual point of view, allowing for observation of some peculiarities of field coverage hardly visible from the ground. One of the main advantages of UAV compared to HP is that it allows data acquisition in shorter time periods (10–15 min) over large areas, resulting in homogeneous data.

Remotely-sensed information offers a unique way to obtain large-scale mapping of forest attributes from Canopy Height Models (CHMs) based on different types of sensors such as Aerial Laser Scanning (ALS) and aerial photogrammetry. Recently, several studies indicated that UAVs deployed with RGB cameras can be used for accurate construction and computation of several forest attributes from CHMs, using Structure from Motion (SfM) algorithms [17–19]. UAV platforms coupled with imaging, ranging and positioning sensors can offer centimeter-level resolution, improving the resolution

of photogrammetric point clouds and the acquisition of three-dimensional (3D) structural data from the forest [20] and therefore offering great possibilities in the precision of forest management applications [21] and the geosciences [22]. Indeed, UAV surveys allow us to work with remote images at very small pixel sizes (Ground Sample Distance-GSD), often in the order of centimeters, a value that greatly improves the normal resolution of an aerial platform.

Regarding the possibility of modelling the amount of solar radiation from UAV-based sensors, several studies that have been found in the literature are mainly focused on either traditional meteorological datasets [23–27] or combined works (point cloud analysis derived from UAV images and meteorological solar radiation datasets) [28]. The results of this study [28] verified the potential of indirect methods for estimating the amount of solar radiation, providing input to ecosystems that are dependent on solar radiation.

Other studies have tested the use of UAV-derived vegetation indices to monitor photosynthetic radiation of canopy attributes [29,30]. The majority of the adopted indices involve the use of Near-Infrared (NIR) bands, because the near-infrared portion of the electromagnetic spectrum provides strong information on both the physiological status and the geometric properties of vegetation [31,32]. On the contrary, no work has been found on the evaluation of the amount of solar irradiance transmitted during the growing seasons that is solely based on UAV RGB imagery.

A limiting factor in passive imaging has been the dependence on sunlight and the high impact of the changing and different illumination conditions on the radiometry of the data [33]. Shadowing and brightening of individual tree crowns cause the pixels of a single tree crown to scale from very dark pixels to very bright pixels. However, only a few methods have been developed to reduce the effect of the changing illumination in the forest canopy [34].

In general, we may say that all information produced from UAV surveys provides a unique method of obtaining estimates over spatially extensive areas. It can help in reliable decision making and planning processes, in improving forest production and in optimizing the resource utilization with quick and cheap regular flights based on inexpensive RGB camera sensors.

In this study, we developed a novel method for data acquisition based on close-range photogrammetric techniques, to determine the correlation between different environmental variables (potential solar radiation, soil and crown openness) and regeneration of *Pinus sylvestris* L., in temperate forest conditions. We compared the UAV-based data with fisheye imagery as the standard method for potential solar radiation characterization for botany and forestry purposes. Our objectives were to answer the following: (i) how the proposed methods for measurement of solar radiation (HP and UAV) are related to each other in terms of overall performance (i.e., data acquisition and analysis); and (ii) to what extent can the methodology be used as an efficient and accurate source for forest management plans and strategies?

2. Materials and Methods

2.1. Characterization of the Study Area

The Doksy territory lies on the shores of Lake Mácha in northern Bohemia in the Czech Republic (Figure 1). It is mainly surrounded by dense forest area, covering 300 km². The area is sandstone pseudokarst in the last stage of development, while the soils are either sandy or peaty, with shallow peaty basins prevailing over rocky sandstone hummocks and sporadic volcanic hills [35]. The whole study area is six hectares and it extends geographically from 50°33′49.48″N; 14°43′36.27″E to 50°33′42.40″N; 14°43′21.75″E, while the Coordinate Reference System (CRS) is WGS84. The study area is located at the northeast part of the city of Doksy as can be seen in Figure 1. The research was carried out in a monoculture stand of *Pinus sylvestris* L., with an age of approximately 140 years. The vegetative period is rather warm and dry. Mean annual air temperature is 7.3 °C and average maximum temperature is 31.5 °C. Mean annual precipitation is 635 mm, with only 354 mm during the

growing season. Mean annual number of cloudy days is 120–150, of which 50–60 occur in the growing season (1 April to 30 October).

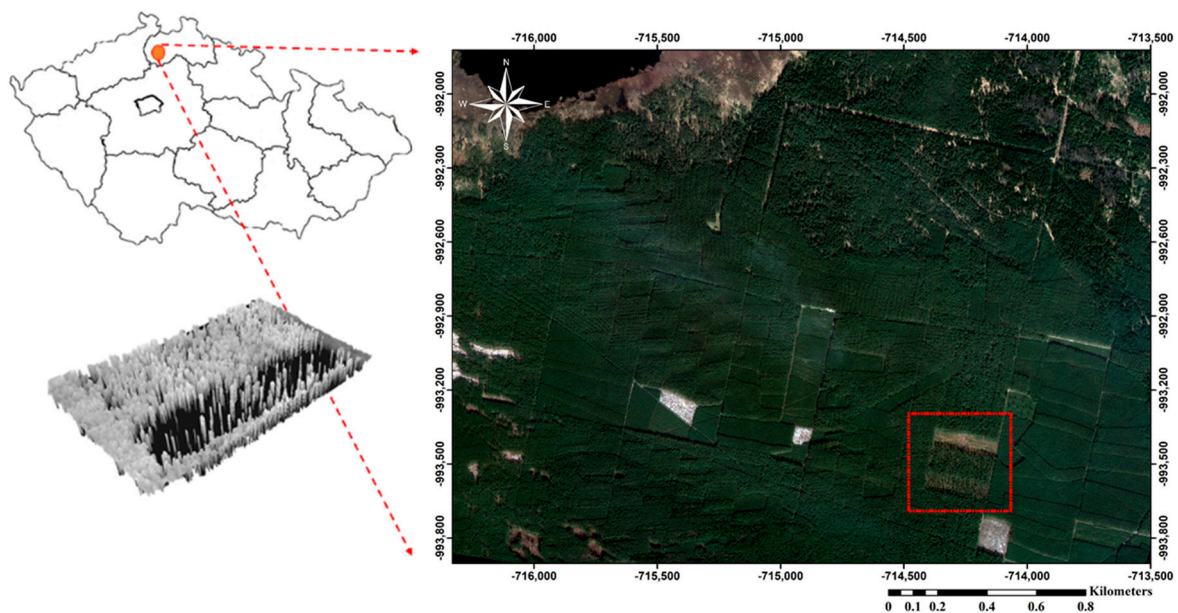


Figure 1. Location of the study area in Doksy, Czech Republic in a local coordinate system (S-JTSK/Krovak East North) source: Pléiades-1A satellite image taken in 2017.

2.1.1. Experimental Plot Design

Different thinning intensities were applied to differentiate the whole area into four different compartments. Each compartment contains rectangular strips with an area of 1.5 ha (rectangle of dimensions $60 \times 250 \text{ m}^2$). Using a thinning control, the stocking percentage was reduced to 0%, 50%, 70% and 30% of crown canopy cover, as can be seen in Figure 2.

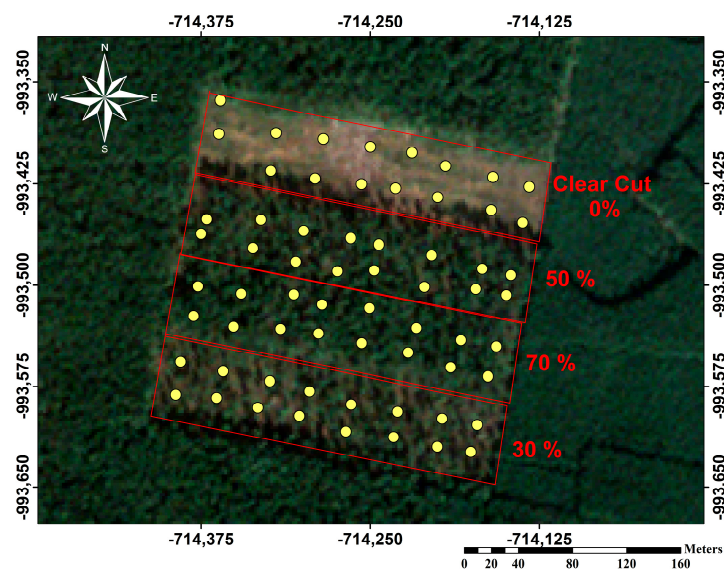


Figure 2. Scheme of the study area showing the locations of the plots in different percentages of canopy cover. Within the whole plot, 64 measuring points was ascertained.

A total station Real Time Kinematic (RTK) TS LEICA TCRP1201 and a portable GNSS LEICA system VIVA with centimeter accuracy were used to measure the position of 10 reference points on the perimeter of the study area.

Additionally, for georeferencing the position of each plot (center), Postex technology was used with 0.1° vertical angle accuracy and distance accuracy of one centimeter, respectively (Figure 3). The positions of sub-sample plots were measured using the azimuth-distance approach based on the central position of each plot. Overall, 512 subplots were measured for natural regeneration, whereas 64 subplots were measured for artificial regeneration.



Figure 3. Illustration of Postex technology for georeferencing the position of each plot.

2.1.2. Pine Regeneration Measurement

The circular areas presented in Figures 2 and 4 had a radius $r = 12.6$ m (500 m²) in the total number of plots $N = 64$. In the center of each circular plot, a homocentric circle with a radius of 3 m was made in order to measure the number of regenerations of *Pinus sylvestris* L. (Figure 4). Smaller circles with a diameter of $d = 0.625$ m, namely 1 to 8, represent sub-sample plots for natural regeneration, whereas the central part of each plot was used for artificial regeneration evaluation and 200 seeds were sown there. The number of seedlings in each plot was counted every month during the vegetation season and at the end of vegetation season in October.

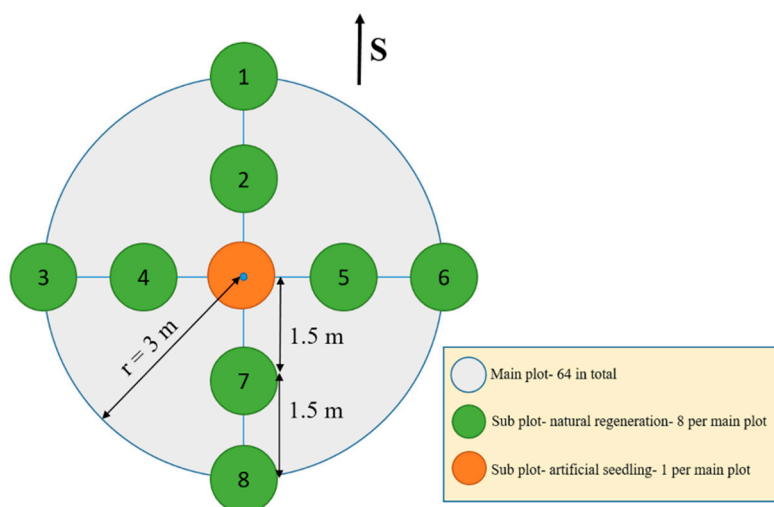


Figure 4. The plots' and sub-plots' position and size.

2.2. Acquisition of Data

2.2.1. Hemispheric Photography

Potential light conditions within canopies were assessed by obtaining hemispherical images (Figure 5) on a cloudy day, 12 November 2016, in the morning (8:00–10:00 a.m.). A Sigma circular fisheye lens (180° diagonal angle of view), focal length 4.5 (Equivalent 35 mm) f/2.8 and a Canon EOS 1100D digital camera, ISO 100; aperture priority were used. Three images (4272 × 2848 pixel) with shifted exposition were taken using the so-called bracketing function (underexposed -2 , -1 , 0) to ensure higher contrast between canopy and sky, at 1.3 m height in the center of each sampling plot (Figure 4). Geographical orientation was ascertained prior to taking the photo (with the upper part of the camera pointing to the north), which is necessary to correct the computation of light parameters by WinSCANOPY 2012a Pro version (Regent Instruments Inc., Ville de Québec, QC, Canada).

The calculation employs slope (in this case none), slope aspect (to the north), sky conditions (cloudy) and altitude, latitude and longitude (50°33′43.20″N; 14°43′34.35″E) as entry parameters of the site and three levels of sharpening intensity as parameters for photos' preliminary processing.

Afterwards, the best (the most contrasting) images from the series were adjusted in Photoshop 8 image editor to a black and white version (Figure 5). To calibrate the process of the photos' conversion from colored to black and white, we used a random sample of 10 photo triads (one with automatic exposure and two underexposed). The standard deviation of crown openness detection was 4.7% for the different light aspects calculated from triads of photos.

The selected light parameters were computed using standard procedures in WinSCANOPY pro 2012a software for canopy analysis, by estimating the following: (i) sky openness (fraction of open sky unobstructed by vegetation above the lens in three-dimensional space); (ii) relative direct and diffuse photosynthetically active flux density under canopy average for the growing season (in MJ/m²·day); and (iii) its total (direct + diffuse flux density).

The calculation employs slope (in this case none), aspect (to the north) and given altitude, latitude and longitude (PPFD direct 3.88 MJ/m²·day, PPFD diffuse 5.08 MJ/m²·day, PPFD total 38.96 MJ/m²·day). The growing season period was defined from 1 April to 30 October.

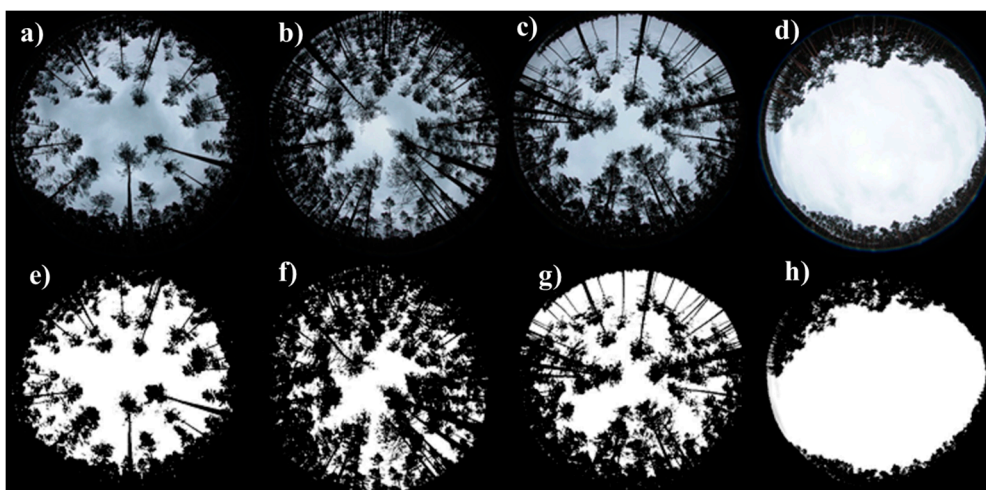


Figure 5. Fisheye images before (a–d) and after processing (e–h) in four different variants of (a,e) 30%, (b,f) 70%, (c,g) 50% and (d,h) 0% (clear-cut).

2.2.2. Acquisition of UAV Imagery

The UAV platform which was used was octocopter SteadiDrone EI8HT RTF (ROBOTS IN SEARCH INC.: Aurora, Canada). As can be seen in Figure 6, it was embedded with an uncalibrated camera, Sony Alpha 6000 high resolution 24.3 megapixels APS-C CMOS sensor and had an adjusted focal

length of 25 mm. For better performance during the data acquisition, the copter was guided by a Dà-Jiāng Innovations Science and Technology Co., Ltd (DJI) ground station to ensure autonomous stabilization and way-point based navigation. The route was planned for a height of 60 m with 85% frontal overlapping and 70% side overlapping, while it needed approximately 8 min to complete a flight, based on the predefined parameters (i.e., number of waypoints). The route was uploaded to the driving unit of the copter (Naza V2), though for security reasons the take-off and landing of the copter were guided by manual radio control. The camera was in automatic mode and a point cloud of images comprising the flight path lines covered the entire study area.



Figure 6. Scaled image that illustrates the actual size of the octocopter Steadi Drone EI8HT RTF that was used for the acquisition of aerial images.

In total, 2341 images were acquired for the SfM image reconstruction process. Due to the low image quality, one of the original 2341 images was excluded from the alignment process using Agisoft Photoscan© V. 1.2.7 (Agisoft LLC: St. Petersburg, Russia). For optimization of the 3D model and before the alignment process, we set the accuracy to high. This step is important because it is the process where the algorithm (SfM) is trying to find and match points between overlapping images and refine the image positions for each photograph separately (Figure 7). The georeferencing of the point cloud was done using a combination of direct-georeferencing and Ground Control Points (GCPs). Even though an on-board GPS module can offer significant accuracy for short base-lines, we additionally used four GCPs randomly distributed within the plot. The GCPs were marked and measured with geodetic RTK GPS (the same one used for the terrestrial measurements) with centimeter accuracy in the field, to further enhance the accuracy of the output model.

2.3. Study Area Model Reconstruction

We also used the same software to construct the Digital Terrain Model (DTM) and Digital Surface Model (DSM) from the dense 3D point cloud model with a 5-cm cell size ($0.05 \times 0.05 \text{ m}^2$). In order to classify the dense point cloud, we used the automatic division tool in Agisoft Photoscan© V. 1.2.7 for classifying all the points into two classes—ground points and the rest—through the Triangulated Irregular Network (TIN) method. For this case study, the parameters that we used for the automatic ground classification was based on the following: (i) maximum angle set to 14 degrees; (ii) max distance 2.5 m; and (iii) cell size in meters remained the same. After this process, each class was colored with a unique color.

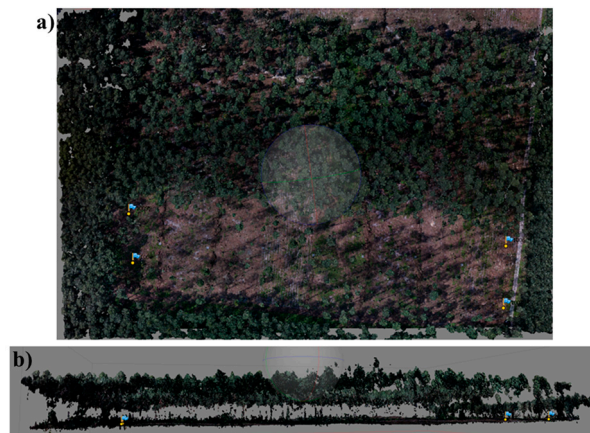


Figure 7. Illustration of different perspectives of the 3D model reconstruction of the study area in Photoscan: nadir (down-looking) camera (a) and side view (b).

2.4. Solar Radiation Analysis

Before any process, we calculated the CHM by subtracting the DSM from the DTM using ArcGIS 10.3.1 by ESRI©. For the sake of process simplicity, we resized the CHM from 5 cm to 1 m, so each single pixel covers a 1 m^2 area. Afterwards, we constructed the sun skymap (Figure 8), during the growing season between 1 April to 30 October, using spatial analyst toolbox in ArcGIS \geq solar radiation toolset \geq area solar radiation tool. We used the CHM ($1 \times 1 \text{ m}^2$) as the input layer by using the default options of the area solar radiation tool except the time configuration option, which was set to multiple days in a year (growing season).

In addition, we constructed different types of potential solar radiation layers: (a) direct radiation (b) diffuse radiation (c) direct duration (d) total solar radiation, in units of $\text{W}\cdot\text{h}/\text{m}^2$ as can be better seen in Figure 9, using spatial analyst toolbox. Diffuse radiation examination is important; because of its characteristic of being reflected in almost all kinds of surfaces, diffuse light reaches deeper and penetrates well into the lower strata of a forest canopy.

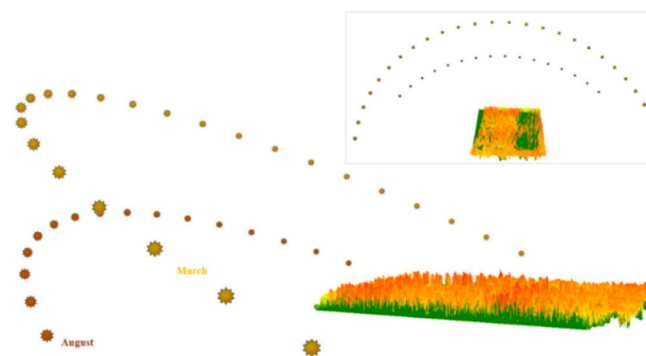


Figure 8. Illustration of the azimuth and sun elevation in two different seasons at the beginning and at the end of the growing season.

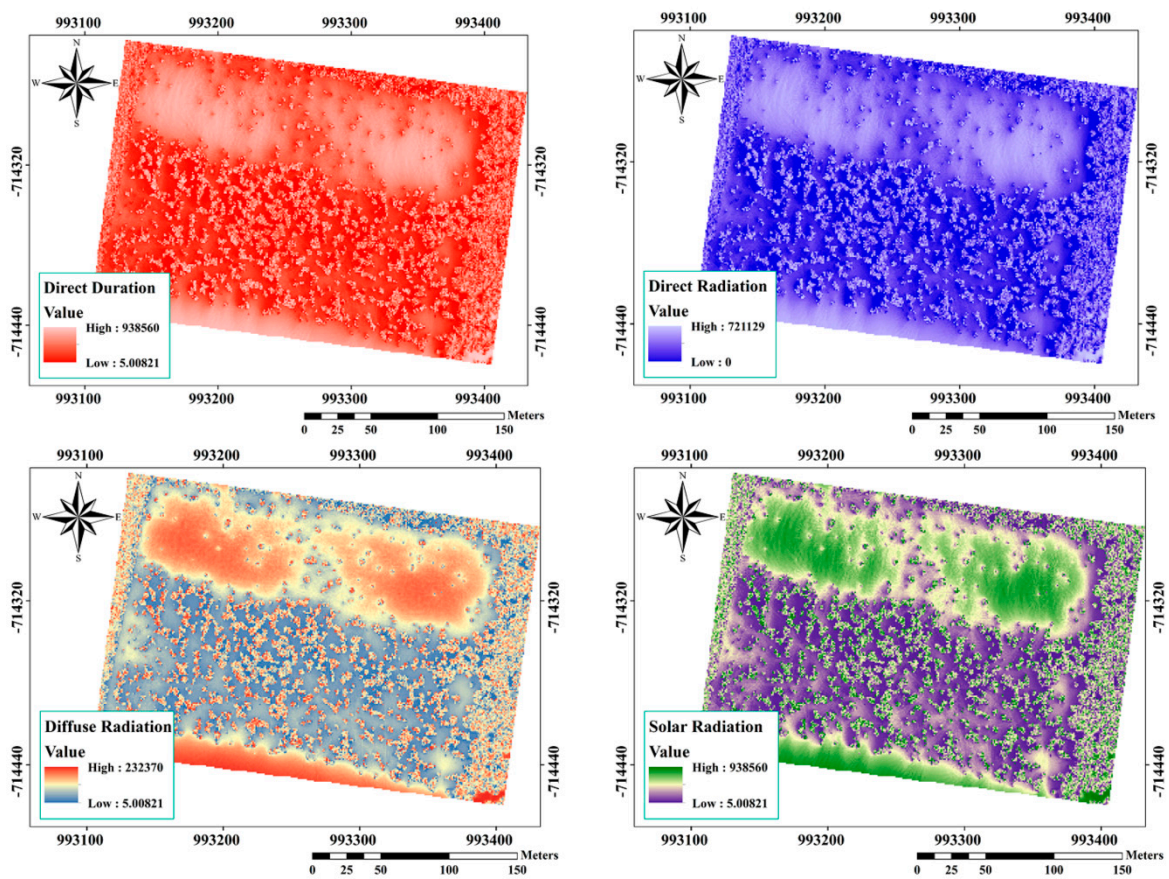


Figure 9. Produced layers of different types of potential solar radiation from the Unmanned Aerial Vehicle (UAV) RGB imagery in units of $W \cdot h/m^2$ (average of growing season).

2.5. Statistical Analysis

In order to be able to compare the results of different sources of RS data, initially we converted the different units in an equivalent measurement result. We used Equation (1) to convert the unit of HP data from $MJ/m^2 \cdot day$ to $W \cdot h/m^2$.

$$1 \text{ W/m}^2 = 1 \text{ J/m}^2/\text{s} \quad (1)$$

where W refers to watt and J to joule.

After the conversion, linear regression was used in order to model the relationship between light intensity and regeneration. Box plots, bias and bias% indicators were used to better illustrate the differences between measured data from HP and the UAV.

$$Bias = \frac{\sum_{i=1}^n (y_i - \hat{y}_j)}{n} \quad (2)$$

where n presents number of plots; y_i represents measured solar radiation data by HP for plot number i ; \hat{y}_j represents the sum of the measured solar radiation data by the UAV for whole plots.

$$Bias\% = \frac{Bias}{\bar{y}_j} \times 100 \quad (3)$$

where \bar{y}_j is the mean of measured solar radiation data by the UAV. All statistical analyses were conducted in IBM SPSS V.24 (IBM Corporation, Armonk, NY, USA) and Excel (Microsoft Corporation, Albuquerque, NM, USA).

3. Results

As we already mentioned in the methodology, first we constructed a map that shows sun positions in different months of the growing season in order to help us to understand the concept of different amounts of solar radiation (Figure 8).

For further analyses, potential solar radiation maps, such as direct duration, potential direct, diffuse and total radiation were prepared using the produced CHM in ArcGIS, as can be seen in Figure 9.

We evaluated the correlation between the crown openness as a structural parameter of our study area and potential solar radiation components (Figure 10). The results showed a significant correlation of 0.96 at a significance level of $\alpha = 0.05$ (Table 1). The same table and Figure 10 show a strong uphill coefficient of determination for the amount of total potential solar radiation and openness (0.93), as expected.

Table 1. Correlation between crown openness and solar radiation derived from Hemispherical Photography (HP).

Regression Statistics	Direct	Diffuse	Total
R	0.96	0.95	0.96
R Square	0.92	0.89	0.93
Adjusted R Square	0.92	0.89	0.92
Standard Error	3.7	4.31	3.6
sig.	0.00	0.00	0.00
df.	63	63	63
Pearson Correlation	0.96	0.95	0.96

The adjusted R-squared is a modified version of R-squared that has been adjusted for the number of inputs in the model and compares the explanatory power of regression models that contain different numbers of predictors.

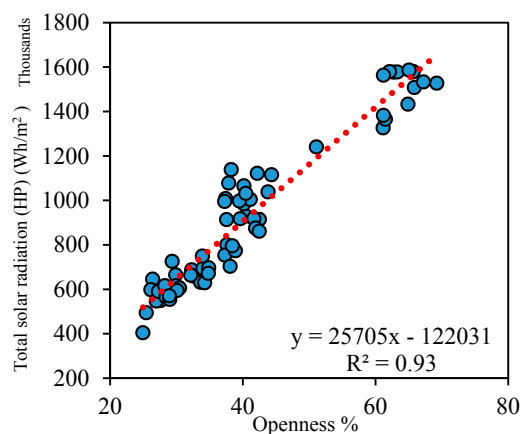


Figure 10. The correlation between the openness and potential total (direct + diffuse) solar radiation derived from HP in $W \cdot h/m^2$ for an entire growing season (64 plots).

3.1. Investigation of a Correlation between HP and the UAV

The main aim of our work was the examination of the correlation between the two imagery techniques, in order to be able to assess the precision of mapping potential solar radiation factors. The correlation between the ascertained potential total (direct + diffuse) solar radiation parameter using HP and the UAV was statistically high resulting in $R^2 = 0.74$, as can be seen in Figure 11 and Table 2. This correlation was calculated for 64 main plots using the average estimated amount of total solar radiation per square meter. Concerning the investigation of correlation between the potential direct and diffuse radiation derived from the UAV and HP, the results showed a significant correlation of 0.83, with higher relation ($R^2 = 0.78$) for the diffuse solar radiation and lower for the direct ($R^2 = 0.68$),

as can be seen in Table 2 and Figure 12. The same table and figure also show that there is a significant difference between potential diffuse and direct radiation in terms of standard error, indicating better a model fit in the case of diffuse radiation.

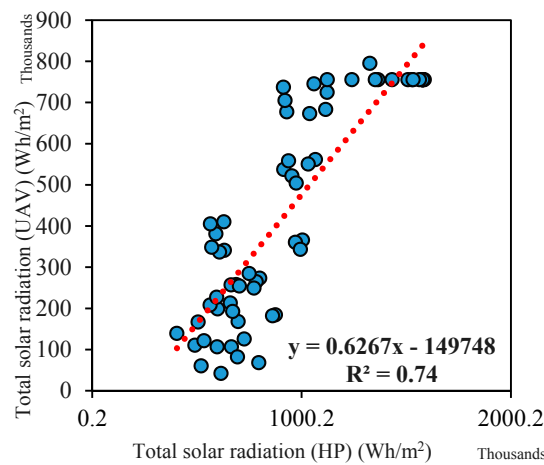


Figure 11. The correlation between the potential total solar radiation derived from HP and the UAV (64 plots) in $W \cdot h/m^2$ for an entire growing season.

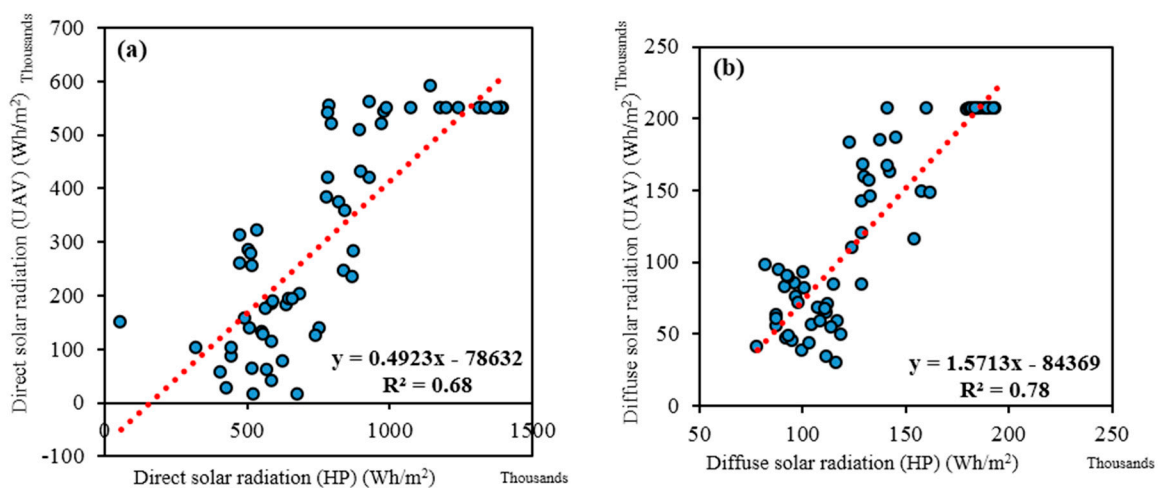


Figure 12. The correlation between the direct (a) and diffuse (b) potential solar radiation as derived from HP and the UAV (64 main plots) in $W \cdot h/m^2$ for an entire growing season.

Table 2. Summary table of regression statistics using HP and the UAV.

Regression Statistics	Total Solar Radiation	Direct Solar Radiation	Diffuse Solar Radiation
Multiple R	0.86	0.83	0.88
R Square	0.74	0.68	0.78
Adjusted R Square	0.73	0.68	0.77
Standard Error	182.41	186.40	16.98
Pearson Correlation	0.86	0.83	0.83
Number of plots	64	64	64

Multiple R is the square root of R-squared and this correlation coefficient shows how strong the linear relationship is. The adjusted R-squared is a modified version of R-squared that has been adjusted for the number of inputs in the model and compares the explanatory power of regression models that contain different numbers of predictors.

In general, as shown in Figure 13, and in accordance with the above results, based on median and mean of values, UAV data underestimated potential solar radiation parameters compared to

HP. The underestimation of values in the case of potential direct solar radiation was significant and had a direct influence on the underestimation of the total amount of potential solar radiation. The performance of HP and the UAV in estimation of potential diffuse solar radiation was not significantly different due to the small amount of bias (-2.70) and bias% (-8.06).

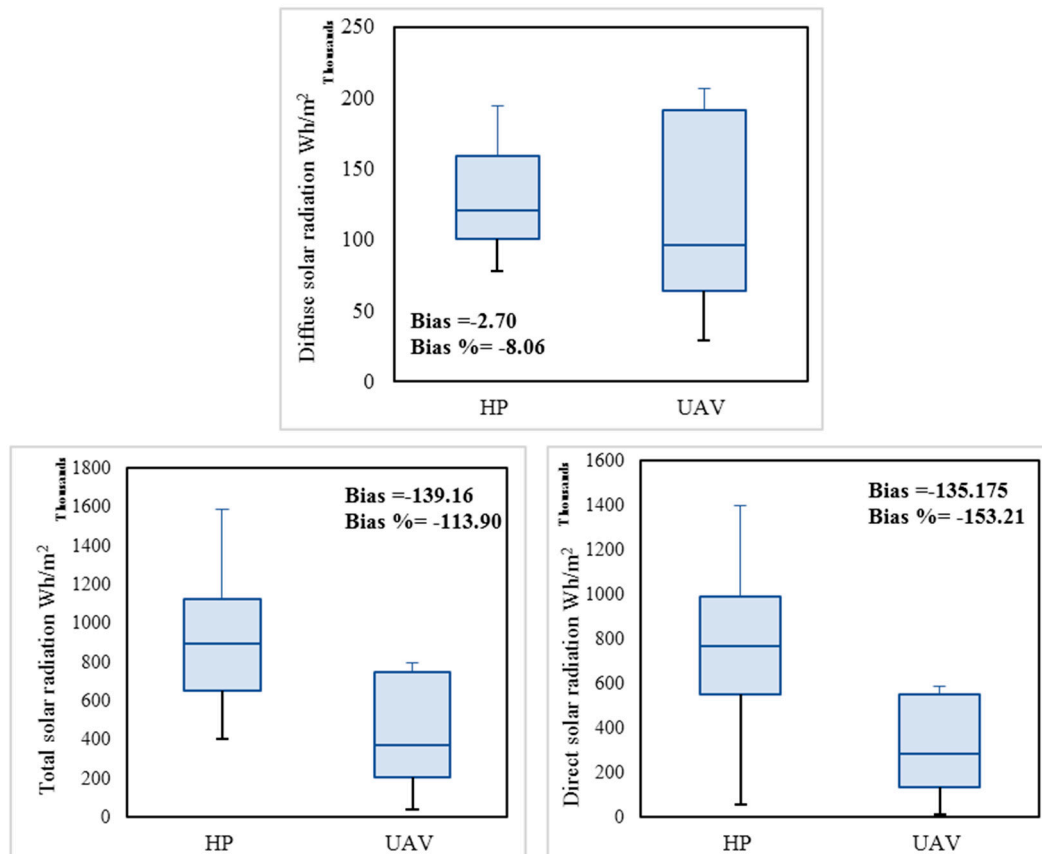


Figure 13. The differences between the mean of potential solar radiation parameters in $W \cdot h/m^2$ for an entire growing season estimated by HP and the UAV. The median of measured values is marked by a horizontal line inside the boxes. The end of the boxes is the interquartile range (the upper: Q3 and lower: Q1). The whiskers show the highest and lowest observations.

3.2. Solar Radiation and Its Components Derived from HP

The investigation of the correlation between the potential solar radiation parameters and regeneration showed that there is an inverse correlation between natural regeneration ($-0.35 \sim -0.37$) and artificial regeneration ($-0.28 \sim -0.30$) and the potential amount of light that can reach the young seedling (Table 3). The results of regression analysis showed that the measured crown openness and potential diffuse solar radiation had the highest relation with natural regeneration (Table 3). In addition, there was a significant relation between the number of artificial regeneration and potential solar radiation parameters at a significance level of $\alpha = 0.05$ (Table 3).

3.3. Solar Radiation and Its Components Derived from the UAV

The assessment of the correlation between the potential solar radiation parameters derived from the UAV and regeneration showed that there is a negative correlation between natural regeneration ($-0.31 \sim -0.34$) and the amount of light that reaches the young seedling. The results of regression analysis showed a significant relation between the solar radiation layers and natural regeneration (Tables 4 and 5). Conversely, there was no significant relation between artificial regeneration and potential solar radiation parameters at a significant level of $\alpha = 0.05$ (Tables 4 and 5). According to the

results in Table 5, we can clearly see that the amount of artificial regeneration is mostly affected by potential direct duration (0.71) and natural regeneration by potential diffuse radiation (0.74).

Table 3. Statistical summary table for the number/amount of natural and artificial regeneration as dependent variables.

Regression Statistics	Natural				Artificial			
	Direct	Diffuse	Total	Openness	Direct	Diffuse	Total	Openness
R	0.35	0.37	0.35	0.38	0.28	0.30	0.28	0.30
R Square	0.12	0.14	0.13	0.14	0.08	0.09	0.08	0.09
Adjusted R Square	0.11	0.12	0.11	0.13	0.06	0.08	0.06	0.08
Std error of the estimate	8.86	8.80	8.85	8.80	7.34	7.28	7.33	7.28
sig.	0.00	0.00	0.00	0.00	0.02	0.01	0.02	0.01
Pearson Correlation	−0.35	−0.37	−0.35	−0.38	−0.28	−0.30	−0.28	−0.30

df = 63. The adjusted R-squared is a modified version of R-squared that has been adjusted for the number of inputs in the model and compares the explanatory power of regression models that contain different numbers of predictors. Estimated Solar radiation in W·h/m² for an entire growing season.

Table 4. Model summary, which describes the relation between the two different regeneration methods as dependent variables and potential solar radiation parameters as independent variables derived from the UAV.

Regeneration Type	R	R ²	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R ² Change	F Change	df1	df2	Sig. F Change
Artificial	0.292 ^a	0.085	0.040	7.427	0.085	1.869	3	60	0.144
Natural	0.349 ^a	0.122	0.078	9.016	0.122	2.776	3	60	0.049

^a: Predictors such as (Constant), mean solar radiation, mean direct duration, mean diffuse radiation; Excluded Variables: Mean direct radiation. The adjusted R-squared is a modified version of R-squared that has been adjusted for the number of inputs in the model and compares the explanatory power of regression models that contain different numbers of predictors. df1: the degree of freedom for predictors. df2: degree of freedom for dependent variables. F change: F-test, which evaluates the changes in R-square while adding the new predictor. Sig. F Change: *p*-value at α level of 0.05.

Table 5. Illustration of the Pearson correlation between the regeneration and potential solar parameters derived from the UAV.

Variable	Artificial		Natural	
	Sig.	Coefficient	Sig.	Coefficient
Diffuse radiation	0.06	−0.188	0.74	−0.319
Direct duration	0.71	−0.165	0.24	−0.341
Solar radiation	0.08	−0.144	0.04	−0.310

Excluded variable: Direct radiation. Estimated Solar radiation in W·h/m² for an entire growing season.

4. Discussion

The main aim of forest management is to provide sustainable forest production and secure the non-productive function of the forests. For this purpose, forest ecosystems need to be regularly monitored and inventoried for different characteristics.

In this work, we investigated two alternative methods for potential light conditions assessment in forests, which are available today—the ground-based (terrestrial imagery using fisheye lens, HP) and the aerial-based (UAV) approaches—and we compared the accuracy and precision of both, and with the ground-based inventory of artificial and natural regeneration. Based on our results (Table 1), estimated potential solar radiation had a significant correlation with the amount of openness (0.96), which is statistically strong evidence of the efficiency of the applied method, where the amount of potential direct and diffuse radiation was estimated based on canopy gap fractions (black (canopy) and white (sky) pixels on photography) and site position on the globe.

The main result of this study shows that the UAV-based “potential light inventory” is comparable with the ground-based inventory using a fisheye camera, according to the high correlation coefficient between the potential direct and diffuse solar radiation derived from the UAV and HP (>0.83). The uphill R-square between the potential direct (0.68) and potential diffuse (0.78) solar radiation indicates that there was a strong relationship between the solar data derived from HP and the UAV.

In general, the generation of intercepted radiation maps is demonstrated to be a feasible solution using low-cost UAVs deployed with spectral cameras, which is in accordance with the results of Guillen-Climent et al. [27].

The amount of natural and artificial regeneration was evaluated in relation to the light conditions (amount of potential diffuse and direct radiation). In the results, there was a higher correlation between potential diffuse radiation and natural regeneration than potential direct radiation; this was probably because of the symmetrical distribution of diffuse radiation around the trees. That indicates that potential diffuse radiation correlates more strongly with other parameters, affecting seedling appearance and survival. This result is in accordance with the results of Strand et al. [36], and Pukkala et al. [37].

In addition, we found an inverse relationship between the amount of natural regeneration and potential solar radiation variables. Essentially, this means that regeneration is lower under a high light regime because, in open crown canopy conditions and especially in clear-cut areas, there are usually less seeds and thus less regeneration potential. We assume that an equal amount of seeds (200 per plot) used in all 64 plots is the main reason for the smaller connection between the crown openness and its components (i.e., potential solar radiation) and artificial seedlings. Regarding the relation between potential solar radiation and its components and the methods of regeneration, natural regeneration is affected more by potential diffuse radiation (0.74). The explanation for such a result could be related to the fact that the majority of seeds that fall directly below the parent plant are more eligible to obtain this type of radiation which is necessary for their growth.

Comparing the performance of HP and the UAV for the estimation of potential solar radiation as can be seen in Figure 13, the amount of potential direct solar radiation was significantly underestimated by the UAV. We assume that the nature of the used method has the highest impact on the underestimation of values in a UAV. In the UAV approach, the gap area within and between canopy is specified by an image-based 3D point cloud which has a close relation with the point cloud density, while in the HP method, the contrast between the sky and leaf values is the main parameter for identifying the amount of crown openness. We also assume that the down-sampling process on CHM from 5 cm to 1 m could be one of the reasons for this underestimation. Since estimation of potential solar radiation, using CHM with $5 \times 5 \text{ cm}^2$ accuracy, was a time-consuming and labor-intensive process for ArcGIS 10.3.1 by ESRI ©, down-sampling was a necessary step to decrease the processing time.

As a result, the correlation coefficient between potential solar radiation computed by HP and the UAV is rather small in the areas with high crown canopy density, where UAV cannot penetrate the canopy and reach the ground surface (Tables S1 and S2).

Regarding the potential diffuse solar radiation, the performance of used methods was not significantly differentiated based on the small amount of bias (−2.70) and bias% (−8.06).

Overall, potential diffuse solar radiation was the most valuable parameter based on (i) higher correlation between the tested methods; and (ii) modeling of pine regeneration. Since the nature of diffuse solar radiation is quite erratic depending on the geographical characteristics of local area and environmental parameters and, due to its importance in developing the empirical methods, estimation techniques have been widely adopted to quantify solar radiation maps, especially diffuse solar radiation [38,39].

5. Conclusions

This paper has demonstrated the ability of a low-cost image-based UAV for the assessment of potential solar radiation as a potential replacement method for HP. At this point, it is worth mentioning

that our case study was located on a flat terrain, thus, the use of CHM instead of DSM had no significant effect during the solar radiation analysis. However, in any other case (i.e., inclined topographical terrain), the use of CHM implies the removal of all terrain information (i.e., altitudes, slopes, slope orientation etc.) which are necessary for the validity of such case studies.

The main result of this study showed that there was a reasonable correlation between the UAV images and the hemispherical photos with $R^2 = 0.74$, especially in the case of estimation of potential diffuse solar radiation (Table S2). This represents statistically strong evidence that UAVs can be a reliable and ergonomic approach, which will allow forest managers to make decisions in an optimal manner. The suggested methodology for modelling the solar radiation can be used for providing auxiliary data from UAV technology. The potential diffuse solar radiation was the most significant UAV output among other potential solar radiation parameters, which can be used as an environmental map for modeling the forest attributes. Remote sensing approaches are already having a substantial impact on forest management practices and with the advent of precision forestry, these works are focusing on deploying high resolution data to support site-specific tactical and operational decision-making.

The presented method can bring new insight into the management and especially into inventorying of forests with high spatial and temporal resolution, such as the construction of continuous maps of potential solar radiation. In the meantime, UAVs and their onboard sensors can combine the advantage of high resolution imagery with quick turnaround series, being therefore suitable for routine forest stand monitoring and real-time applications.

Supplementary Materials: The following are available online at www.mdpi.com/2072-4292/10/3/423/s1. Table S1. Shows the different correlation coefficients between direct and diffuse solar radiation derived from the HP and UAV methods for different crown canopy density. Table S2. Statistical summary table for the two applied methods HP and UAV in units of $W/m^2/S$.

Acknowledgments: This research was supported by (a) the project of the Internal Grant Agency (IGA) of the Faculty of Forestry and Wood Sciences, Czech University of Life Sciences (CULS) in Prague [No. A01/17]; (b) the Ministry of Agriculture of the Czech Republic, Project [No. QJ1520187]; (c) the Czech University of Life Sciences (CULS) in Prague [NAZV QJ1520037].

Author Contributions: Azadeh Abdollahnejad and Dimitrios Panagiotidis designed and wrote the manuscript. Azadeh Abdollahnejad processed the UAV data and analyzed the statistics and data. Peter Surový and Dimitrios Panagiotidis acquired the UAV data; terrestrial data acquisition and plot establishment were conducted by Iva Ulbrichova. Peter Surový supervised the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fu, P.; Richm, P.M. A geometric solar radiation model with applications in agriculture and forestry. *Comput. Electron. Agric.* **2002**, *37*, 25–35. [[CrossRef](#)]
2. Fournier, R.A.; Mailly, D.; Walter, J.M.N.; Soudani, K. Indirect measurement of forest canopy structure from in situ optical sensors. In *Remote Sensing of Forest Environments*; Springer: New York, NY, USA, 2003; pp. 77–113.
3. Peffer, K.; Pebesma, E.J.; Burrough, P.A. Mapping alpine vegetation using vegetation observation and topographic Attributes. *Lands. Ecol.* **2003**, *18*, 759–776. [[CrossRef](#)]
4. Abdollahnejad, A.; Panagiotidis, D.; Shataee Joybari, S.; Surový, P. Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. *Forests* **2017**, *8*, 42. [[CrossRef](#)]
5. Evans, G.C.; Coombe, D.E. Hemispherical and woodland canopy photography and the light climate. *J. Ecol.* **1959**, *47*, 103–113. [[CrossRef](#)]
6. Anderson, M.C. Light relations of terrestrial plant communities and their measurement. *Biol. Rev.* **1964**, *39*, 425–486. [[CrossRef](#)]
7. Rich, P.M. Characterizing plant canopies with hemispherical photographs. *Remote Sens. Rev.* **1990**, *5*, 13–29. [[CrossRef](#)]
8. Coombe, D.E.; Evans, G.C. Hemispherical photography in studies of plants. *Med. Biol. Illus.* **1960**, *10*, 68–75. [[PubMed](#)]
9. Brown, H.E.; Wordley, D.P. Some applications of the canopy camera in forestry. *J. For.* **1965**, *63*, 674–680.

10. Madgwick, H.A.I.; Brumfield, G.L. The use of hemispherical photographs to assess light climate in the forest. *J. Ecol.* **1969**, *57*, 537–542. [[CrossRef](#)]
11. Becker, P.; Erhart, D.W.; Smith, A.P. Analysis of forest light environments. I. Computerized estimation of solar radiation from hemispherical canopy photographs. *Agric. For. Meteorol.* **1989**, *44*, 3–4. [[CrossRef](#)]
12. Chan, S.S.; McCreight, R.W.; Walstad, J.D.; Spies, T.A. Evaluating forest vegetative cover with computerized analysis of fisheye photographs. *For. Sci.* **1986**, *32*, 1085–1091.
13. Chazdon, R.L.; Field, C.B. Determinants of photosynthetic capacity in six rainforest *Piper* species. *Oecologia* **1987**, *73*, 222–230. [[CrossRef](#)] [[PubMed](#)]
14. Roxburgh, J.R.; Kelly, D. Uses and limitations of hemispherical photography for estimating forest light environments. *N. Z. J. Ecol.* **1995**, *19*, 213–217.
15. Jonckheere, I.; Fleck, S.; Nackaerts, K.; Muys, B.; Coppin, P.; Weiss, M.; Baret, F. Review of methods for in situ leaf area index determination. Part I. theories, sensors and hemispherical photography. *Agric. For. Meteorol.* **2004**, *121*, 19–35. [[CrossRef](#)]
16. Nebiker, S.; Annena, A.; Scherrerb, M.; Oeschc, D. A light-weight multispectral sensor for micro UAV—Opportunities for very high resolution airborne remote sensing. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37 Pt 1*, 1193–1198.
17. Surovy, P.; Yoshimoto, A.; Panagiotidis, D. Accuracy of Reconstruction of the Tree Stem Surface Using Terrestrial Close-Range Photogrammetry. *Remote Sens.* **2016**, *8*, 123. [[CrossRef](#)]
18. Mikita, T.; Janata, P.; Surovy, P. Forest Stand Inventory Based on Combined Aerial and Terrestrial Close-Range Photogrammetry. *Forests* **2016**, *7*, 165. [[CrossRef](#)]
19. Panagiotidis, D.; Abdollahnejad, A.; Surovy, P.; Chiteculo, V. Determining tree height and crown diameter from high-resolution UAV imagery. *Int. J. Remote Sens.* **2017**, *38*, 2392–2410. [[CrossRef](#)]
20. Lisein, J.; Pierrot-Deseilligny, M.; Bonnet, S.; Lejeune, P. A photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. *Forests* **2013**, *4*, 922–944. [[CrossRef](#)]
21. Grenzdorffer, G.J.; Engel, A.; Teichert, B. The photogrammetric potential of low-cost UAVs in forestry and agriculture. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 1207–1214.
22. Westoby, M.; Brasington, J.; Glasser, N.F.; Hambrey, M.J.; Reynolds, M.J. Structure from Motion photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology* **2012**, *179*, 300–314. [[CrossRef](#)]
23. Mohammadi, K.; Shamshirband, S.; Tong, C.W.; Arif, M.; Petkovic, D.; Ch, S. A new hybrid support vector machine-wavelet transform approach for estimation of horizontal global solar radiation. *Energy Convers. Manag.* **2015**, *92*, 162–171. [[CrossRef](#)]
24. Olatomiwa, L.; Mekhilef, S.; Shamshirband, S.; Petkovic, D. Adaptive neuro-fuzzy approach for solar radiation prediction in Nigeria. *Renew. Sustain. Energy Rev.* **2015**, *51*, 1784–1791. [[CrossRef](#)]
25. Moghaddamnia, A.; Remesan, R.; Kashani, M.H.; Mohammadi, M.; Han, D.; Piri, J. Comparison of LLR, MLP, Elman, NNARX and ANFIS Models-with a case study in solar radiation estimation. *J. Atmos. Sol. Terr. Phys.* **2009**, *71*, 975–982. [[CrossRef](#)]
26. Feng, Y.; Cui, N.; Zhang, Q.; Zhao, L.; Gong, D. Comparison of artificial intelligence and empirical models for estimation of daily diffuse solar radiation in North China Plain. *Int. J. Hydrogen Energy* **2017**, *24*, 14418–14428. [[CrossRef](#)]
27. Huaiwei, S.; Gui, D.; Yan, B.; Liu, Y.; Liao, W.; Zhu, Y.; Lu, C.; Zhao, N. Assessing the potential of random forest method for estimating solar radiation using air pollution index. *Energy Convers. Manag.* **2016**, *119*, 121–129.
28. Park, J.K.; Das, A.; Park, J.H. Estimating distribution of precision solar radiation using unmanned aerial vehicle. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 6718–6721.
29. Hunt, E.R.; Cavigelli, M.; Daughtry, C.S.T.; McMurtrey, J.E.; Walthall, C.L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
30. Guillen-Climent, M.L.; Zarco-Tejada, P.J.; Berni, J.A.J.; North, P.R.J.; Villalobos, F.J. Mapping radiation interception in row-structured orchards using 3D simulation and high resolution airborne imagery acquired from a UAV. *Precis. Agric.* **2012**, *13*, 473–500. [[CrossRef](#)]

31. Houborg, R.; Boegh, E. Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data. *Remote Sens. Environ.* **2008**, *112*, 186–202. [[CrossRef](#)]
32. Breunig, F.M.; Galvao, L.S.; Formaggio, A.R.; Epiphanyo, J.C.N. Influence of data acquisition geometry on soybean spectral response simulated by the prosail model. *Eng. Agrícola* **2013**, *33*, 176–187. [[CrossRef](#)]
33. Korpela, I.; Mehtätalo, L.; Seppänen, A.; Kangas, A. Tree species identification in aerial image data using directional reflectance signatures. *Silva Fenn.* **2014**, *48*, 1080. [[CrossRef](#)]
34. Kane, V.R.; Gillespie, A.R.; McGaughey, R.; Lutz, J.A.; Ceder, K.; Franklin, J.F. Interpretation and topographic compensation of conifer canopy self-shadowing. *Remote Sens. Environ.* **2008**, *112*, 3820–3832. [[CrossRef](#)]
35. Demek, J. *Hory a Nížiny. Zeměpisný lexikon ČSR*; Academia: Praha, Czech Republic, 1987; p. 584. (In Czech)
36. Strand, M.; Löfvenius, M.O.; Bergsten, U.; Lundmark, T.; Rosvall, O. Height growth of planted conifer seedlings in relation to solar radiation and position in Scots pine shelterwood. *For. Ecol. Manag.* **2006**, *224*, 258–265. [[CrossRef](#)]
37. Pukkala, T.; Kuuluvainen, T.; Stenberg, P. Below-Canopy distribution of photosynthetically active radiation and its relation to seedling growth in a boreal *Pinus sylvestris* L., stand. *Scand. J. For. Res.* **1993**, *1*, 313–325. [[CrossRef](#)]
38. El-Sebaei, A.A.; Al-Hazmi, F.S.; Al-Ghamdi, A.A.; Yaghmour, S.J. Global direct and diffuse solar radiation on horizontal and tilted surfaces in Jeddah, Saudi Arabia. *Appl. Energy* **2010**, *87*, 568–576. [[CrossRef](#)]
39. Khorasanizadeh, H.; Mohammadi, K. Diffuse solar radiation on a horizontal surface: Reviewing and categorizing the empirical models. *Renew. Sustain. Energy Rev.* **2016**, *53*, 338–362. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Chapter 6

6. Discussion

The purpose of this dissertation was to detect and model the forest attributes in forests with different densities using RS and auxiliary data.

Using reconstruction of precise 2D (paper II and III) and 3D (Paper IV, V and VI) models derived from spectral and non-spectral (Paper II and III) data as input, it was possible to evaluate the performance of different statistical algorithms in modeling the tree attributes in forests with different densities (Paper IV and VI) in individual and stand level.

In paper I, “*Forest canopy density assessment using different approaches*” different methods for assessment of crown projection were collected. All the studied methods can be divided into two main categories: pixel and object based approaches. The principle of pixel based approach is based on spectral data derived from pixel cells. The object based approach attempts to detect group of pixels as an object, based on the shape, spectral and textural similarities between the pixels in each group. Means object based method uses the segments that are regions specified by one or more indicators of homogeneity in one or more dimensions (Hay and Castilla 2008). Potential of using different spectral and spatial dimensions is the biggest advantage of object based methods, which make them the most popular methods compare to the pixel based method in recent times (Conchedda et al. 2008). In case of crown canopy classification, shape, size and density of crown as well as the aim of study are the most significant factors which can influence the selected methodology and also have a critical impact on the total accuracy of canopy modelling.

The outputs of paper II, “*Investigation of a possibility of spatial modelling of tree diversity using environmental and data mining algorithms*” showed elevation as an primary topographic parameter was the most significant environmental factor, which has the highest impact on the spatial distribution of tree species. Changes in elevation can cause different environmental conditions, such as temperature, precipitation, moisture, solar radiation, air pressure etc. Therefore, they can affect the spatial distribution of trees. This conclusion is also verified by several other studies

-Chapter 6: Discussion -

like Peffer et al. (2003), Marvie Mohajer (2006), Gracia et al. (2007), Guoyu (2011), Shirzad and Tabari (2011), Kymasi (2012), Momeni Moghaddam et al. (2012).

Among secondary topographic layers, solar radiation has the highest significant impact on the spatial distribution of trees (Peffer et al. 2003; Saatchi et al. 2008, Ghanbari et al. 2011). The importance of solar radiation in the growing season on species spatial distribution and density, was agreeable with the results of (Ejtehadi et al. 1998, Ghanbari et al. 2011), but opposed to the results of Wheatley et al. (2000). As a result of this study, we can conclude using topographic (elevation, slope, aspect and solar radiation etc.), edaphic and climatic (precipitation, temperature and evaporation), layers can yield reasonable outputs in the modelling of spatial distribution and diversity of plants (Zimmermann and Kienast 1999).

In paper III, “*Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data*” the result showed very high correlation between the environmental data and forest attributes similarly to paper II. From the paper II and III, it can be concluded that topography (elevation and solar radiation) is one of the most important factor affecting species composition and distribution, which can be used as an auxiliary source of information for the modelling of forest attributes in large areas.

The results in paper III showed, using auxiliary data such as topographic and climatic layers combined with satellite spectral data derived from Quick-bird can increase, the overall accuracy of modelling of tree species distribution, from 54.21% to 63.85%.

Regarding to the comparison between the performance of the used algorithms in paper II and III, the results showed better performance with RF algorithm. The ability of RF in determining important coefficients, weighting the independent variables and its non-necessity of decision tree structure pruning are all parameters that enhance the functionality and effectiveness of this algorithm. On the contrary, *k*-NN and SVM algorithms use the same proportions of weighting for all independent variables (Kernes and Ohmann 2004). Also, RF can compare the effectiveness of each independent variable and considers different weights for independent layers in the modelling process.

-Chapter 6: Discussion -

Paper IV, “*Determining tree height and crown diameter from high-resolution UAV imagery*” and paper V, “*Estimation and Extrapolation of Tree Parameters Using Spectral Correlation between UAV and Pléiades Data*” showed, the potential using of UAV data as a quick and effective alternative approach to the conventional approaches, in modelling individual tree parameters. Regarding to overall performance of UAVs for extraction of tree parameters, such as crown projection, height and volume from the calculated CHMs, the results are in common with many other studies such as Baltsavias et al. (2008), Dandois and Ellis (2010), Næsset (2002).

Also, based on the results of Paper IV and V we can conclude that estimation of tree parameters in homogeneous forests with lower density is preferable for application of operational tools and algorithms, such as local maxima and Inverse Watershed Segmentation (IWS). That is happening because, homogeneous forests allow for higher precision using a single kernel size method for smoothing the CHM (Wang et al. 2004, Jakubowski et al. 2013).

Due to the fact that UAVs are limited for collecting data over large areas (case study: paper II and III) in paper V, calibrated satellite bands (RGB) was used to extrapolate the estimated forest parameters derived from UAV to a larger area. Paper V showed that there is a possibility of extrapolation of precise data derived from UAVs to larger areas based on the correlation between spectral information of UAV and satellite bands (correlation more than 90%). Based on this finding, high-resolution satellite imagery, such as Pleiades with 2 meters’ spatial resolution for multi-spectral bands can be used to estimate forest attributes, not only over larger areas but also in areas which are not accessible by human. However, the calibrated satellite bands are applicable only in areas with high homogeneity and similar structural characteristics. Noted that, in areas with high heterogeneity in terms of crown projection and species diversity, one suggestion is to divide the study area into smaller areas with lower variety. That division will allow us to process each class individually, by applying an optimal kernel size and unique threshold values. In addition, with the advent of hyperspectral sensors, it is possible to lead to better results especially in case of forests with higher species diversity.

-Chapter 6: Discussion -

Based on the results of our previous studies (paper II and III), potential solar radiation is one of the most important environmental parameters which influence the composition and distribution of tree species. In paper VI, we investigated its impact on regeneration of *Pinus sylvestris* L., in four different thinning areas with different densities (0%, 30%, 50% and 70%). The results showed, potential diffuse solar radiation had higher relation with natural and artificial regeneration of *Pinus sylvestris* L., than potential direct solar radiation. This was because of symmetrical distribution of diffuse solar radiation under crown canopy. This result is in accordance with the results of Strand et al. (2006) and Pukkala et al. (1993).

In addition, the UAV performance in estimation of potential solar radiation based on extracted DSM (from 2341 collected images over the study area) was investigated. The results showed, UAV systems can provide the same accuracy in modeling potential diffuse solar radiation with HP approach, based on the small amount of bias (-2.70) and bias% (-8.06). On the contrary, a significant underestimation of direct solar radiation using UAV comparing to HP was observed. The down-sampling process on CHM from 5 cm to 1 m could be one of the main reasons in UAV underestimation. In addition, differences in principle of estimation of potential direct solar radiation in the two used approaches (UAV and HP), can be another reason for increasing the bias. Also, the results showed that the correlation coefficient between the estimated solar radiation by HP and UAV is rather small in areas with higher density (50% and 70%), because of the inability of RGB camera to penetrate the crown canopy and reach the ground surface. Use of ALS can improve the accuracy of potential solar radiation estimation in dense forests.

Chapter 7

7. Conclusion

In this dissertation, the performance of modern RS approaches through the estimation and modelling of several forest attributes in both individual tree and stand level was investigated.

Based on our findings in Paper II and III, combination of different source of data, such as environmental and spectral can lead to better results in modelling forest parameters. Topography, soil, and climate variables, influence the distribution and composition of tree species and can be used as an auxiliary data during the modelling process in order to enhance the ecological habitat of each group of species.

As we know, forests are integrated complex social-ecological systems, therefore application of multiple-input algorithms (non-parametric), such as RF, SVM and k -NN, allow us to process hundreds of different independent variables simultaneously. Consequently, these algorithms have a better understanding of the relationship and the boundaries, than logistic regression models and are more applicable in such complex case studies.

Regarding the UAV performance in data acquisition and in modelling forest attributes, such as individual tree parameters and mapping potential solar radiation components, the results of papers IV, V and VI showed the possibility of using UAV as a replacement method to the time consuming and costly conventional approaches. Based on our results, we can conclude that uncalibrated mid-level commercial cameras, can provide image based 3D models which can be used for modeling forest parameters, with acceptable accuracy and satisfactory correlation between the observed and estimated values.

Also, results of suggested methodology in paper VI showed, UAV method can be used for providing auxiliary maps, such as potential solar radiation especially in areas with high variety of topographic conditions. That means UAVs and their onboard sensors can combine the advantage of high-resolution imagery with quick turnaround series, being therefore suitable for routine forest stand monitoring and real-time applications and also collecting environmental data over the area of interest.

-Chapter 6: Conclusion -

The provided spatial, temporal and spectral resolution of images by different sensors brings unlimited advantages to natural resource managers and academic researchers for the classification, monitoring, and management of natural ecosystems. 1) availability of required digital data, 2) collection of up-to-date information from the area of interest, 3) selection of the appropriate methodology and materials, 4) application of the correct thresholds values during processing, with respect to the specific characteristics of the study area, aim of study and commercial acceptance and finally 5) producing accurate and useful outputs are the basic requirements for successful RS-based monitoring.

References

- Anderson, M.C., 1964. Light relations of terrestrial plant communities and their measurement. *Biological Reviews* 39, 425–486.
- Abdollahnejad, A., Shataee, S. H., 2014. The study of tree and shrub species diversity changes in the parameters of a physiographic, soil and vegetation. District one of DR. Bahramnia forestry plan. *Journal of Wood and Forest Science and Technology* 21, 61–84. (in Persian)
- Abdollahnejad, A., Panagiotidis, D., Shataee Joybari, S., Surový, P., 2017. Prediction of Dominant Forest Tree Species Using QuickBird and Environmental Data. *Forests* 8, 42.
- Agresti, A., Coull, B.A., 1998. Approximate is Better than "Exact" for Interval Estimation of Binomial Proportions. *The American Statistician* 52, 119–126.
- Applegate, K.E., Crewson, P.E., 2002. An introduction to biostatistics. *Radiological Society of NORTH America*. 225(2), 318–22.
- Atkinson, P.M., 2013. Downscaling in remote sensing. *International Journal of Applied Earth Observation and Geoinformation* 22, 106–114.
- Baatz, M., Benz, U.C., Dehghani, S., Heynen, M., Holtje, A., Hoffmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M., Willhauck, G., 2003. eCognition object oriented image analysis user guide. Definiens Inc., Munich.
- Baltsavias, E., Gruen, A., Eisenbeiss, H., Zhang, L., Waser, L.T., 2008. High-Quality Image Matching and Automated Generation of 3D Tree Models. *International Journal of Remote Sensing* 29, 1243–1259.
- Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding* 110, 346–359.
- Bierkens, M.F.P., Finke, P.A., De Willigen, P., 2000. Upscaling and Downscaling Methods for Environmental Research. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Birth, G.S., Mcvey, G.R., 1968. Measuring colour of growing turf with a reflectance spectrophotometer. *Agronomy Journal* 60, 640–649.
- Boardman, J.W., 1994. Geometric Mixture Analysis of Imaging Spectrometry Data. International Geoscience and Remote Sensing Symposium (IGARSS '94), pp. 2369–2371.

-References -

- Boxwell, M., 2012. Solar Electricity Handbook: A Simple, Practical Guide to Solar Energy, p. 41–42.
- Breidenbach, J., Nothdurft, A., Kandler, G., 2010. Comparison of nearest neighbor approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. *European Journal of Forest Resources* 129, 833–846.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32.
- Breiman, L., 2002. Using Models to Infer Mechanisms. IMS Wald Lecture 2. Accessed on 19 September 2002. Available online: <http://statistics.berkeley.edu/>
- Buschmann, C., Nagel, E., 1993. In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *International Journal of Remote Sensing* 14, 711–722.
- Byroodyan, M., 1990. Weather and Climatology (Ghare Sou River Watershed Studies). Gorgan, Agriculture Publication: 300. (in Persian)
- Camargo, J.A., 1992. New diversity index for assessing structural alterations in aquatic communities. *Bulletin of Environmental Contamination and Toxicology* 48, 428–434.
- Carleer, A.P., Debeir, O., Wolff, E., 2005. Assessment of very high spatial resolution satellite image segmentations. *Photogrammetric Engineering and Remote Sensing* 71, 1285–1294.
- Colomina, I., Blázquez, M., Molina, P., Parés, M.E., Wis, M., 2008. Towards a new paradigm for high-resolution low-cost photogrammetry and remote sensing. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, 37 (B1), 1201–1206.
- Conchedda, G., Durieux, L., Mayaux, P., 2008. An object-based method for mapping and change analysis in mangrove ecosystems. *ISPRS Journal of Photogrammetry and Remote Sensing* 63, 578–589.
- Cortez, P., Morais, A., 2007. A data mining approach to predict forest fires using meteorological data. In Proceedings of the EPIA 2007—Portuguese Conference on Artificial Intelligence, Guimarães, Portugal, 3–7 December 2007.
- Dandois, J.P., Ellis, E.C., 2010. Remote Sensing of Vegetation Structure Using Computer Vision. *Remote Sensing* 2, 1157–1176.

-References -

- Dandois, J.P., Ellis, E.C., 2013. High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sensing of Environment* 136, 259–276.
- Dandois, J.P., Olano, M., Ellis, E.C., 2015. Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure. *Remote Sensing* 7, 13895–13920.
- De Jong, B.H.J., Cairns, M.A., Haggerty, P.K., Ramirezmarcial, N., Ochoa-Gaona, S., Mendoza-Vega, J., Gonzalez-Espinosa, M., March-Mifsut, I., 1999. Land-use change and carbon flux between 1970's and 1990's in central highlands of Chiapas, Mexico. *Environmental Management* 23, 373–385.
- Delcourt, H.R., Delcourt, P.A., Webb, T., 1983. Dynamic plant ecology: the spectrum of vegetation change in space and time. *Quaternary Science Reviews* 1, 153–175.
- Del Río, S., López, V., Benítez, J.M., Herrera, F., 2014. On the use of MapReduce for Imbalanced Big Data using Random Forest. *Information Sciences* 28, 112–137.
- Demek, J., Balatka, B., Kirchner, K., Mackovčín, P., Pánek, T., Slavík, P., 2009. Gemorphological Conditions 1: 500 000. In Hrančiarová, T., Mackovčín, P., Zvara, I. (Eds.). *Landscape Atlas of the Czech Republic*. Výzkumný ústav Silva Taroucy pro krajinu a okrasné zahradnictví Průhonice a Ministerstvo Životního prostředí, Praha.
- Du Prel, J.B., Med, M.P. H., Röhrig, B., Nat, R., Blettner, M., 2009. Critical Appraisal of Scientific Articles. *Deutsches Ärzteblatt International* 6, 100–105.
- Durbha, S.S., King, R.L., Younan, N.H., 2007. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer. *Remote Sensing of Environment* 107, 348–361.
- Du, L., Zhou, T., Zou, Z., Zhao, X., Huang, K., Wu, H., 2014. Mapping Forest Biomass Using Remote Sensing and National Forest Inventory Data. *Forests* 5, 1267–1283.
- Edson, C., Wing, M.G., 2011. “Airborne Light Detection and Ranging (LiDAR) for Individual Tree Stem Location, Height, and Biomass Measurements.” *Remote Sensing* 3, 2494–2528.

-References -

- Ejtehadi, H., Sepehry, A., Horvath, F., 1998. Separability of forest vegetation type using environmental variables including elevation, slop, aspect and direct incoming solar radiation: A GIS application. In Proceedings of the 5th Conference on Geographic Information System, National Cartographic Center (NCC), Tehran, Iran, 10 May 1998.
- Erikson, M., Karin, V., 2003. Finding tree-stems in laser range images of young mixed stands to perform selective cleaning. In: Proceedings of the Scandlaser Scientific Workshop on Airborne Laser Scanning of Forest, pp. 244–250.
- Eisenbeiss, H., 2008. UAV photogrammetry in plant sciences and geology, in: 6th ARIDA Workshop on "Innovations in 3D Measurement, Modeling and Visualization, Povo (Trento), Italy.
- Evans, G.C., Coombe, D.E., 1959. Hemispherical and woodland canopy photography and the light climate. *Journal of Ecology* 47, 103–113.
- Foody, G.M., Darby, S., Wu, F., 2004. GeoDynamics. Taylor and Francis, Boca Raton 440, ISBN 9780849328374.
- Forsman, M., Borlin, N., Holmgren, J., 2016. Estimation of tree stem attributes using terrestrial photogrammetry with a camera rig. *Forests* 7, 61.
- Fournier, R.A., Mailly, D., Walter, J.M.N., Soudani, K., 2003. Indirect measurement of forest canopy structure from in situ optical sensors. In Remote Sensing of Forest Environments; Springer: New York, NY, USA, 2003, pp. 77–113.
- Franco Lopez, H., Ek, A.R., Bauer, M.E., 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbor's method. *Remote Sensing of Environment* 77, 251–274.
- Freixenet, J., Munoz, X., Raba, D., Marti J., Cuff, X., 2002. Yet another survey on image segmentation. In Proceedings of European Conference on Computer Vision, pp. 408–422.
- Fritz, A., Kattenborn, T., Koch, B., 2013. UAV-Based Photogrammetric Point Clouds - Tree Stem Mapping in Open Stands in Comparison to Terrestrial Laser Scanner Point Clouds. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL (1/W2), 141–146.
- Fu, P., Richm, P.M., 2002. A geometric solar radiation model with applications in agriculture and forestry. *Computer and Electronics in Agriculture* 37, 25–35.

-References -

- Gatziolis, D., Liénard, J.F., Vogs, A., Strigul, N.S., 2015. 3D tree dimensionality assessment using photogrammetry and small unmanned aerial vehicles. *PloS one*, 10, e0137765.
- Ghanbari, F., Shataee, S.H., Mohseni, A., Habashi, H., 2011. Using a logistic regression model to predict the spatial characteristics of topography and forest type (case study of a forest series Shastklath Gorgan). *Journal of Forest and Poplar Research* 19, 27–41. (in Persian)
- Gibson, P.J., 2000. “Introductory Remote Sensing Principles and Concepts” Routledge, London.
- Gitelson, A.A., Kaufman, Y. J., Stark, R., Rundquist, D., 2002. Novel algorithms for remote estimation of vegetative fraction. *Remote Sensing of Environment* 80, 76–87.
- Gracia, M., Montané, F., Piqué, J., Retana, J., 2007. Overstory structure and topographic gradient determining diversity and abundance of understory shrub species in temperate forest in central Pyrenees (NE Spain). *Forest Ecology and Management* 242, 391–397.
- Graf, W.L., 1999. Dam nation: A geographic census of American dams and their large-scale hydrologic impacts. *Water Resources* 35, 1305–1311.
- Grenzdörffer, G.J., Engel, A., Teichert, B., 2008. The photogrammetric potential of low-cost UAVs in forestry and agriculture. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Beijing, China, 2008, 37(B1), 1207–1213.
- Guoyu, L., 2011. Topography related spatial distribution of dominant tree species in a tropical seasonal rain forest in China. *Forest Ecology and Management* 262, 1507–1513.
- Hay, G.J., Marceau, D.J., Dube, P., Bouchard, A., 2001. A multiscale framework for landscape analysis: Object-specific analysis and upscaling. *Landscape Ecology* 16, 471–490.
- Hay, G.L., Castilla, G., 2008. Geographic object-based image analysis (GEO BIA): A new name for a new discipline. In: Blaschke T., Lang S., Hay G.J. (Eds.). *Object-Based Image Analysis – Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Berlin, Heidelberg, Springer-Verlag, pp. 93–112.

-References -

- Hill, T., Lewicki, P., 2007. Statistics Methods and Applications. StatSoft, Tusla, OK. 830 p.
- Houghton, R.A., Hackler, J.L., 2003. Sources and sinks of carbon from land-use change in China, *Global Biogeochemical Cycles* 17, 1034.
- Huete, A.R., 1988. A soil-adjusted vegetation index (sAvi). *Remote sensing of environment* 25, 295–309.
- Hufkens, K., Bogaert, J., Dong, Q. H., Lu L., Huang, C. L., Ma, M. G., Che, T., Li, X., Veroustraete, F., Ceulemans, R., 2008. Impacts and uncertainties of upscaling of remote-sensing data validation for a semi-arid woodland. *Journal of Arid Environments* 72, 1490–1505.
- Hunter, J.T., 2003. Factors affecting range size differences for plant species on rock outcrops in eastern Australia. *Diversity and Distributions* 9, 211–220.
- Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting* 22, 679–688.
- Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S., Zhu, Y.H., 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. *Forest Ecology and Management* 128, 109–120.
- Immitzer, M., Stepper, C., Böck, S., Straub, C., Atzberger, C., 2016. Forest ecology and management use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. *Forest Ecology and Management* 359, 232–246.
- Jaakkola, A., Hyyppä, J., Kukko, A., Yu, X. Kaartinen, H., Lehtomäki, M., Lin, Y., 2010. A low-cost multi-sensoral mobile mapping system and its feasibility for tree measurements. *ISPRS Journal of Photogrammetry and Remote Sensing* 65, 514–522.
- Jakubowski, M.K., Li, W., Guo, Q., Kelly, M., 2013. Delineating Individual Trees from Lidar Data: A Comparison of Vector- and Raster-based Segmentation Approaches. *Remote Sensing* 5, 4163–4186.
- Kardgar, N., 2012. Accuracy assessment of soil maps in Dr. Bahramnia forestry plan. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 150. (in Persian)
- Katila, M., Tomppo, E., 2001. Selecting estimation parameters for the Finnish multisource national forest inventory. *Remote Sensing of Environment* 76, 16–32.

-References -

- Kerle, N., Heuel, S., Pfeifer, N., 2008. Real-time data collection and information generation using airborne sensors, In: Geospatial information Technology for Emergency Response, Ed.: Zlatanova S and Li, J., Taylor & Francis, London, UK, 43-74.
- Kernes, B.K., Ohmann, J.L., 2004. Evaluation and prediction of shrub cover in coastal Oregon forests (USA). *Catena* 55, 341–365.
- Krebs, C.J., 1999. Ecological Methodology. 2nd Ed. Menlo Park, Addison-Welsey Educational Publishers, Inc., 620.
- Kymasi, F., 2012. Spatial distribution of tree and shrub species diversity in forests in Golestan province using GIS. [MSc Thesis.] Gorgan, Gorgan University of Agricultural Sciences and Natural Resources: 170. (in Persian)
- Lisein, J., Deseilligny, M.P., Bonnet, S., Lejeune, P.A., 2013. Photogrammetric workflow for the creation of a forest canopy height model from small unmanned aerial system imagery. *Forests* 4, 922–944.
- Lowe, D.G., 2004. Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image, US Patent 6, 711, 293.
- Maack, J., Kattenborn, T., Fassnacht, F. E., Enßle, F., Hernández, J., Corvalán, P., Koch, B., 2015. Modelling forest biomass using Very-High-Resolution Data Combining textural, spectral and photogrammetric predictors derived from spaceborne stereo images. *European Journal of Remote Sensing* 48, 245–261.
- Marvie Mohajer, M.R., 2006. Silviculture. Tehran, Tehran University Press: 387. (in Persian)
- Masek, J., Hayes, D., Turner, D., Hughes, M. J., Healey, S., 2015. The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management* 355, 109–123.
- McRoberts, R.E., Tomppo, E.O., 2007. Remote sensing support for national forest inventories. *Remote Sensing of Environment* 110, 412–419.
- McRobert, R., Tomppo, E., Finley, A., Heikkinen, J., 2007. Estimating aerial means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. *Remote sensing of environment* 111, 466–480.
- McRoberts, R.E., Hansen, M.H., Smith, W.B., 2010. United States of America. In: Tomppo E, Gschwantner T, Lawrence M, McRoberts R.E (Eds.). National

-References -

- forest inventories, pathways for common reporting. Springer, Heidelberg, 567–582.
- Melillo, J.M., McGuire, A.D., Kicklighter, D.W., Moore, B., Vorosmarty, C.J., Schloss, A.L., 1993. Global climate change and terrestrial net primary production. *Nature* 363, 234–239.
- Mikita, T., Janata, P., Surový, P., 2016. Forest Stand Inventory Based on Combined Aerial and Terrestrial Photogrammetry. *Forests* 7, 165.
- Mohammadi, J., Shataee, S.H., Yaghmaee, F., Mahiney, A. 2010. Modelling forest stand volume and tree density using Landsat ETM+ data. *International Journal of Remote Sensing* 7, 2959–2975.
- Mohammadi, J., Shataee, S.H., Babanezhad, M., 2011. Estimation of forest stand volume, tree density and biodiversity using Landsat ETM + Data, comparison of linear and regression tree analyses. *Procedia Environmental Sciences* 7, 299–304.
- Momeni Moghaddam, T., Sagheb Talebi, K.H., Akbarinia, M., Akhavan, M., Hosseini, S.M., 2012. Impact of physiographic and edaphic factors on some of qualitative and quantitative characteristics of Juniperus trees. Case study: Layn region – Khorasan. *Iranian Journal of Forest* 4, 143–156. (in Persian)
- Moore, D., McCabe, G., 2006. Introduction to the practice of statistics. 4th ed. New York: Free man.
- Næsset, E., 2002. Predicting Forest Stand Characteristics with Airborne Scanning Laser Using a Practical Two-Stage Procedure and Field Data. *Remote Sensing of Environment* 80, 88–99.
- Næsset, E., Gobakken, T., Holmgren, J., Hyypä, H., Hyypä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å., Söderman, U., 2004. Laser scanning of forest resources: The nordic experience. *Scandinavian Journal of Forest Research* 19, 482–499.
- Nelson, W.B., 1982. Applied life data analysis, newyork; Wiley.
- Neter, J., Wasserman, W., Kutner, M. H., 1990. Applied Linear Statistical Models, 3rd Edition, Irwin, Boston, MA.
- Newman, J.F., Bonin, T. A., Klein, P. M., Wharton, S., Newsom, R. K., 2016. Testing and validation of multi-lidar scanning strategies for wind energy applications, Wind Energy, doi:10.1002/we.1978, 2016.

-References -

- Ohmann, J.L., Gregory, M.J., 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A. *Canadian Journal of Forest Research* 32, 725–741.
- Ostapowicz, K., Lakes, T., Kozak, J., 2010. Modelling of land cover change using support vector machine. In Proceedings of the 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal, 2010.
- Ozdemir, I., Karnieli, A., 2011. Predicting forest structural parameters using the image texture derived from WorldView-2 multispectral imagery in a dryland forest, Israel. *International Journal of Applied Earth Observation and Geoinformation* 13, 701–710.
- Peffer, K., Pebesma, E.J., Burrough, P.A., 2003. Mapping alpine vegetation using vegetation observation and topographic Attributes. *Landscape Ecology* 18, 759–776.
- Pelgrum, H., Schmugge, T., Rango, A., Ritchie, J., Kustas, W., 2000. Length-scale analysis of surface albedo, temperature, and normalized difference vegetation index in desert grassland. *Water Resources Research* 36, 1757–1765.
- Pickles, J., 1995. Ground Truth: The Social Implications of Geographical Information Systems. 179 p.
- Pinty, B., Verstraete, M. M., 1992. GeMi: A non-linear index to monitor global vegetation from satellites. *Vegetation* 101, 15–20.
- Popescu, S.C., Wynne, R.H., Nelson, R. F., 2003. “Measuring Individual Tree Crown Diameter with Lidar and Assessing Its Influence on Estimating Forest Volume and Biomass.” *Canadian Journal of Remote Sensing* 29, 564–577.
- Posten, H.O., 1978. The Robustness of the Two-Sample *T*-test over the Pearson System. *Journal of Statistical Computation and Simulation* 6, 295–311.
- Powers, D.M.W., 2011. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" *Journal of Machine Learning Technologies* 2, 37–63.
- Pukkala, T., Kuuluvainen, T., Stenberg, P., 1993. Below-Canopy distribution of photo synthetically active radiation and its relation to seedling growth in a boreal *Pinus sylvestris* L., stand. *Scandinavian Journal of Forest Research* 1, 313–325.
- Pyysalo, U., Hyypä, H., 2002. Reconstructing Tree Crowns from Laser Scanner Data for Feature Extraction. International Society for Photogrammetry and

-References -

- Remote Sensing - ISPRS Commission III Symposium (PCV'02). September 9–13, 2002, Graz, Austria.
- Quattrochi, D.A., Goodchild, M.F., 1997. Scale in Remote Sensing and GIS. CRC Press, Boca Raton, FL.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., Sarazzi, D., 2011. UAV photogrammetry for mapping and 3D modeling - Current status and future perspectives. In: Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(1/C22). ISPRS Conference UAV-g, Zurich, Switzerland.
- Röhrig, B., Du Prel, J.B., Wachtlin, D., Blettner, M., 2009. Types of study in medical research: part 3 of a series on evaluation of scientific publications. *Dtsch Arztebl Int.* 2009 April 106(15), 262–8.
- Rosnell, T., Honkavaara, E., 2012. Point cloud generation from aerial image data acquired by a quadcopter type micro unmanned aerial vehicle and a digital still camera. *Sensors* 12, 453–480.
- Rouse, J.W., Haas, R.H., Schell, J.A., Deering, D.W., 1973. Monitoring vegetation systems in the Great Plains with erts. in: fraden s.c., Marcanti e.P., Becker M.A. (Eds.): 3rd erts-1 symposium, nAsA sP-351, 309–317. December 10–14, 1973, Washington, D.c., U.S.A.
- Saatchi, S., Buermann, W., Ter Steege, H., Mori S.A., Smith, T.B., 2008. Modeling distribution of Amazonian tree species and diversity using remote sensing measurements. *Remote Sensing of Environment* 112, 2000–2017.
- Salajanu, D.M., Jacobs, D., 2008. Predicting Spatial Distribution of Privet (LIGUSTRUM SPP) In South Carolina from MODIS and forest Inventory Plot data, forest inventory and analysis. The imaging & geospatial information society, Pacora 17, Denver Colorado, ASPRS 2008.
- Salehi, M.M., Seber, G.A.F., 1997. Two-stage adaptive sampling. *Biometrics* 53, 959–970.
- Shafri, H.Z.M., Ramle, F.S.H., 2009. A comparison of support vector machine and decision tree classifications using satellite data of Langkawi islands. *Information Technology Journal* 8, 64–70.
- Shannon, C.E., Weaver, W., 1949. The Mathematical Theory of Communication. Urbana, University of Illinois Press: 163

-References -

- Shataee, S.H., 2004. Improved classification of forest types by combining spectral data and help establish a method to determine the probability of occurrence of classes of models. *National Mapping Agency* 83, 1–6.
- Shataee, S., Kalbi, S., Fallah, A., Pelz, D., 2012. Forest attribute imputation using machine-learning methods and ASTER data: Comparison of k-NN, SVR and RF regression algorithms. *International Journal of Remote Sensing* 19, 6254–6280.
- Sheeren, D., Fauvel, M., Josipovic, V., Lopes, M., Planque, C., Willm, J., Dejoux, J. F., 2016. Tree species classification in temperate forests using Formosat-2 satellite image time series. *Remote Sensing* 8, 734.
- Shirzad, M.A., Tabari M., 2011. Effect of some environmental factors on diversity of woody plants in Junipeus excels habitat of Hezarmasjed mountains. *Iranian Journal of Biology* 24, 800–808. (in Persian)
- Sironen, S., Kangas, A., Maltamo, M. 2010. Comparison of different non-parametric growth imputation methods in the presence of correlated observations. *Forestry* 83, 39–51.
- Simpson, E.H., 1949. Measurement of diversity. *Nature* 163, 688.
- Sivanpilai, R., Smith, C.T., Srinivasan, R., Messina, M.G., Benwu, X., 2006. Estimation of managed loblolly pine stands age and density with Landsat ETM+ data. *Forest Ecology and Management* 223, 247–254.
- Smith, B., Bastow Wilson, J., 1996. A consumer's guide to evenness indices. *Oikos* 76, 70–82.
- Snaveley, N., Seitz, S, Szeliski, R., 2008. Modeling the world from internet photo collections. *International Journal of Computer Vision* 80, 189–210.
- Snedecor, G.W., Cochran, W.G., 1989. Statistical methods, 8th end. Iowa State University Press. Ames.
- Sripada, R.P., Heiniger, R.W., White, J.G., Meijer, A.D., 2006. Aerial colour infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal* 98, 968–977.
- Strand, M., Löfvenius, M.O., Bergsten, U., Lundmark, T., Rosvall, O., 2006. Height growth of planted conifer seedlings in relation to solar radiation and position in Scots pine shelterwood. *Forest Ecology and Management* 224, 258–265.
- STATISTICA, 2010. Electronic Text Book, Stat Soft Inc. Accessed on 20 November 2010. Available online: <http://statistica.io/>

-References -

- Stehman, S.V., 1997. "Selecting and interpreting measures of thematic classification accuracy". *Remote Sensing of Environment* 62, 77–89.
- Stigler, S.M., 1978. "Mathematical Statistics in the Early States," *The Annals of Statistics* 6, 239–265.
- Tahar, K.N., Ahmad, A., 2011a. Capability of Low Cost Digital Camera for Production of Orthophoto and Volume Determination. CSPA 2011 7th International Colloquium on Signal Processing and Its Applications IEEE. Malaysia.
- Tahar, K.N., Ahmad, A., Wan Mohd Akib, W.A.A., 2011b. UAV-Based Stereo Vision for Photogrammetric Survey in Aerial Terrain Mapping. Proceedings of the IEEE International Conference on Computer Applications and Industrial Electronics. December 4-7, 2011 Penang, Malaysia.
- Tatjana, K., Suppan, F., Schneider, W., 2007. The impact of relative radiometric calibration on the accuracy of k -NN predictions of forest attributes. *Remote Sensing of Environment* 110, 431–437.
- Tomppo, E., 1987. Stand delineation and estimation of stand variates by means of satellite images. Remote Sensing-Aided Forest Inventory. University of Helsinki, Department of Forest Mensuration and Management Research Notes 19, 60–76.
- Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment* 112, 1982–1999.
- Turner, M.G., O'Neill, R.V., Gardner, R.H., Milne, B.T., 1989. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3, 153–162.
- Wallace, L., Lucieer, A., Malenovský, Z., Turner, D., Vopěnka, P., 2016. Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests* 7, 1–16.
- Walton, J.T. 2008. Sub pixel urban land cover estimation: comparing cubist, random forests, and support vector regression. *Photogrammetric Engineering and Remote Sensing* 74, 1213–1222.
- Wang, L., Gong, P., Biging, G.S. 2004. Individual Tree-Crown Delineation and Treetop Detection in High-Spatial-Resolution Aerial Imagery. *Photogrammetric Engineering and Remote Sensing* 70, 351–357.

-References -

- Wheatley, J.M., Wilson, J.P., Redmond, R.L., Ma, Z., Dibenedetto, J., 2000. Automated land cover mapping using Landsat thematic mapper images and topographic attributes. In *Terrain Analysis, Principles and Applications*; Wilson, J.P., Gallant, J.C., (Eds.), John Wiley and Sons: Hoboken, NJ, USA, 2000.
- Wilson, J.P., Gallant, J.C., (Eds.). John Wiley and Sons: Hoboken, NJ, USA, 2000.
- Whittaker, R.H., 1977. Evolution of species diversity in land communities. *Evolutionary Biology* 10, 1–67.
- Wilson, M.F.G., O’Connell, B., Brown, Connell, B., Brown, C., Guinan, J.C., Grehan, A.G., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. *Marine Geodesy* 30, 3–35.
- Wolf, P.R., Dewitt, B.A., 2000. *Elements of Photogrammetry, with Applications in GIS*. McGraw–Hill, Boston, MA.
- Wood, B.D., 2008. The role of scaling laws in up-scaling. *Advances in Water Resources* 32, 723–736.
- Wulder, M. A., White, J.C., Goward, S.N., Masek, J.G., Irons, J.R., Herold, M., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2008. Landsat continuity: Issues and opportunities for land cover monitoring. *Remote Sensing of Environment* 112, 955–969.
- Yarbrough, L.D., 2005. Quickbird 2 Tasseled Cap Transform Coefficients: A Comparison of Derivation Methods. Pecora 16 “Global Priorities in Land Remote Sensing” October 23–27, 2005, Sioux Falls, South Dakota, U.S.A.
- Yu, X., Hyypä, J., Vastaranta, M., Holopainen, M., 2011. Predicting Individual Tree Attributes from Airborne Laser Point Clouds Based on Random Forest Technique.” *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 28–37.
- Yu, X., Hyypä, J., Karjalainen, M., Nurminen, K., Karila, K., Kukko, A., Jaakkola, A., Liang, X., Wang, Y., Hyypä, H., 2015. Comparison of laser and stereo optical, SAR and InSAR point clouds from air- and space-borne sources in the retrieval of forest inventory attributes. *Remote Sensing* 7, 15933–15954.
- Zadeh, L.A., 1965. Fuzzy sets. *Information and Control* 8, 338–353.
- Zhang, H., Fritts, J.E., Goldman, S.A., 2008. Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding* 110, 260–280.

-References -

- Zhang, R., Ma, J., 2008. An improved SVM method P-SVM for classification of remotely sensed data. *International Journal of Remote Sensing* 29, 6029–6036.
- Zimmermann N.E., Kienast F., 1999. Predictive mapping of alpine grassland in Switzerland: Species versus community approach. *International Journal of Vegetable Science* 10, 469–482.