INSTITUTO POLITÉCNICO DE BRAGANÇA Escola Superior Agrária

Modeling height-diameter relationship in *Pinus pinaster* Ait. in the Forest Intervention Zone of Lomba, NE-Portugal

Khaoula Msaouak

Dissertation presented to the School of Agriculture of Bragança in partial fulfilment of the requirements for the degree of Master of Science in Forest Resources Management

Advisors

Luís Filipe Nunes (PhD)

School of Agriculture of Bragança (PORTUGAL)

Fernando Pérez-Rodríguez (PhD)

School of Agriculture of Bragança (PORTUGAL)

Bouhaloua Mhammed (PhD)

Agronomic and Veterinary Institute Hassan II (MOROCCO)

This dissertation does not include the comments and suggestions of the Jury

BRAGANÇA

JULY 2016

Il faut toujours remercier l’arbre à Karité sous lequel on a ramassé le bon fruit pendant la bonne saison

De Ahmadou Kourouma

This work was supported by the SIMWOOD Project (Sustainable Innovative Mobilisation of Wood), EU FP7 Collaborative Project 2013-2017 Grant Agreement Nº 613762.

ACKNOWLEDGEMENT

First of all, I would like to thank deeply my thesis advisor Professor Luís Filipe Nunes of Escola Superior de Agrária at Instituto Politécnico de Bragança. His office door was always open whenever I ran into a trouble spot or had a question about my research or writing. He steered me in the right direction whenever he thought I needed it.

I would also like to acknowledge Doctor Fernando Pérez-Rodríguez of Escola Superior de Agrária at Instituto Politécnico de Bragança as my Co-Adviser. I am gratefully indebted to his help but also for his very valuable comments on this thesis.

My special thanks to Professor Bouhaloua Mhammed from Institut Agronomique et Veterinaire Hassan II, Rabat for his thoughts and advices toward the work.

My sincere thanks to Doctor Manuel Arias-Rodil from Escuela Politécnica Superior de Lugo, Universidade de Santiago de Compostela, Spain, for his advices and help with the R program for calibration of mixed models.

My thanks to Dr. Sara Sarmento (MsC) and Dr. Cristina Patrício (MsC) from ARBOREA for their big help in the field work and also to SIMWOOD team from of Escola Superior de Agrária at Instituto Politécnico de Bragança for the facilitation of the logistics of the transportation and field work in Lomba ZIF.

An immense acknowledgement to Erasmus+ Project, namely to Professor Luis Pais, Vice-President of IPB, and to the International Relations Office (IRO), and Dra Joana Aguiar for their help and backness, to offer up this chance to do my Master Degree in Portugal. A big thank to Professor Chtaina, Director of Agronomic department from Institut Agronomique et Vétérinnaire Hassan II in Morocco, for their encouragement to participate in this project.

My special thanks to all the teachers of the units courses I received during the Academic Year of 2015/2016, related to the Master Degree on Management of Forest Resources in the School of Agiculture of Polytechnic Institute of Bragança in Portugal, for their big support and also their love and respect .

Finally, I must express my very profound gratitude to my parents and to my family, my friends Chaimaâ Echtaibi and Mohammed Maghfoul but also my class mates for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Khaoula Msaouak

Abstract

In this work, the relationship between diameter at breast height (d) and total height (h) of individual-tree was modeled with the aim to establish provisory height-diameter (h-d) equations for maritime pine (*Pinus pinaster* Ait.) stands in the Lomba ZIF, Northeast Portugal. Using data collected locally, several local and generalized *h-d* equations from the literature were tested and adaptations were also considered. Model fitting was conducted by using usual nonlinear least squares (nls) methods. The best local and generalized models selected, were also tested as mixed models applying a first-order conditional expectation (FOCE) approximation procedure and maximum likelihood methods to estimate fixed and random effects. For the calibration of the mixed models and in order to be consistent with the fitting procedure, the FOCE method was also used to test different sampling designs. The results showed that the local *h-d* equations with two parameters performed better than the analogous models with three parameters. However a unique set of parameter values for the local model can not be used to all maritime pine stands in Lomba ZIF and thus, a generalized model including covariates from the stand, in addition to *d*, was necessary to obtain an adequate predictive performance. No evident superiority of the generalized mixed model in comparison to the generalized model with nonlinear least squares parameters estimates was observed. On the other hand, in the case of the local model, the predictive performance greatly improved when random effects were included. The results showed that the mixed model based in the local *h-d* equation selected is a viable alternative for estimating *h* if variables from the stand are not available. Moreover, it was observed that it is possible to obtain an adequate calibrated response using only 2 to 5 additional *h-d* measurements in quantile (or random) trees from the distribution of *d* in the plot (stand). Balancing sampling effort, accuracy and straightforwardness in practical applications, the generalized model from nls fit is recommended. Examples of applications of the selected generalized equation to the forest management are presented, namely how to use it to complete missing information from forest inventory and also showing how such an equation can be incorporated in a stand-level decision support system that aims to optimize the forest management for the maximization of wood volume production in Lomba ZIF maritime pine stands.

Keywords: *Pinus pinaster*; *h-d* equations; nonlinear least squares; mixed models; maximum likelihood; calibration; R software.

Resumo

Neste trabalho modelou-se a relação entre o diâmetro à altura do peito (d) e a altura total da árvore individual com vista ao estabelecimento de equações hipsométricas provisórias para povoamentos de pinheiro bravo (*Pinus pinaster* Ait.) na ZIF da Lomba, Nordeste de Portugal. Com dados recolhidos localmente, testaram-se várias equações locais e generalizadas apresentadas na literatura, tendo-se considerado igualmente algumas adaptações. No ajustamento das equações utilizou-se o método dos mínimos quadrados não lineares (MQNL). Os modelos selecionados, local e generalizado, foram testados na forma de modelos mistos, tendo-se utilizado um método de estimação condicional de primeira ordem e estimadores de máxima verosimilhança para estimar os efeitos fixos e aleatórios. Para a calibração dos modelos mistos usou-se o mesmo método de aproximação linear para testar diferentes tipologias de amostragem. Os resultados revelaram melhores performances das equações hipsométricas locais com dois parâmetros relativamente às análogas com três parâmetros. Contudo, o mesmo conjunto de valores dos parâmetros não pode ser usado em todos os povoamentos de pinheiro bravo da ZIF da Lomba e, portanto, uma equação hipsométrica generalizada com inclusão de variáveis do povoamento para além de *d* torna-se necessária para uma capacidade preditiva adequada. Não se observaram evidências de superioridade do modelo misto generalizado relativamente ao obtido por MQNL. Por outro lado, no caso da equação local, a performance preditiva melhorou claramente após a inclusão de efeitos aleatórios. Os resultados mostraram que o modelo misto baseado na equação local selecionada é uma alternativa viável para a estimação de *h* na ausência de informação sobre variáveis do povoamento. Além disso, observou-se que é possível obter uma calibração adequada com medições adicionais de apenas 2 a 5 pares *h*-*d* em árvores correspondentes a quantis (ou ao acaso) da distribuição de *d* na parcela (povoamento). Ponderando o esforço de amostragem, exatidão e facilidade de uso em aplicações práticas, o modelo generalizado resultante do ajustamento por MQNL é recomendado. São apresentados exemplos de aplicação da equação generalizada selecionada em gestão florestal, nomeadamente como usar a equação para completar informação em falta num contexto de inventário e também para mostrar como uma equação deste tipo pode ser incorporada num sistema de suporte à decisão ao nível do povoamento com o objetivo de otimizar a gestão para a maximização do volume lenhoso produzido em povoamentos de pinheiro bravo da ZIF da Lomba.

Palavras-chave: *Pinus pinaster*; equações hipsométricas; mínimos quadrados não lineares; modelos mistos; máxima verosimilhança; calibração; Software R.

Index

Index of Figures

Index of tables

1. Introduction

Forest management is essential in order to enable, in a context of forest multifunctionality and sustainability, the supply of wood and non-wood products as well as ecosystem services, tailored to the physiographic, climatic and soil characteristics, in line with the aspirations and socio-economic needs of the populations.

Important instruments for the forest management in Portugal are the Regional Plans of Forest Management ("Planos Regionais de Ordenamento Florestal" – PROF – in Portuguese) concluded by the end of 2006. The elaboration of the PROF was contemplated, from the outset, in the law that establishes the bases of the forest policy in Portugal (Law 33/1996 of August 17). PROF, which are currently being reviewed, are main guidelines for planning and managing the forest regionally. In local scale PROF are technically implemented by the Forest Management Plans ("Planos de Gestão Florestal" – PGF – in Portuguese). These plans are articulated with National and Municipality Plans of Defense Against Forest Fires (PNDFCI and PMDFCI) and Special Plans of Forest Intervention (PEIF). There is then the integration with territory planning instruments.

The forest ownership in Portugal causes problems to the implementation of efficient management strategies. The state forests represent only 2% of the total forest area, 8% are communal forests, private forests owned by industries are 13% and the majority is individual private forest, representing 77% (Mendes, 2002). Thus, most of the forest belongs to small owners that are not professionals, owners who are both small farmers and local communities (Coelho, 2003). Concerning property dimension, more than 60% of forest owners have small (1 to 5 ha) or very small (<1ha) properties, mainly located in the North and Centre and related to the species: maritime pine (*Pinus pinaster* Ait.), eucalyptus (*Eucalyptus, sp.)* and also chestnut (*Castanea sativa* Mill.) (Baptista and Santos, 2005). According to the authors the profile of these forest owners is one of no investment and low or no-management.

Given the small size of the majority of the forest properties, with particular focus on the North and Central regions, obtaining minimum areas for adequate management is very dependent on the attitudes and way of being of the owners and necessarily involves grouping scenarios (ENF, 2006). These scenarios may consist in promote the association for common management through the creation of Forest Intervention Zones (Decree-Law 127/2005 of August 5) (Zonas de Intervenção Florestal – ZIF – in Portuguese), whose key objectives are: i) promoting sustainable management of forest areas under the ZIF; ii) planned coordination for the protection of forests and natural spaces; iii) reducing conditions for ignition and spread of forest fires; iv) recovery of these spaces. Thus, the creation of scale for the management is stimulated allowing gains in efficiency through the ordering and reparcelling of forest properties, and discourages land fracturing.

The forest inventory has a key role in supporting forest management as it provides crucial information about the state of the forests. Depending on the objectives, inventories can have more or less complexity. Nowadays, additionally to the traditional information about volume of wood growing stocks, biomass and carbon, health and vitality, are examples of the kind of information collected in the framework of national forest inventories. Tree and stand data obtained from inventories are very useful for obtaining new or recalibrate existent forest growth and yield models. These models are valuable tools and can be incorporated in forest simulators or decision support systems for helping forest planning and management (e.g., Diéguez-Aranda et al., 2009; Rojo-Alboreca et al., 2015). One equation that can be included in a growth and yield modeling framework is the one that relates individual tree total height (h) with individual tree diameter at breast height (d), usually known by *h-d* relationship.

In the Northeast (NE) of Portugal, regionally developed modeling tools to support forest management are lacking. In this work the *h-d* relationship in *Pinus pinaster* stands from a forest intervention zone located in NE Portugal (Lomba ZIF) was modeled in order to establish provisory local and generalized *h-d* equations for supporting forest management. The usual nonlinear ordinary least squares approach was used as well as the more recent mixed models approach. Examples of application height-diameter equation to the forest management are presented.

2. State of the art

2.1. Modeling height-diameter relationship

The breast-height diameter (d) and total height (h) of individual-tree are two variables frequently measured in forest inventories and used in forest management plans (Adame et al, 2008; Avery and Burkhart, 2011). Despite theoretically the variable height could be measured in all trees from a stand, from a practical point of view, it can be time-consuming and expensive to do so. Total height is usually measured indirectly with height measuring instruments based on angle and distance measures. So, also, many times it is complicated to measure height with a very high precision as in very dense stands, where the top of the trees is difficult to visualize (Larjavaara and Muller-Landau, 2013). On the other hand, tree diameter can be quickly and simply measured with high accuracy and little cost. Thus, it is usual in forest inventories to measure the diameter of all the trees in the plots and take a subsample of trees to measure height or not take any measurements of this variable at all.

As a result of the difficulty in measuring tree height and the cost associated with field inventories, and as *h* and *d* are correlated, it is common practice to fit height-diameter (h-d) models to predict *h* from measured *d* (Crecente-Campo et al 2010). Development of simple and accurate models that allow forest managers to determine with reliability the height of the trees in a stand from diameter data is a prime objective in forest management. Knowledge of the relation between these variables permits managers to obtain, without investing large amounts of money in height measurement, the input values needed to estimate individual-tree volume, dominant height of the stands, competition indices for individual-tree growth, height/diameter ratio, and structural diversity indexes (Calama and Montero, 2004). Parresol (1992) refers that *h-d* equations can be helpful for damages appraisal.

Following Adame et al. (2008), several practical examples of these applications can be mentioned, including i) to assess individual-tree volume (Larsen and Hann, 1987; Jayaraman and Lappi, 2001), ii) to determine the social position of the tree within the stand (Colbert et al., 2002), iii) to find the dominant height and from this, calculate the index of site productivity (Huang and Titus, 1993; Vanclay, 1994; Jayaraman and Lappi, 2001), iv) to describe stand growth dynamics and succession (Curtis, 1967; Peng et al., 2001). Other examples (Crecente-Campo et al., 2010) include characterizing canopy height diversity and wildlife habitat relationships (Spies and Cohen, 1992; Morrison et al., 1992). Applications cover a wide variety of forest species and structures from pure even-aged (e.g., López Sánchez et al., 2003) to mixed uneven-aged stands (e.g., Corral-Rivas et al., 2014).

Growth and yield models are useful tools for forest management. Many growth and yield-projection systems also used height and diameter as the two basic input variables, with all or part of the tree heights predicted from measured diameters (Curtis et al., 1981; Wykoff et al., 1982; Arney, 1985). Because of their importance for a number of forest stand modeling applications, *h-d* equations have received considerable attention and, in addition to predicting average heights associated with *d* classes in diameter distribution systems, *h-d* relationships have been also employed in stand-table projection and individual-tree growth and yield simulators (Burkhart and Tomé, 2012). According to Tomé (1989), two types of *h-d* equations can be considered: local and generalized; in the first type, *h* is usually only dependent on tree diameter and these equations can be applied to the stand where the data were collected (local application). In the second type, *h* is a function of tree diameter, age, and other stand variables and the equations can be applied at the regional level.

Yuancai and Parresol (2001) listed the following desirable characteristics for functions used to model *h-d* relationships: (1) increase monotonically, (2) have an upper asymptote, and (3) have an inflection point. According to these authors, the Schnute function and the Richards function are probably the most flexible and versatile functions available for modeling *h-d* relationships. These functions have outperformed other functions in several studies (Huang et al., 1992; Zhang, 1997; Peng, 1999; Peng et al., 2001). In contrast, in other studies no uniformly best function or model formulation were found (e.g., Mehtätalo, 2004; Mehtätalo et al., 2015). Paulo et al. (2011) questioned the third requirement of Yuancai and Parresol's list mentioned before. The authors worked with a dataset containing diameter values in a large range (from young to old individuals) and no evidence of an inflection point was found when plotting height against diameter. According to Burkhart and Tomé (2012), both diameter and height growth curves have an inflection point, but this may not be necessarily so for the relationship between height and diameter. Crecente-Campo et al. (2010) stated also that, a sigmoidal-shaped tendency (which includes characteristics 1-3 above), as observed for example in a dominant height growth model, is not necessary in a *h-d* relationship, as it only expresses the relationship between two variables at a given point in time, and not any trends in growth.

The height–diameter relationship varies from stand to stand, and even within the same stand the relationship is not constant over time (e.g., Curtis, 1967; Pretzsch, 2009). Zeide and Vanderschaaf (2002) refer that stand density is one important obvious factor that may modify the height-diameter relationship and should be included in *h-d* models to increase accuracy of predictions. Staudhammer and Lemay (2000) note that introducing stand density variables into the base (local) *h-d* equations resulted in increased accuracy for predicting heights. Therefore, a single curve cannot be used to estimate all the possible relationships that can be found within a forest (Castedo Dorado et al., 2006). So, the use of stand-level information in addition to *d* could provide more accurate estimates of height than using only *d* (Sharma and Zhang, 2004). This is particularly important in mixed uneven-aged stands in which different species, ages, structures and levels of competition coexist (Vargas-Larreta et al., 2009).

Several stand variables have been explicitly used in generalized *h-d* equations to cope with the need to represent all possible conditions in forest stands. As some examples, stand density measures were used by Larsen and Hann (1987), Staudhammer and Lemay (2000), Temesgen and Gadow (2004), Tesmegen et al. (2007), Newton and Amponsah (2007) and Corral-Rivas et al. (2014); relative tree position variables were used by Temesgen and Gadow (2004) and Tesmegen et al. (2007); site quality variables were used by Bennet and Clutter (1968), Larsen and Hann (1987) and Wang and Hann (1988); and the average *h* and *d* of the top height trees (similar to dominant height and dominant diameter, *hdom* and *ddom*, respectively) were used by Krumland and Wensel (1978) and Hanus et al. (1999). Harrison et al. (1986) only used dominant height in addition to *d*. Quadratic mean diameter (dg) was used for example in Mirkovich (1958), Gaffrey (1988), and Hui and Gadow (1993). Age (t) has also been used (e.g., Lenhart, 1968; Burkhart and Strub, 1974; Soares and Tomé, 2002).

When revising literature about generalized *h-d* equations, Tomé (1989) realized that not much variability existed, concerning the use of this type of equations. The author selected a few equations from the literature but also deduced new equations by restricting the local or base models of Michailoff (1943), Stoffels and Van soest (1953) and Prodan (1965) to pass in the point (ddom, hdom). For example, Petras et al. (2014) adopted similar procedure with equations of Michailoff (1943) but restrict them to the point (dg, hg), where *hg* is mean stand height. Nowadays, more variability of generalized *h-d* equations exist, resulting from new deductions or simple parameter expansion of base models with stand variables, and lists of such equations can be found for example in Soares and Tomé (2002), López Sánchez et al. (2003), Adame et al. (2005), Crecente-Campo et al. (2010), and Stankova and Diéguez-Aranda (2013).

When modeling the height–diameter relationship, measurements of both variables taken from trees growing in sample plots located in different stands or regions are usually used (Calama and Montero, 2004; Castedo Dorado et al., 2006). The hierarchical (nested) structure of *h-d* data (i.e., trees grouped in plots and plots grouped in stands) can result in a lack of independence between measurements because the observations in each sampling unit will be correlated (Gregoire, 1987). Mixed models approach has been successfully used to address this type of problem. This relatively recent approach was applied in several studies (e.g., Hökkä, 1997; Eerikäinen, 2003; Calama and Montero, 2004; Mehtätalo, 2005; Sharma and Parton, 2007; Trincado et al., 2007; Adame et al., 2008; Budhathoki et al., 2008; Crecente-Campo et al., 2010; Stankova and Dieguez-Aranda, 2013, Corral-Rivas et al., 2014; Mehtätalo et al., 2015). Mixed models simultaneously estimate fixed parameters (parameters that are common to the entire population) and random parameters (parameters that are specific to each plot) within the same model and enables the variability between plots of the same population to be modeled (Corral-Rivas et al., 2014). Thus, mixed models allow prediction of a typical response, using only fixed effects, and a calibrated response (calibration) where random effects are predicted and included in the model if measurements of additional heights from a sample of trees are available (Burkhart and Tomé, 2012). According to Sharma and Breidenbach (2015), inclusion of both stand variables as covariates and possible sources of subject-specific variation (e.g., plot-level variation) as random effects into the model potentially can better describe the variation in the height-diameter relationship on stand level than a model with no random effects included. Because computational power is enabling huge simulation calculus, unblocking practical applications in some areas of statistics, we end this literature review section, referring also two recent papers that used mixed models in a Bayesian framework as an alternative to classical frequentist framework of maximum likelihood mixed effects models (Li et al., 2012; Zhang et al., 2014).

2.2. Measuring *h* and *d* in forest inventory

The diameter at breast-height (d) and total height (h) of individual-tree are two variables frequently measured at plot level in forest inventories. These two variables have a unique importance in forest management and planning (Adame et al, 2008; Avery and Burkhart, 2011). There are numerous tools that can be used to collect information about these variables. Timber cruising has been the preferred method that foresters use for forest inventory and thus, new measurement instruments, to replace more traditional ones, were developed and improved in that context. On the other hand, new concerns with biomass production, health and vitality of forests in the context of climate change as well as rapid changes of land use, lead to the need of a continuous update inventory of forest resources. As a consequence, for example, it is necessary to have better assessments of canopy characteristics. As technology has advanced, new remote sensing techniques have been developed in order to help facing these challenges.

Diameter measuring tools

In the absence of other tools, diameter *d* can be approximately measured using a simple tape to obtain the girth at breast height and then dividing by π (3.1415). Graduated sticks, as the Biltmore stick and similar, are also early tools that were used for measuring *d*. To improve accuracy of estimates, specific diameter tapes that directly measure *d* and other expedite and efficient instruments like calipers, quickly become available.

The caliper is commonly used today to measure stem diameters over bark both in standing and felled trees. It consists of a fixed arm mounted perpendicularly to a graduated beam and a movable arm, parallel to the former and sliding along the fixed beam (Figure 1a). Aluminum calipers have increased in popularity and are frequently used. They should be regularly checked for their accuracy and, if necessary, calibrated at least once annually (van Laar and Akça, 2007). A new generation of calipers, the digital (sometimes called electronic) calipers are available (Figure 1b) with wireless Bluetooth, infrared (IR) and USB technology, including computational capabilities for pre-treatment of data. These kind of calipers are designed for easy collection and storing information of many thousands of trees. Data can also be sent to handheld computers and Android mobiles or tablets.

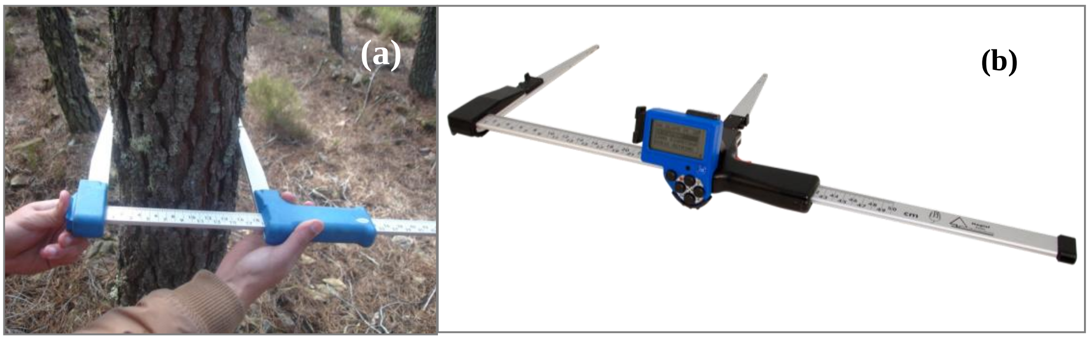


Figure 1. (a)- Haglöf Mantax Caliper; (b)- Haglöf DP II digital Caliper(source: http://www.haglofcg.com/).

Other tool commonly used is the Diameter tape. It is a measuring tape used to estimate the diameter of a cylinder object, as is the stem of a tree. A diameter tape has either metric or imperial measurements reduced by the value of pi. Thus, the tape measures the diameter of the object. It is assumed that the cylinder object is a perfect circle.

Height measuring tools

The tree height is subject to more error and harder to measure than diameter. Approximate estimates of tree height can be obtained with more rudimental methods such as the stick methods. It is also possible to measure *h* using a clinometer and a tape. The standard practice is to measure tree heights with hypsometers. The conventional hypsometers are classified according to the principle of their construction based on the trigonometric relationship or geometric basis. Modern hypsometers are based in ultrasonic technologies (like the Haglöf Vertex IV in Figure 2) or laser technologies (e.g., Haglöf VL5 Vertex Laser; Nikon Forestry Pro Laser Rangefinder/Height Meter).

The Vertex IV, which was used in this study, is a relatively new ultrasonic digital instrument for measuring heights, distances, vertical angles, slopes and current temperatures. To perform measurements, the instrument must be equipped with an external unit the transponder (Figure 2). The principle of height measurement with the Vertex is very simple. The transponder is placed on a tree to be measured, at a calibrated height which, for metric system, is usually 1.30 m from the ground. When the measurer aims at the transponder the instrument registers the angle and the distance to the transponder, from which the horizontal distance and the height from the transponder to the isohypse are calculated. When aimed at tree top the instrument calculates the height from the isohypse to the tree top using the previously established horizontal distances and the angle.

Figure 2. The vertex IV with the transponder (source: http://www.haglofcg.com/)

This ultrasonic hypsometer is based on the trigonometric principle measure, from eye level, the vertical angles between the baseline and the top and base of the tree, respectively (Figure 3). The tree height is obtained from measurements of the angle subtended by the top and base of the tree with the horizontal.

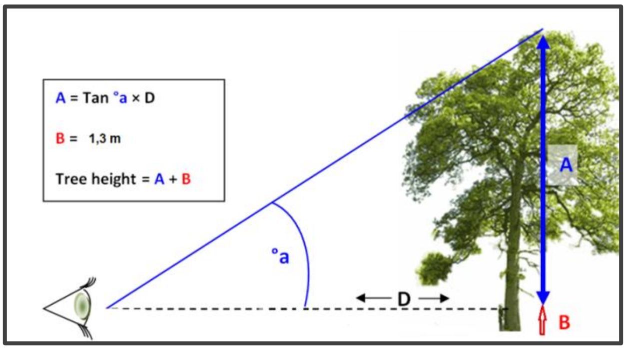


Figure 3. The trigonometric principle of the hypsometer

Other emerging tools

Laser technologies seem to be promise in the development of new measurement instruments. These technologies allow to develop tools in order that a single device can measure height, diameter and take other information such as location and distance. As example we can refer the Laser-relascope (Kalliovirta et al., 2005). Also usage of LiDAR technology is becoming a more common practice for forest inventory analysis (Magnusson et al., 2007).

LiDAR, which stands for Light Detection and Ranging, is a remote sensing method that uses light in the form of a pulsed laser to measure ranges (variable distances) to the objective we want to measure. A LiDAR instrument principally consists of a laser, a scanner, and a specialized GPS receiver. LiDAR originated in the early 1960s, shortly after the invention of the laser, and combined laser-focused imaging with radar's ability to calculate distances by measuring the time for a signal to return. Its first applications were in meteorology. The National Center for Atmospheric Research used it to measure clouds (Goyer and Watson, 1963). The general public became aware of the accuracy and usefulness of LiDAR systems in 1971 during the Apollo 15 mission, when astronauts used a laser altimeter to map the surface of the moon. It was introduced in the 1990s as a tool for topographic mapping (Flood, 2001). Despite the initial focus on topographic mapping, research quickly demonstrated the high potential of LiDAR for forestry applications (Nilsson, 1996; Naesset, 1997; Magnussen and Boudewyn, 1998; Lefsky et al., 1999; Lim and Treitz, 2004; Naesset et al., 2004; Hyyppä et al., 2008). Since 2001, as a complement to traditional measurements, airborne LiDAR technology has been used to rapidly describe forest structure over large areas (Dassot et al., 2011).
Because, airborne LiDAR scanning provides limited information at the tree scale or under the canopy, which is required for certain forest applications, complementary terrestrial LiDAR (or terrestrial laser scanning (TLS)) technologies have therefore been implemented to obtain detailed information at the tree or plot scales. Today it is possible to obtain plot level variables (namely stem diameter and height in some conditions), with high accuracy with TLS. According to Newnham et al. (2015), despite the initial intimation that these instruments could replace manual measurement methods, they cannot be viewed as a logical progression of existing plot-based measurement. TLS must be viewed as a disruptive technology that requires a rethink of vegetation surveys and their application across a wide range of disciplines.

3. *Pinus pinaster* Ait.

3.1. Botanical description

The maritime pine (*Pinus pinaster* Ait.) is a medium-size tree that can reach until 20-40 m tall and, usually, 40 to 50 cm of diameter at breast height when adult (Correia et al., 2007). The bark is thick and fissured, presenting reddish slots. When the diameter growth becomes dominant compared to growth in height, the bark becomes less thick and less fissured, more compact and grayish (Correia et al., 2007). The leaves (needles) are in groups of two but, occasionally, groups of three can be observed. The first needles (primary) are bluish, presenting tolerance to shade, allowing the young seedlings to develop in shadow conditions the first year. Secondary needles are darker with an average duration of 2 to 3 years, depending on the climate conditions (Correia et al., 2007). According to Rushforth (1986), the cones are conic, 10-20 cm long and 4-6 cm broad at the base when closed, green at first, ripening glossy red-brown when the tree is 24 months old. They open slowly over the next few years, or after being heated by a forest fire, to release the seeds, opening to 8-12 cm broad. The seeds are 8-10 mm long, the wing is a 20-25 mm, and are wind-dispersed. Correia et al. (2007) refer that in normal conditions the maximum distance seeds can reach is some hundreds of meters but in extreme conditions it is possible to reach a few kilometers.

3.2. Distribution

The maritime pine is widely spread over the western Mediterranean region, the High Atlas and Tunisia in North Africa (Carrión et al, 2000). Its main populations are located in the Iberian Peninsula, where the species has adapted to extremely cold winters of central Spain and to the milder temperate climate of the Atlantic coast (Blanco et al., 1997). Population genetics studies have identified three main refugia across its range: the Atlantic coast of Portugal, southwestern Iberia, and Pantelleria and Sardinia in Italy (Ribeiro et al., 2001; Abad Viñas et al., 2016) (Figure 4). Meanwhile, 200 000 hectares are located in other areas of reforestation (Australia, South Africa, New Zealand, Chile, Argentina and Uruguay).

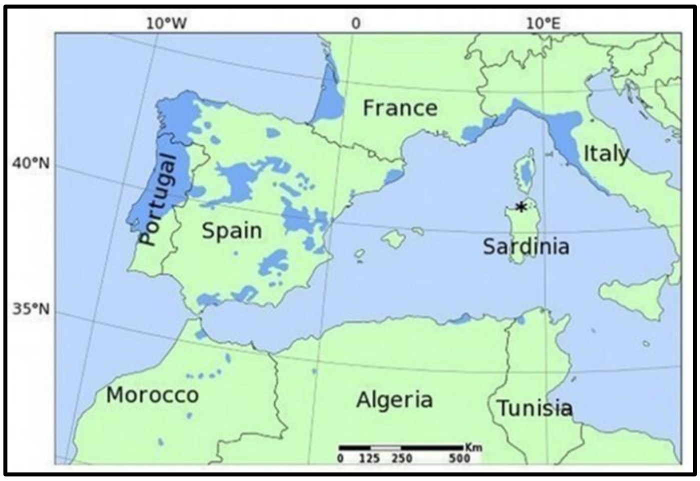


Figure 4. Main distribution areas of *Pinus pinaster* (Mazza, 2014).

In Portugal, where forests represent 35%, corresponding to about 3.2 million ha, the maritime pine is the most important conifer species sharing 23% of total forested area (around 714 000 ha) (ICNF, 2013). In the area of the regional forest management plan of Nordeste (where our study area is located), the maritime pine covers about 43 000 ha of land (AFN, 2010).

3.3. Ecology

The maritime pine is an ecologically versatile species, showing a wide range of expressive traits regarding growth characteristics, frost resistance and adaptation to summer drought and limestone substrates (Abad Viñas et al., 2016). Naturally, it grows in warm temperate regions with an oceanic influence on climate, mainly in humid and sub-humid areas, where annual rainfall is greater than 600 mm. It welcomes mean annual rainfall of 800 mm with at least 100 mm in drought period (Oliveira et al., 2000). In spite of that, it is possible for trees to survive in areas with only 400 mm annual precipitation, providing there is sufficient atmospheric moisture. Maritime pine is not tolerant to shade and exhibits preference for siliceous soils with a coarse texture, especially sandy soils, dunes and other poor substrates (Abad Viñas et al., 2016). It inhabits from sea level in coastal lowlands to moderate elevations in the Iberian Peninsula (1600 m) and inland Corsica, up to around 2000 m in Morocco (Wahid et al., 2006; Farjon, 2010). In Portugal the average conditions for maritime pine include an annual rainfall of 800 mm and an annual mean temperature between 13 and 15 ºC, with the mean temperature of coldest month being 8-10 ºC and the mean temperature of warmest month being equal or inferior to 20ºC (Oliveira et al., 2000).

4. Methodology

4.1. Study area

In order to promote the association for common management of forest areas in Portugal, the Decree-Law 127/2005 of August 5 (altered by Decree-laws 15/2009 from January 14, 2/2011 from January 6, and 27/2014 from February 18) enabled the possibility of creating Forest Intervention Zones (ZIF). Since then many ZIF areas have been constituted in Portugal. From 2005 till the beginning of 2016, 178 ZIF were constituted representing more than 900 000 ha of land and corresponding to 28% of the national forest areas (ICNF, 2016). *Pinus pinaster*, *Quercus suber* and *Eucalyptus globulus* are the main species in ZIF territories. Our study area is the Lomba ZIF, located in northern Portugal in the district of Bragança, municipality of Vinhais, in the parish of Vilar de Lomba e S. Jomil (Figure 5). This ZIF was created in 2008 (ZIF 37, process 154/07-AFN) by the Ordinance 1369, from November 28, Journal of Portuguese Republic 232, Series I. It has a total area of 2142 hectares. In the census of Portugal 2011, the population was 237 in an area of 29.48 km². Lomba ZIF is under the Nordeste PROF as can be observed in Figure 5. The forest management of the ZIF is assured by ARBOREA, an association of forest producers.

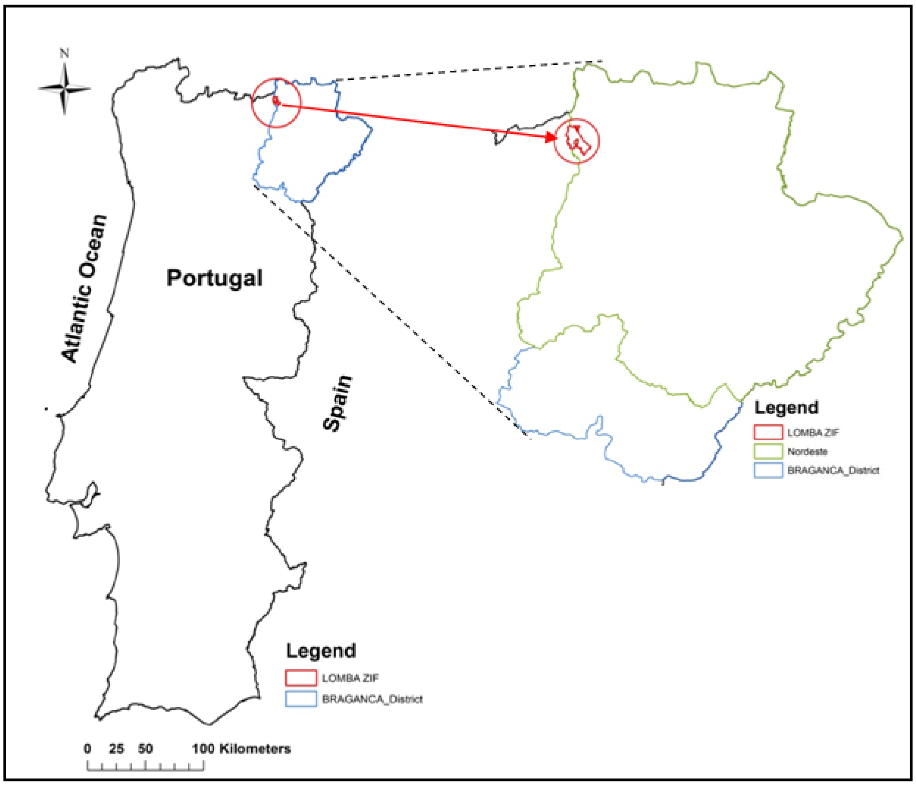


Figure 5. Location of the Lomba ZIF

4.1.1. Climate characteristics

The climate of the Lomba ZIF is similar to the climate of Bragança. It is classified as Csb according to the Köppen classification, temperate with warm dry summer. In the period 1981-2010, the average maximum temperature in January for Bragança was around 8.8 °C (47.8 °F) while in July it was around 29.2 °C (84.5 °F) (IPMA, 2016). The January minimum temperature hovers around the freezing point. It has been known to snow in May, and winter temperatures can fall to as low as -11.6 °C (11.1°F). The annual mean temperature is around 13 °C (55.4 °F). The average of total rainfall was around 800 mm per year in the period 1981-2010 (IPMA, 2016). The year of 2005 was particularly dry in Portugal and Bragança suffered from water shortages and devastating forest fires in the rural areas (IFFN, 2006).

We present here the ombrothermic diagram of Vinhais municipality from the period 1981-2010 (Figure 6) using temperature values from the meteorologic station of Bragança (IPMA, 2016) and precipitation values from meteorologic station (udographic) of Vinhais (APA, 2016a).

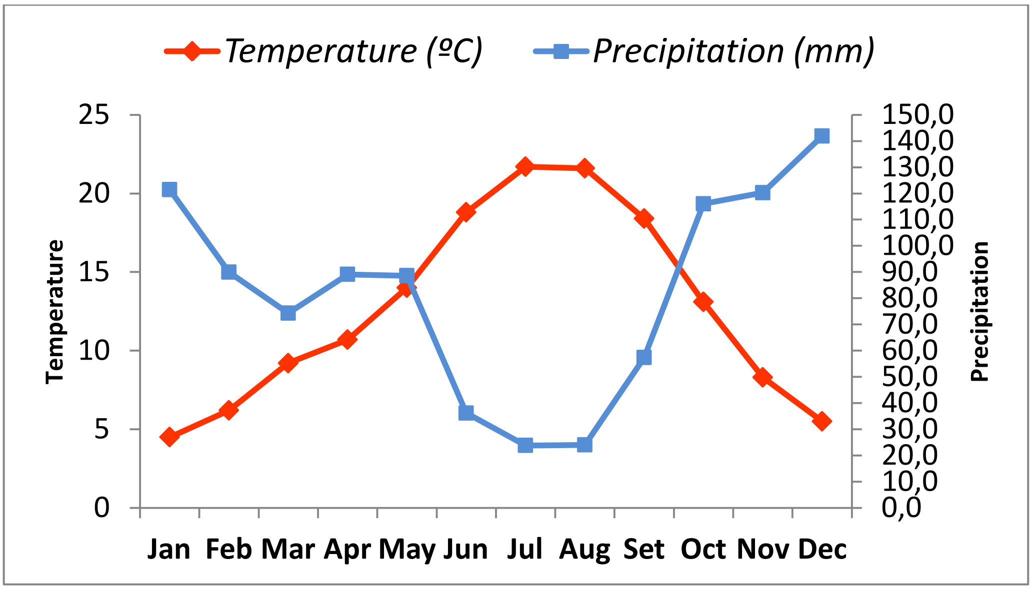


Figure 6. Ombrothermic diagram of Vinhais (period 1981-2010)

Specific climatic characteristics for our study area are following presented, basing on APA (2016b). Insolation and mean annual temperature differs more or less from the north to the extreme southeast as indicated in Figure 7. The real evapo-transpiration is the same in all the area (500 to 600 mm). The frost is lasting almost 80 days in all our area with air humidity equal to 70-75%.

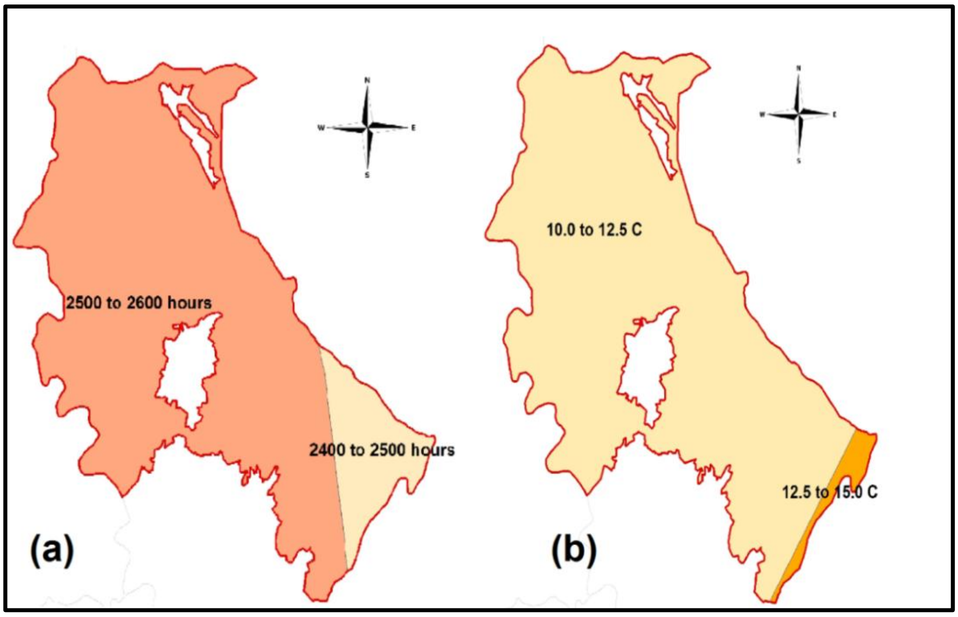


Figure 7. Maps of insolation (a) and mean annual temperature (b) in Lomba ZIF

4.1.2. Soil and physiographic characteristics

The soil type and condition may greatly influence the site productivity. Two types of soil are present in our study area (Figure 8): Cambissols Humic with average pH in the range 4.5-5.5 (acid), occupying only a small area compared to the low acid (pH in the range 5.5-6.5) Litossols Eutric, which are known by their shallow depht (APA, 2016b).

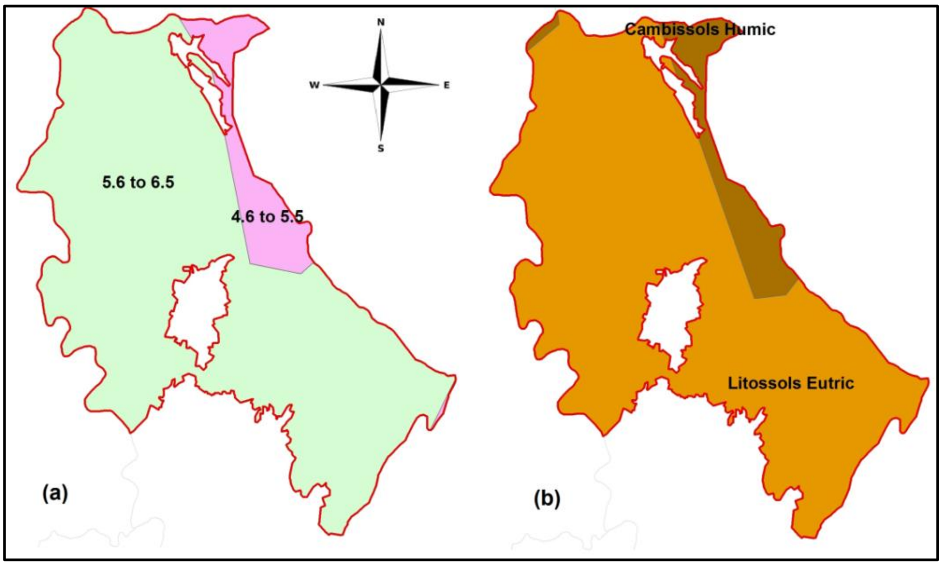


Figure 8. Maps of soil pH (a) and soil type (b) in Lomba ZIF

In Figure 9 a digital model of terrain with elevation and a map of slopes in our study area are shown. Elevation varies from 400 to 900 meters above the sea level and the slopes can reach more than 25 degrees in some areas turned to the rivers Mente and Rabaçal which are limits of the ZIF area by the West and South, respectively. The area is in the Douro Watershed (APA, 2016b).

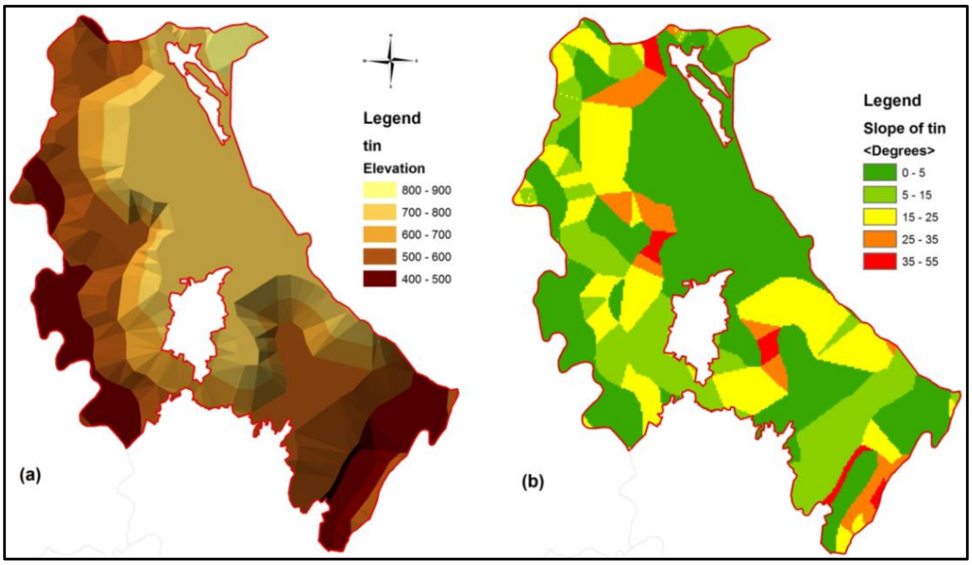


Figure 9.Digital model of terrain (a) and map of slopes (b) in Lomba ZIF

4.1.3. Vegetation

The map of soil use capacity (Figure 10b) shows that our study area is more suited for forest than for agriculture.

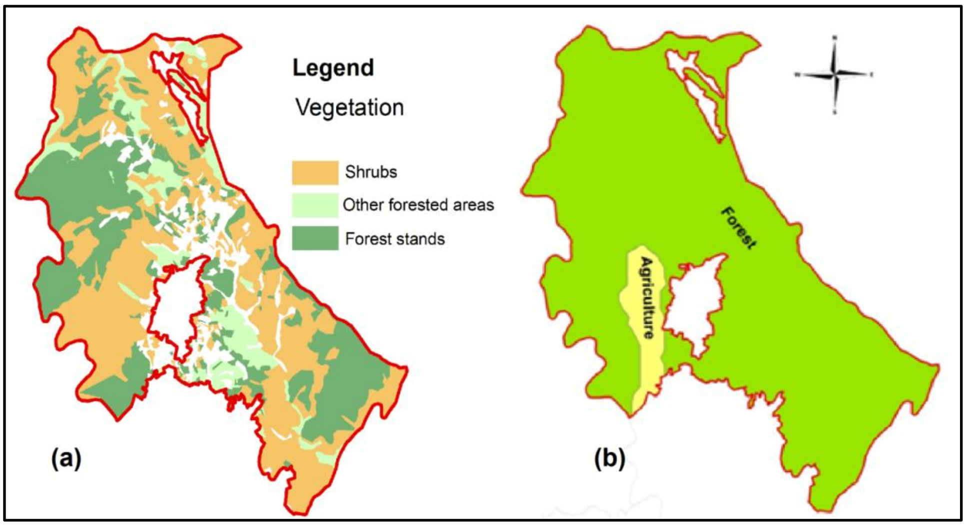


Figure 10. Vegetation map (a) and soil use capacity map (b) in Lomba ZIF

Concerning vegetation cover, and according to the cartography of the Forest Management Plan (PGF) for the Lomba ZIF, shrubs occupy more or less half of the ZIF area. The other half is covered with forest stands. The main forest species are the maritime pine and sweet chestnut (*Castanea sativa* Mill.), the last one mainly for fruit production. Other forested areas are covered with diverse broadleaves and low density *Pinus pinaster* stands. The area of maritime pine is 662 hectares, consisting mainly in pure regular stands (524 hectares). About 138 hectares were classified as other forested areas. Figure 11 shows the map with the distribution of the maritime pine in the Lomba ZIF (PGF 2008, Courtesy of ARBOREA).

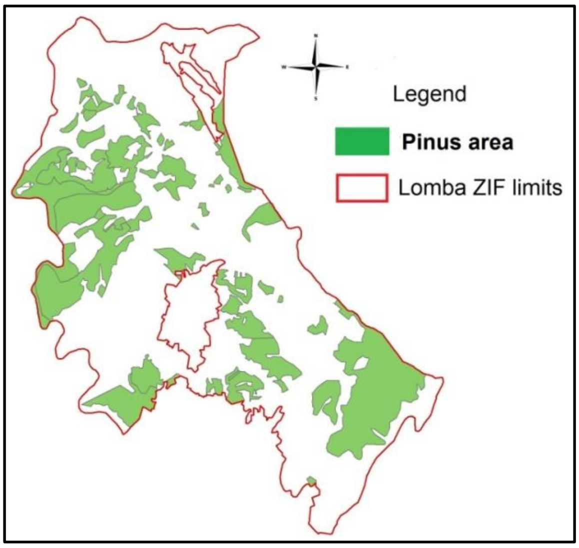


Figure 11. Distribution of the maritime pine in the Lomba ZIF

4.2. Data and measurements

In mid-November 2015 under a workshop, five circular 500 m^2 plots were established in maritime pine stands of the Lomba ZIF, to demonstrate the use of terrestrial LiDAR in forest inventory. These plots were located according to the coordinates of the systematic 2 x 2 km plots from Portuguese national forest inventory falling inside Lomba ZIF. Because some of these coordinates were coincident with clearings in the middle of or very close to forest areas, for those cases we decided to move the location just to the nearest area covered with forest. One of the plots was just over the limit of the ZIF and another one was located out the limit but less than 2 km of the ZIF border, thus being representative. At that time the maritime pine areas of the ZIF were visited, to understand actual conditions after recurrent wild fires in the region. It was observed that there are many areas of natural regeneration with very young stands of maritime pine that do not have attained yet the minimum sizes to be inventoried. During the field visits we registered on Google Earth maps the maritime pine stands that could be possible measured according to the objective of this study. Then we randomly allocated additional plots to maritime pine stands of Lomba ZIF covering different combinations of ages and densities. A total of 18 circular 500 m^2 plots were measured in this study, which are presented in Figure 12.

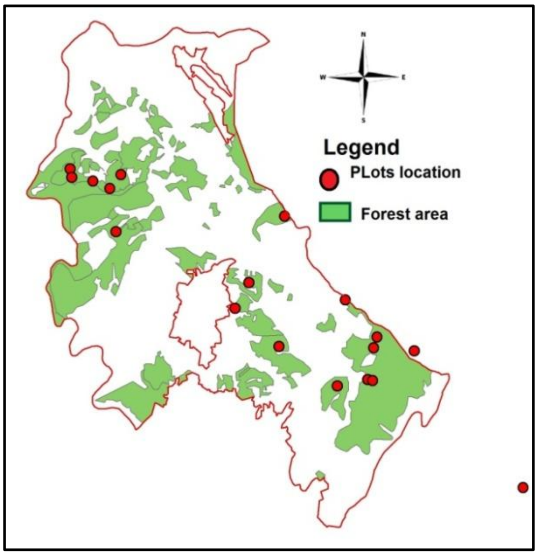


Figure 12. Map with plots location

In each plot, the coordinates of the centre were registered using a Trimble GeoXM GPS. The GPS was also used for plot location. The slope was measured in degrees with the Haglöf Vertex IV using the angle menu. This measurement was necessary to correct the radius of the circular plot for gradients higher than 5º (the radius is 12.62 m without correction), according to the formula:

——— , (1)

where r is the radius of the plot in meters and α is the slope in degrees (º).

All the trees inside the plot were numbered with appropriate spraying paint. In plantation stands the numbering was made in order that the first tree was the one on the upper left or most north and west point of the plot, progressing along the planting lines and in serpentine (Figure 13a). In the other cases, the first tree was the one on the right and more close to the north direction (azimuth equal to 0) given by a compass, progressing then according to the movement of the clock pointers 360º (Figure 13 b). The limits of the plot were defined with the help of the Vertex IV and the corresponding transponder that was fixed in a stick in the plot centre for controlling the distances.

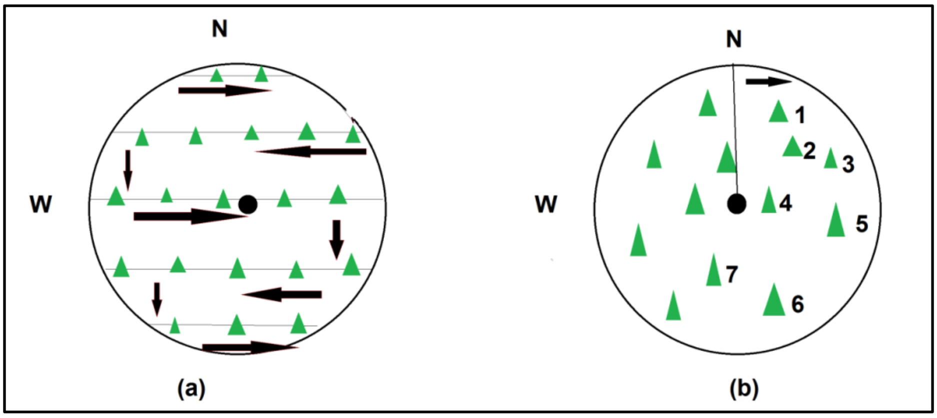


Figure 13. Illustration of tree's numbering inside the plots

Measurements of diameter at breast height (d) in centimeters, and total tree height (h) in meters were taken in all the trees from the plots. Diameter was measured using a Haglöf Mantax Caliper. Two measurements were taken, the first one pointing the main axis of the caliper to the plot centre and the second in the transverse direction. The average of these two measurements defines *d*. Height was measured with Vertex IV.

Stand variables where computed after, in the office, using information of individual-tree measurements. Dominant height (hdom) and dominant diameter (ddom) where computed as the average height and average diameter of the 100 thickest trees per hectare (5 trees in our plots). Number of living trees per hectare (Nv) was obtained by expanding the number of living trees inside the plot to the hectare. This expansion was also done to obtain basal area per hectare (G) after computing basal area of the plot (g) from individual-tree diameter (d_i) as:

– **(2)**

Quadratic mean diameter (dg) was obtained by the expression:

(3)

Site Index (*SI*) was obtained by using Tomé (2001) equation. Age of the stand (*t*) was obtained in the field using a Haglöf Increment Borer in two of the dominant trees.

4.3. Model fitting and selection

4.3.1. Data preparation and outliers removal

Before proceeding with model fit, an initial exploratory data analysis was made, looking for possible measurement errors or other anomalies in the data. Scatter plots of the *h-d* relationship were observed for each sample plot. Also an approach, similar to the one proposed by Bi (2000) for detecting abnormal data points, was applied. This procedure has been widely used in the framework of stem taper equations (e.g., Rojo et al., 2005; López-Sánchez et al., 2016) and has been used recently for *h-d* equations by Corral-Rivas et al. (2014). A local regression curve with a smoothing parameter of 0.25 was fitted for each plot and also to the total dataset. The *loess* function from package *stats* of statistical program R (R Development Core Team, 2016). These procedures led to the identification and removal of outlying observations representing 2% of the total initial data. Due to the very low correlation between *h* and *d* observed in one plot, it was decided to exclude it from the posterior analysis. Finally, the available dataset for model fit and selection contained 1087 observations of *h-d* pairs on 17 plots. The summary statistics for this dataset are presented Table 1.

Table 1. Summary statistics of the dataset for model fitting and selection

	Tree level (n=1087)		**Plot level (n=17)**						
	h (m)	*d* (cm)	*hdom* (m)	*ddom* (cm)	*dg* (cm)	N_v (trees ha^{-1})	*G* (m^2 ha^{-1})	*t* (years)	*SI* (m)
Mean	11.0	16.3	12.5	24.7	17.9	1336	31.6	25.5	18
Max	20.9	33.8	19.1	32.5	24.1	2880	56.5	40.0	20
Min	5.6	6.5	8.7	18.1	13.0	380	8.0	16.0	14
SD	2.8	5.4	3.0	4.2	3.8	709	13.9	7.9	2

SD – standard deviation

4.3.2. Local models

Several local models were tested, including power, exponential and hyperbolic equations, as well as sigmoidal equations, based on Huang et al. (1992; 2000), Fang and Bailey (1998) and Leduc and Goelz (2009). Not all models in these references were tested because that was not our objective and moreover, many of them are derivations of the same basic model as pointed in Leduc and Goelz (2009). Only models with 3 or less parameters where tested, which are presented in Table 2.

Table 2. Local *h-d* models tested in this study

Two-parameter equations	**References**
[L1]	Stoffels and Van Soest (1953)
[L2]	Wykoff et al. (1982)
[L3]	Burkhart and Strub (1974)
[L4]	Meyer (1940)
[L5]	Bates and Watts (1980)
[L6]	Curtis (1967)
[L7]	Huang et al (2000)
Three-parameter equations	**References**
[L8]	Richards (1959)
[L9]	Weibull-type (Yang et al, 1978)
[L10]	Gompertz (Zeide, 1993)
[L11]	Korf (Zeide, 1993)
[L12]	Curtis et al. (1981)
[L13]	Pearl and Reed (1920)
[L14]	Ratkowsky (1990),
[L15]	Ratkowsky and Reedy (1986)

The equations were fitted in R by nonlinear least squares using the *nls* function of package *stats*. Each equation was fitted plot by plot. To evaluate the goodness-of-fit, the root mean square error (RMSE), the bias and the fit index (FI_{adj}), a statistic similar to the adjusted coefficient of determination in linear regression were used. Also the Akaike Information Criterion (AIC; Akaike, 1974) and the Bayesian information criterion (BIC; Schwarz, 1978) were computed as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2}{n-k}} \qquad \textbf{(4)}$$

$$Bias = \frac{\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)}{n} \qquad \textbf{(5)}$$

$$FI_{adj} = 1 - \frac{(n-1)\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2}{(n-k)\sum_{i=1}^{n} \left(h_i - \bar{h}\right)^2} \qquad \textbf{(6)}$$

$$AIC = n\ \ln(n)\left(\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2/(n-k)\right) + 2k \qquad \textbf{(7)}$$

$$BIC = n\ \ln(n)\left(\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2/(n-k)\right) + k\ln(n) \qquad \textbf{(8)}$$

In the above equations h_i, and are the observed heights, the estimated heights and the mean of the observed heights, respectively; *n* is the number of observations used in the fitting; *k* is the number of parameters in the equation, and *ln* is the natural logarithm. These statistics were averaged for all the plots in order to obtain a unique value for model comparison.

4.3.3. Generalized models

Several generalized *h-d* equations were selected from literature (Soares and Tomé, 2002; López Sánchez et al., 2003; Crecente-Campo et al., 2010; Corral-Rivas et al., 2014; Gómez-García et al., 2015). Some adaptations were also tried and additionally it was decided not to use age as explanatory variable in the models tested (Table 3).

Table 3. Generalized h-d models tested in this study

One-parameter equations		**References**
[G1]		Canãdas et al. (1999)
[G2]		Canãdas et al. (1999)
[G3]		Canãdas et al. (1999)
Two-parameter equations		**References**
[G4]		Sloboda et al. (1993) Adapted
[G5]		Canãdas et al. (1999) Adapted
[G6]		Crecente-Campo et al. (2010) (Gaffrey, 1988 adapted)
Three-parameter equations		**References**
[G7]		Gómez-García et al. (2015) (Tomé, 1989, adapted)
[G8]		López Sánchez et al (2003) (Castedo et al., 2001 adapted)
[G9]		Harrison et al. (1986)
[G10]		Crecente-Campo et al. (2010) (Harrison et al., 1986 adapted)
[G11]		Crecente-Campo et al. (2010) (Pienaar et al., 1990 adapted)
[G12]		Stoffels and Van Soest modified by Tomé (1989) Adapted
[G13]		Prodan modified by Tomé (1989) Adapted
[G14]		Crecente-Campo et al. (2010) (Krumland and Wensel, 1988 adapted)
[G15]		Richards (1959) adapted

(continue)

Table3. Generalized h-d models tested in this study (*continue)*

Four-parameter equations		**References**
[G16]		Schröder and Álvarez (2001) (Mirkovich, 1958 adapted)
[G17]		Hui and Gadow (1999)
[G18]		Cox (1994)
[G19]		Harrison et al. (1986) adapted
[G20]		Adame et al. (2008)

The generalized equations were also fitted by nonlinear least squares using the *nls* function of package *stats* from R using all the data. Similarly to local models, the goodness-of-fit was evaluated by using the root mean square error (RMSE), the bias, the fit index (FI_{adj}), *AIC* and *BIC*. For these models the *PRESS* residuals were obtained using function *PRESS* from the package *qpcR* in order to get an idea of the predictive capacity of the models. The mean of *PRESS* residuals (Mpr), the mean of absolute PRESS residuals (MApr), and the model efficiency (EF) were computed as:

$$Mpr = \frac{1}{n}\sum_{i=1}^{n}\left(h_i - \hat{h}_i^*\right) \qquad \textbf{(9)}$$

$$MApr = \frac{1}{n}\sum_{i=1}^{n}\left|h_i - \hat{h}_i^*\right| \qquad \textbf{(10)}$$

$$EF = 1 - \frac{\sum\left(h_i - \hat{h}_i^*\right)^2}{\sum\left(h_i - \bar{h}\right)^2}, \qquad \textbf{(11)}$$

where h_i is a real observation and $\hat{h}_i^*$ is the corresponding predicted value calculated with the model fitted with the observation *i* deleted from the original dataset. *Mpr* gives an indication of the bias in predictions and *MApr* gives an indication of the error that in average is produced when using the model for prediction.

4.3.4. Mixed models

The best local and generalized equations were fitted as mixed models and tested. For each type of model, initial fit was made with all the parameters expanded with random effects and in case of non-convergence combinations of two or one parameter as random effects were tried (where applicable). As suggested by Pinheiro and Bates (1998) and Fang and Bailey (2001) all parameters in the model should be considered mixed if convergence is possible. In the modeling procedure just one hierarchical level (trees in plots) was considered because no repeated measurements from the same plot were used. Thus, the variability between the sampling units can be explained by including random parameters, which are estimated at the same time as the fixed parameters.

Basically, the parameter vector of a non-linear mixed model can be defined as follows (Pinheiro and Bates, 1998):

(12)

where $\boldsymbol{\Phi j}$ is the parameter vector $r \times 1$ (where r is the total number of parameters in the model) specified for the j^{th} plot, $\boldsymbol{\lambda}$ is the vector p × 1 of the common fixed parameters for the whole population (p is the number of fixed parameters in the model), $\boldsymbol{b_j}$ is the vector $q \times 1$ of the random parameters associated with the j^{th} plot (q is the number of random parameters in the model), $\boldsymbol{A_j}$ and $\boldsymbol{B_j}$ are matrices of size $r \times p$ and $r \times q$ for specific and random effects for the j^{th} plot, respectively. The random effects are assumed to follow a multivariate normal distribution with mean zero and a positive-definite variance-covariance matrix $\boldsymbol{D}$. The error term of a mixed-effects model is assumed to follow a multivariate normal distribution with mean zero and a positive-definite variance-covariance matrix $\boldsymbol{R_j}$. The residual vector represents within subject (plot in this study) variability and the random effects vector represents between subject variability. Detailed information about nonlinear mixed-effects modeling for *h-d* relationships can be found in Calama and Montero (2004) and Castedo Dorado et al. (2006). Application of mixed-effects modeling involves three steps (Yang and Huang, 2013): (i) estimating model parameters, (ii) predicting random effects, and (iii) making subject-specific predictions. The first step corresponds to the fitting phase while the second and third are known as calibration and subject-specific prediction respectively.

To the fit procedure the first-order conditional expectation (FOCE) approximation of Lindstrom and Bates (1990) in which random effects are set to the current estimated best linear unbiased predictor (EBLUP) was used. The maximum likelihood (ML) was used to compare the models by goodness-of-fit, given that it provides asymptotic efficient estimates, whereas the restricted maximum likelihood (REML) was used to obtain final parameter estimates, as it yields unbiased estimates of variance components (Litell et al., 2006).The mixed-effects modeling framework enables localization of the *h-d* function to a new plot with at least one additional height measurement. This process is known as calibration and involves prediction of random effects (Arias-Rodil et al., 2015). In the calibration the FOCE method was used in order to be consistent with the expansion method used in the fitting step. For the FOCE method, the random effects for a plot ( ) are calculated as follows (Vonesh and Chinchilli, 1997):

, **(13)**

where   is the estimated variance-covariance matrix for the random effects $\boldsymbol{b_j}$,   is the estimated variance-covariance matrix for the error term, $\boldsymbol{Z_j}$ is the partial derivatives matrix with respect to random effects. Taking into account that   appears on both sides of equation 13, the calibration process must be solved iteratively (Lindstrom and Bates, 1990). After obtaining , the function can be localized at plot level, by adding plot-level random effects to the corresponding fixed parameters. In this study we used the *nlme* package from R for model fitting. The calibration was also implemented in R adapting the script of Arias-Rodil et al. (2015) for the purpose. Concerning the trees to use in calibration, three approaches were compared: (i) total height of 1-10 smallest diameter trees per plot; total height 1–10 average-size trees per plot (closest to dg); Total height of 1-10 quantile trees of the diameter distribution per plot-inventory (as in Gómez-García et al., 2015, i.e., 1 tree = median-diameter tree; 2 trees = tercile-diameter trees; 3 trees = quartile-diameter trees; 4 trees = quintile-diameter trees, etc.).

5. Results and Discussion

5.1. Local *h-d* models

Upon fitting the local *h-d* equations by each individual plot it was verified that models with three parameters recurrently presented no-significant parameters and lack of convergence in some situations. This was systematically observed for all the equations with three parameters tested. Concerning the models with two parameters, at the exception of model L5 from Bates and Watts (1980), all the coefficients in the equations were significant. The three models with best goodness-of-fit statistics are identified in bold in Table 4.

Table 4. Goodness-of-fit statistics for the local *h-d* models

Model	Statistics	*RMSE*	*Bias*	FI_{adj}	*AIC*	*BIC*
[L1]	Mean	**0.807**	**-0.0003**	**0.570**	**153.7**	**159.7**
	Min	0.546	-0.0017	0.281	36.6	39.4
	Max	1.293	0.0033	0.847	439.0	447.9
	SD	0.254	0.0011	0.147	99.9	101.4
[L2]	Mean	**0.808**	**0.0008**	**0.571**	**153.9**	**159.9**
	Min	0.549	-0.0012	0.268	35.0	37.8
	Max	1.272	0.0105	0.799	440.5	449.4
	SD	0.243	0.0027	0.138	98.5	100.0
[L3]	Mean	0.810	0.0009	0.569	154.2	160.2
	Min	0.546	-0.0011	0.266	34.8	37.6
	Max	1.271	0.0116	0.786	441.3	450.2
	SD	0.242	0.0029	0.136	98.4	99.9
[L4]	Mean	**0.805**	**0.0017**	**0.572**	**153.6**	**159.6**
	Min	0.545	-0.0039	0.254	34.8	37.6
	Max	1.271	0.0182	0.828	442.5	451.4
	SD	0.245	0.0052	0.143	99.0	100.5
[L5]	Mean	0.805	0.0006	0.573	153.5	159.5
	Min	0.547	-0.0026	0.270	35.4	38.3
	Max	1.277	0.0121	0.833	439.6	448.6
	SD	0.248	0.0034	0.143	99.0	100.5
[L6]	Mean	0.809	0.0008	0.570	154.0	160.0
	Min	0.547	-0.0012	0.267	34.9	37.7
	Max	1.272	0.0110	0.792	440.9	449.8
	SD	0.243	0.0028	0.137	98.4	100.0

Continue

Table 4 (*cont.*): Goodness-of-fit statistics for the local *h-d* models

Model	**Statistics**	***RMSE***	***Bias***	FI_{adj}	***AIC***	***BIC***
[L7]	Mean	0.811	0.0047	0.566	154.7	160.7
	Min	0.535	-0.0043	0.213	34.0	36.9
	Max	1.270	0.0237	0.823	450.3	459.2
	SD	0.245	0.0076	0.147	100.2	101.7

As can be observed in Table 4, the performance of the three best models was very similar. No one of the models was far superior to the others including the performance of L3, L5, L6 and L7 that was also very close to the remaining. Nevertheless, analyzing the values of the goodness-of-fit statistics, the model presenting lower *RMSE* and lower *AIC* and *BIC* was L4, the model from Meyer (1940). Model L5 presented similar performance but presented no-significant parameters in the fit procedure. All the models presented low Bias. Our results show that three parameter models and namely sigmoidal models are not necessarily the ones that better fit *h-d* data. We do not have a considerable large range in our data (from very young to very old stands) and a sigmoidal-shaped tendency was not necessary to model *h-d* relationship. This is in accordance to what was observed and stated by other authors (e.g., Crecente-Campo et al., 2010; Paulo et al., 2011; Burkhart and Tomé, 2012). The evidence of no clear superiority of a particular model form as been referred (e.g., Mehtätalo, 2004) and our results seems to corroborate this.

In this study, the equation L4 from Meyer (1940) was selected as the best local *h-d* equation. A modification of this exponential model was found to have a suitable behavior by Fang and Bailey (1998). The Meyer (1940) equation is flexible and, for example, was used by Harrison et al. (1986) to deduce their *h-d* generalized equation (Soares and Tomé, 2002). The results from the fit of model L4 to the total dataset follow:

, **(14)**

with $FI_{adj}=0.42$; $RMSE=2.15$ and all parameters significant ($p<0.001$).

When fitting model L4 to each individual plot as well as to all the dataset, no evidence of clear violation of the assumptions of homogeneous variance of the errors and normality of the errors were observed (Figure 14). Thus, no additional measures were adopted after fitting the model by nonlinear lest squares.

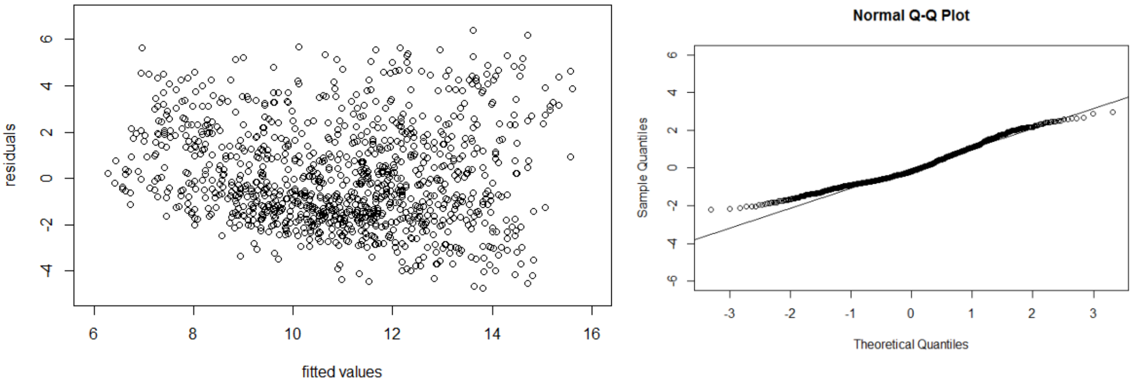


Figure 14. Plot of residuals Vs predicted values and normal Q-Q plot of residuals from fitting model L4 to total dataset.

The modest fit of L4 to the total data set was expected and is related to the variability of stand characteristics. It can be observed that even in a relative small area like in a ZIF (about 2000 ha in Lomba ZIF) variability in stand conditions may hinder the generalized use of a local *h-d* equation. The relationship between the parameters of Meyer (1940) equation and the stand variables were analyzed. This is useful to get a picture of the variability and also to get insight about the possibility to expand the parameters of L4 equation with stand variables to possible develop generalized equations. Interesting relationships were particularly observed between dominant height (hdom) or dominant diameter (ddom) and parameter b_1 (Figure 15A) as well as between the number of living trees per hectare (Nv) and parameter b_2 (Figure 15 B).

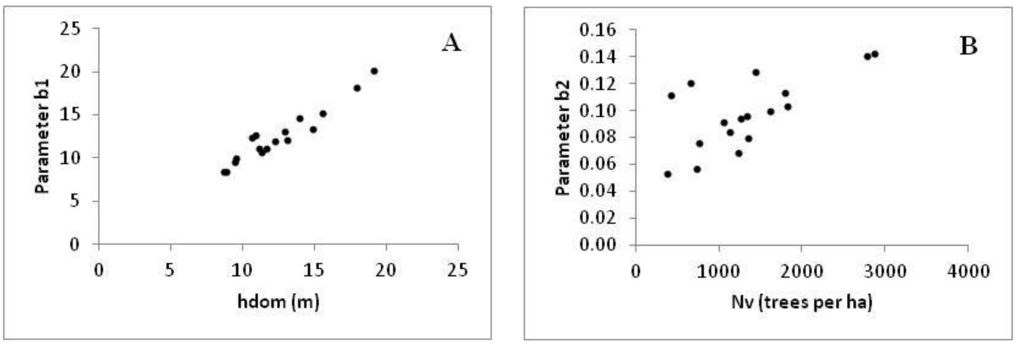


Figure 15. Observed relationship between parameter b1 of model L4 and hdom (A) and between parameter b2 of L4 model and Nv (B).

5.2. Generalized *h-d* models

The results of the nonlinear least squares fitting of the generalized *h-d* equations are presented in Table 5. As before, the three best performing models are identified in bold.

Table 5. Goodness-of-fit and PRESS statistics for the generalized *h-d* models

Model	FI_{adj}	*RMSE*	*Bias*	*AIC*	*BIC*	*Mpr*	*MApr*	*EF*	*Coef*
[G1]	0.889	0.941	0.0372	2956.2	2966.2	0.037	0.743	0.889	1
[G2]	0.879	0.982	0.0122	3047.6	3057.5	0.012	0.771	0.879	1
[G3]	0.885	0.957	-0.0418	2992.6	3002.6	-0.042	0.754	0.885	1
[G4]	0.859	1.061	-0.0123	3216.9	3231.9	-0.013	0.820	0.858	2
[G5]	0.890	0.937	0.0473	2946.8	2961.8	0.047	0.737	0.889	2
[G6]	0.885	0.959	0.0429	2997.6	3012.5	0.043	0.754	0.884	2
[G7]	**0.896**	**0.909**	**-0.0033**	**2881.5**	**2901.5**	**-0.003**	**0.711**	**0.896**	**3**
[G8]	0.849	1.096	-0.0010	3289.9	3309.8	-0.002	0.854	0.848	3
[G9]	0.873	1.005	0.0052	3100.9	3120.9	0.005	0.794	0.873	3
[G10]	0.889	0.942	-0.0124	2960.3	2980.2	-0.012	0.746	0.888	3
[G11]	0.882	0.968	-0.0571	3020.1	3040.1	-0.057	0.768	0.882	3
[G12]	0.893	0.922	0.0678	2913.4	2933.3	0.067	0.721	0.893	3
[G13]	**0.896**	**0.910**	**0.0236**	**2884.4**	**2904.4**	**0.024**	**0.713**	**0.896**	**3**
[G14]	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	3
[G15]	**0.896**	**0.911**	**0.0374**	**2887.8**	**2907.7**	**0.037**	**0.714**	**0.895**	**3**
[G16]	0.889	0.942	-0.0007	2961.4	2986.3	-0.001	0.741	0.888	4
[G17]	0.881	0.974	-0.0002	3033.8	3058.7	-0.001	0.768	0.880	4
[G18]	0.721	1.490	0.0014	3958.3	3983.2	0.002	1.104	0.720	4
[G19]	0.885	0.957	0.0083	2996.3	3021.3	0.009	0.750	0.884	4
[G20]	0.889	0.940	0.0000	2956.2	2981.1	0.000	0.737	0.888	4
a – presented convergence problems									
[G21]	0.890	0.937	0.0009	2949.2	2969.2	0.001	0.741	0.889	3

Several formulations of generalization of Meyer (1940) equation were tried by expanding parameters with stand variables. One of these formulations, presenting good performance is presented also in Table 5 (model G21), having the following mathematical expression:

(15)

The inclusion of stand variables such as dominant height (hdom), dominant diameter (ddom) and number of living trees per hectare (Nv) improved the fit and predictive capacity of the models. The adjusted fit index (FI_{adj}) which, in average, was around 0.57 for local models, now increased to 0.88. As pointed by Zeide and Vanderschaaf (2002), Staudhammer and Lemay (2000) and verified by many other authors, inclusion of density and other variables of the stand in *h-d* equations, improve predictions. For example, Calama and Montero (2004) refer that density and dominant height are usually positively correlated with height. Dominant height correlation means that a relation exists between site index and stand height and the relation between density and height is shown by the fact that for the same height, trees located in denser stands tend to have higher height/diameter ratio than those located in less-occupied stands. Model G8 presented no-significant parameters and model G14 revealed convergence problems and were discarded. It was also observed that models with four parameters did not perform better than three parameters models. The three best models were G7 from Gómez-García et al. (2015), G13 that is an adaptation of Prodan (1965) model modified by Tomé (1989), and G15 which is also an adaptation of Richards (1959) model. These three models presented very similar performance with only slight superiority of model G7 from Gómez-García et al. (2015). This was the generalized *h-d* model selected. The model is simply an adaptation from a modified form of Michailoff (1943) model proposed by Tomé (1989), restricting the Michailoff's model to pass in the point (ddom, hdom). The Parameters of G7 model are presented in Table 6.

As a registry note, we refer two additional models that presented inferior performance than model G7 but may be interesting when possible lack of information about stand density or only information on dominant height exist. These two models are G1 from Canãdas et al. (1999) with just one parameter and *hdom* and *ddom* as regressors and G17 from Hui and Gadow (1999), a four parameters model having only *hdom* as explanatory variable in addition to *d*. Parameters of these two models are also presented in Table 6 (standard errors in parenthesis).

Table 6. Parameters of the best model (G7) and from models G1 and G17

Model	b_1	b_2	b_3	b_4
G7	-5.76514 *** (0.56001)	-0.24790 *** (0.03936)	1.49714 *** (0.12704)	
G1	0.362383 *** (0.00628)			
G17	0.15614 ** (0.05657)	1.26500 *** (0.14000)	0.86210 ** (0.32956)	-0.37226 * (0.14858)

*** ($p<0.001$); ** ($p<0.01$); * ($p<0.05$)

The scatter plot of residuals and predicted values as well as the normal Q-Q plot of the residuals of model G7 are presented in Figure 16. Again, the analysis of the mentioned plots did not revealed clear violation of the assumptions of homocedastic and normal distributed errors. So, no additional remedial measures were adopted.

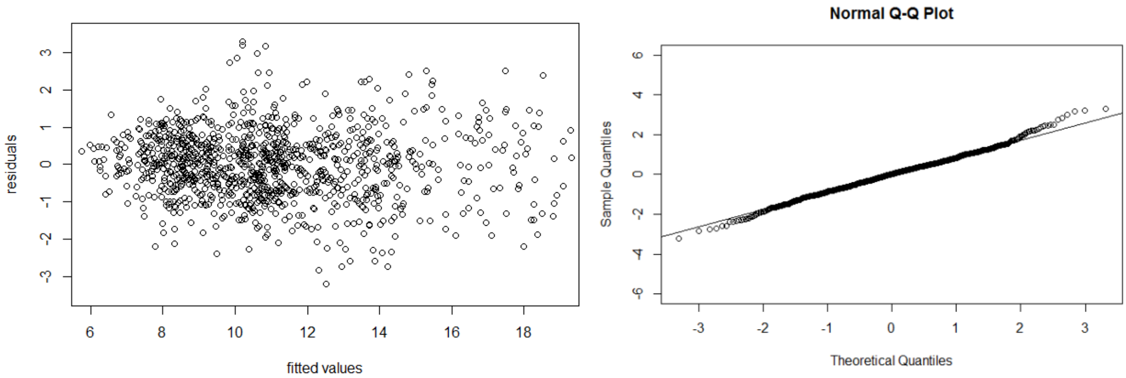


Figure 16. Plot of residuals Vs predicted values and normal Q-Q plot of residuals from fitting model G7 to total dataset

An evaluation of the *Mpr* and *MApr* by diameter class was further undertaken for model G7 in order to better analyze the behavior of the model when predicting heights of trees in different size classes (Figure 17).

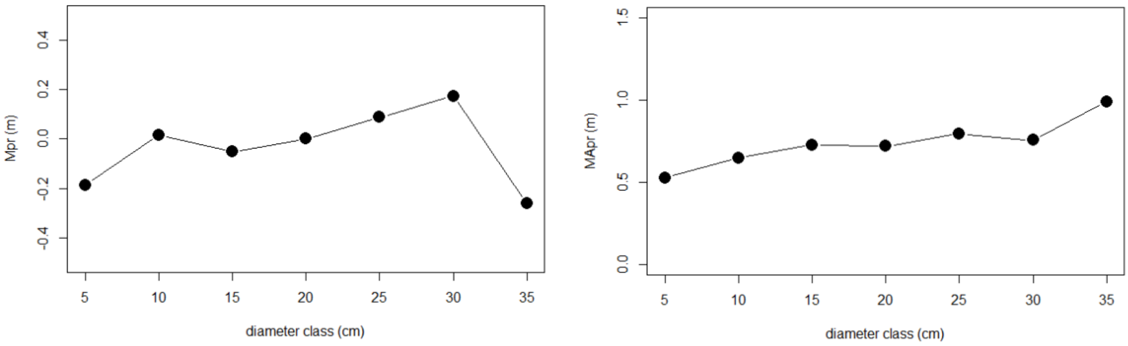


Figure 17. Mean (left) and mean of absolute (right) PRESS residuals by diameter class

Concerning the bias (Mpr), model G7 overestimates height in diameter class 5 cm by 0.2 m, in average. In diameter classes 10 to 25 cm, the model presents very low bias. In diameter class of 30 cm there is a tendency to underestimate height by about 0.2 m. With the due caution of diameter class 35 cm being low represented in the dataset, an overestimation of height by 0.2 to 0.3 m is observed. The error that in average is committed when using the model (given by MApr) stays between 0.5 to 0.7 m in the diameter class range of 5-30 cm, approaching 1.0 m in the diameter class 35 cm.

Finally we want to refer that was our aim from the start of this study to obtain a generalized *h-d* equation not having age as explanatory variable. Nevertheless, an exploratory analysis revealed that should the age be included, a model similar to G7, with additional inclusion of *t* as regressor, presented only a very small improvement to model G7. We think that in this case, the effort to accurately measure age and include it in the model is not worthy.

5.3. Mixed models approach

The selected local (L4) and generalized (G7) *h-d* equations were fitted as mixed models. Both the models were initially fitted by allowing all the parameters to include fixed and random effects. In the case of basic model from Meyer (1940), convergence was obtained but for the generalized model G7, containing three parameters, this was not verified. Then, for this particular model, all combinations of two parameters expanded with random effects were tested and the inclusion of random effects in the fixed parameters b_1 and b_3 produced the best results. This analysis was done using the maximum likelihood (ML) method to compare the models by goodness-of-fit, given that it provides asymptotic efficient estimates. Then, final parameter estimates were obtained with the restricted maximum likelihood (REML) method, as it yields unbiased estimates of variance components (Litell et al, 2006). In the fit of mixed model version of G7 by REML, convergence was not achieved. It forced to another round of tests considering only one parameter expanded with random effects. Finally, the version with b_1 with fixed and random effects and the other parameters fixed, generated the best results and thus, was adopted. The final local (M1) and generalized (M2) mixed models were, respectively:

[M1] ; **(16)**

[M2] , **(17)**

Estimates for the fixed parameters and variance components for the random effects (*u* and *v*) of models M1 and M2 are presented in Tables 7 and 8, respectively.

Table 7. Fixed parameters estimates and variance components for model M1

Fixed effect: Parameter	**Estimate**	**Approx. std error**	***t-value***	**Approx. *p-value***
b_1	12.360898	0.749367	16.49	< 0.001
b_2	0.097136	0.006625	14.66	< 0.001
Random effect: Variance component	**Estimate**	**Standard deviation CI (95%)** *estimate*	*lower*	*upper*
var(u)	9.045559			
sd(u)		*3.007416*	*2.113655*	*4.279106*
var(v)	0.000551			
sd(v)		*0.023468*	*0.015356*	*0.035865*
cov(u,v)	-0.028585			
σ^2 (error variance)	0.720538			
σ		*0.848845*	*0.813364*	*0.885874*

Table 8. Fixed parameters estimates and variance components for model M2

Fixed effect: Parameter	**Estimate**	**Approx. std error**	***t-value***	**Approx. *p-value***
b_1	-5.497257	1.651203	-3.33	< 0.001
b_2	-0.262210	0.115471	-2.27	< 0.01
b_3	1.516130	0.503056	3.01	< 0.05
Random effect: Variance component	**Estimate**	**Standard deviation CI (95%)** *estimate*	*lower*	*upper*
var(u)	1.613510			
sd(u)		*1.270240*	*0.814820*	*1.980203*
σ^2 (error variance)	0.760238			
σ		*0.871916*	*0.835675*	*0.909728*

From Tables 7 and 8 it can be observed that all the estimates were significant at 5% level (*p-value* for the fixed parameters and 95% confidence intervals for random effects). For the prediction of individual-tree heights within a stand with inclusion of random parameters in *h-d* equations, two possibilities exist (Vonesh and Chinchilli, 1997; Fang and Bailey, 2001; Fang et al., 2001): i) predict a population mean response when only diameters are measured (and the stand variables are included in the model in the case of generalized models). In this case, the vector of random parameters is assumed to have an expected value of E($\boldsymbol{b_j}$) = $\mathbf{0}$ (using only the fixed parameters); and ii) predict a calibrated response, when the height of a subsample of n_j trees is measured along with diameter measurement in each new plot j (and the stand variables in the case of generalized models). This subsample is subsequently used to calculate the specific random parameters (   ) of the new sampling units (Calibrated Response).

For the calibration of the basic and generalized mixed models, the first-order conditional expectation (FOCE) approximation of Lindstrom and Bates (1990) was implemented in R (R Development Core Team, 2016). Concerning the three strategies analyzed, we applied it to all the dataset firstly. By doing so, it was verified that the total height of 1-10 average-size trees per plot did not produced adequate results both for the local and generalized mixed models.

For the generalized mixed model, calibrating with smallest trees and quantile trees produced very similar results (except when calibrating with just one tree) (Figure 18). The results obtained with the inclusion of random effects in model (M2) were not significantly different from the results obtained with using only the fixed parameters (average response) nor were different from the nonlinear lest squares fit (Figure 18, Left). For the local mixed model (M1), the calibration with quantile trees produced the best results. The higher RMSE values comparing with the generalized model were perfectly expected as we were fitting a local model to all the data. The most important thing to note however is that in relative terms, the RMSE decreases more in the local than in the generalized model when random effects are included in comparison to the average (Nlfixed) response. This means that it makes difference when considering or not random effects in the local model. However this is not evident for the generalized model.

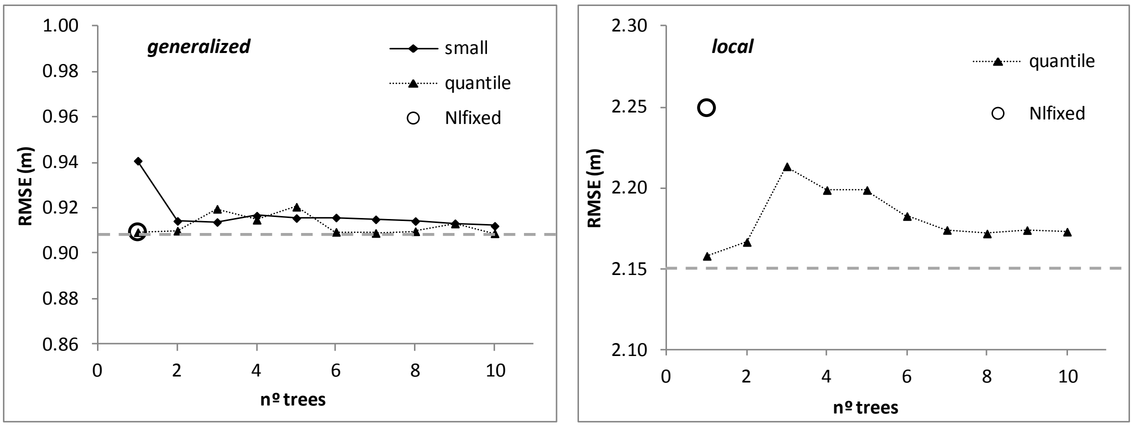


Figure 18.Left: calibrated response of the generalized mixed model (M2) with 1-10 smallest trees and 1-10 quantile trees (dashed thick line represents the simple nls fit). Right: calibrated response of the local mixed model (M1) with 1-10 quantile trees (dashed thick line represents the best possible fit using all trees in calibration). Nlfixed represents the average response, i.e., using only fixed effects with all data.

To get more insight into the question, calibration was made for two plots which are representative of many of the maritime pine stands inside the Lomba ZIF. The characteristics of these plots are represented in Table 9 and Figure 19 shows the results of the calibration.

Table 9. Characteristics of two individual pots used for calibration

Plot	*dg*	*hdom*	*ddom*	*Nv*	*T*	*G*	*SI40*
pn9	13.0	9.6	18.1	1620	16	21.5	19
pn1	21.8	15.6	28.3	1267	30	47.3	19

SI40 is site index at base age 40 years, obtained with equation of Tomé (2001)

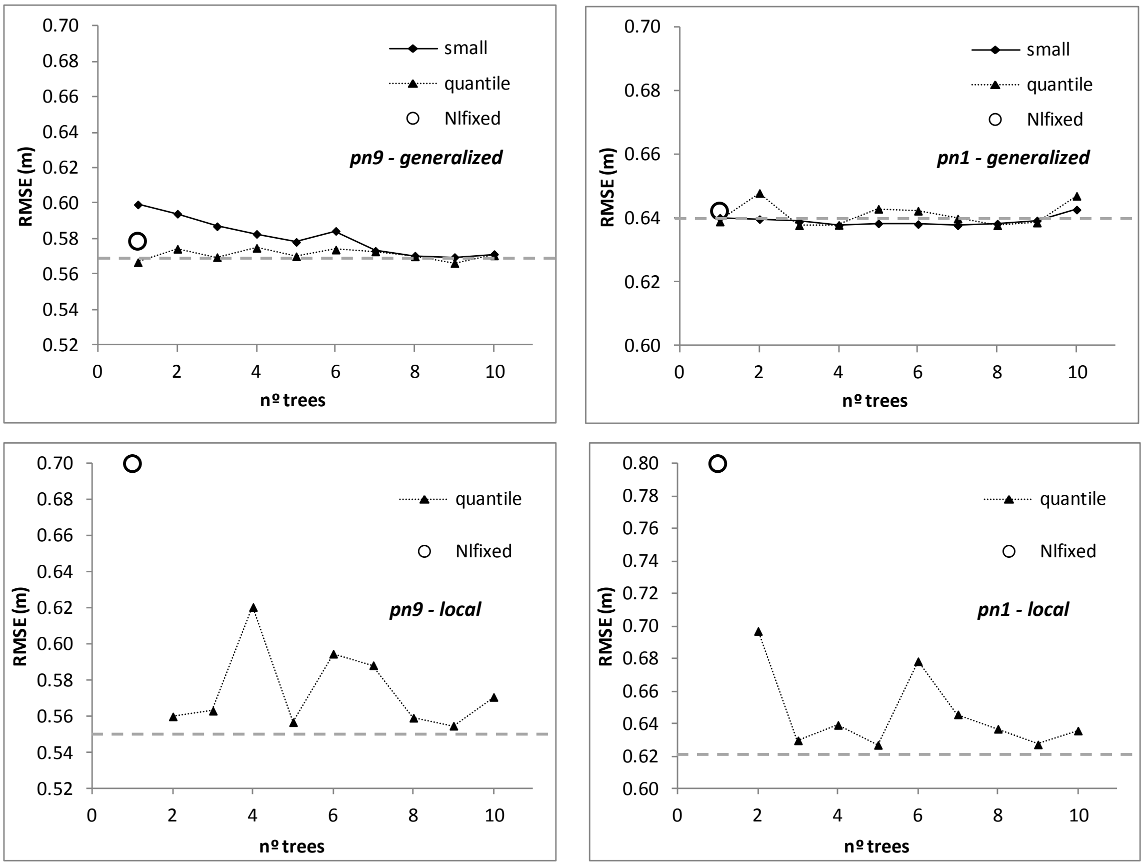


Figure 19.Above: calibrated response of the generalized mixed model (M2) with 1-10 smallest trees and 1-10 quantile trees in inventory plots pn9 (left) and pn1 (right) (dashed thick line represents the simple nls fit). Down: calibrated response of the local mixed model (M1) with 1-10 quantile trees in inventory plots pn9 (left) and pn1 (right) (dashed thick line represents the best possible fit using all trees in calibration). Nlfixed represents the average response.

Again, in the case of the generalized model, no relevant differences in RMSE values were observed when using the three alternatives (mixed model with random effects, mixed model with fixed effects only and model obtained in *nls* fit). Calibrating generalized mixed models with smallest trees in the plot was reported to produce better results than other methods in maritime pine stands (Gómez-García et al., 2015) and also for other species (Castedo Dorado et al., 2006; Crecente-Campo et al., 2010). The method also produced adequate performance in this study but did not produce results much different than those of the quantile trees option (except, perhaps, when using just one tree). The reduction in RMSE with calibration was more evident in the local mixed model rather than on the generalized mixed model (Figure 19). In this study, the use of covariates from the stand in the generalized equation (G7) seems to adequately capture the between-plots variability and using random effects will not explain much more of this variability. The calibration of the local mixed model with just one tree did not work well as was also observed in Gómez-García et al. (2015). Apparently it is not straightforward to clearly recommend a particular number of trees for calibrating this model. Nevertheless, with a relative low sampling effort (2 to 5 trees) it is possible to obtain estimates that closely approach estimates from the generalized model or, in some plots, even slightly better.

In a final test, using information from measurements made in a plot from Lomba ZIF under the framework of Portuguese NFI (IFN5, 2005), we calibrated both local and generalized mixed models with 3 measured heights (roughly corresponding to quantile trees) calculating, afterwards, the bias and mean of absolute differences between real and estimated heights (see Tables 10 and 11 in Appendix 2). The results showed that the generalized model with the parameterization from the *nls* fit produced the best results, close, however, of the results obtained with the corresponding mixed model. The local mixed model performed slightly worst. All these models produced better results than the generalized *h-d* equation used in the IFN5.

From these analyses, balancing sampling effort, accuracy and practical applications, the selected generalized *h-d* model (G7) is recommendable in preference to the other approaches tested. This is in line with the findings in the study of Gómez-García et al. (2015) for maritime pine stands from a region of North Portugal (Vale do Tâmega). Using model G7 requires having information on stand variables that are usually obtained in forest inventories in Portugal. Thus, it will not be a problem. If information on stand variables is not known and the total height of quantile trees from the diameter distribution is available, calibration of the local mixed model can be considered, using up to 5 trees. As it has been typical in NFI to measure height at least in one sample tree by diameter class, we normally will get appropriate data for a calibrate response. As pointed by Trincado et al. (2007), the use of a basic mixed model in forest inventories by selecting a sub-sample of trees for height measurement enables the maintenance of a simple model structure without inclusion of additional predictor variables.

5.4. Comparison of conventional and LiDAR measurements of *h* and *d*

As mentioned in section 4.2, terrestrial laser scanning (FARO Focus 3D) was used to measure diameter (d) and height (h) in a few plots from our study area under a workshop related to forest inventory and new technologies. For these plots that were measured conventionally, using caliper and vertex as well as measured with the LiDAR, comparisons could be made. The number of measured trees by the two approaches was different in most of the plots because no fixed plot area was defined for TLS analysis. The comparisons were done graphically and by using simple *t-tests* (after performing two variances *F-tests*). The results are summarized in Appendix 3. Because it is just a secondary result in our study, here we only report that in the measured plots, no statistical significant differences were detected between conventional and TLS measurements at 5% significance level (except for height in one of the plots, however not significant for 1% level). The notched boxplots obtained with package *ggplot2* from R software show high overlapping of notches for the two approaches in both measured variables, an indication of no significant differences.

6. Applications

In this last part of the work we will present application examples of where and how we can usefully use height-diameter equations. We will mainly focus in the generalized *h-d* equation G7 resulting from nonlinear least squares fit, which, in this study, is recommended for practical applications, demonstrating to present suitable behavior for estimating individual-tree height.

Example 1

The first example shows the applicability of model G7 in the framework of national forest inventory (NFI). Data from a NFI plot located in the Lomba ZIF and measured in the framework of Portuguese NFI (IFN5, 2005) is used. The plot area is 500 m² and at the time of field measurements (2005), the number of living trees per hectare (Nv) was 520, the dominant height (hdom) was 8.2 m, the dominant diameter (ddom) was 21.2 cm, the quadratic mean diameter (dg) was 18.1 cm and stand age (t) was 17 years.

All the diameters of the trees were measured but only three of them had both height and diameter measured. In the Portuguese NFI's it has been common to measure height only in "sample" trees inside each inventory plot. Sample trees are the ones that have its diameter closest to the central value of diameter classes (classes of 5 cm). The stand variables needed to apply G7 for *h* estimation were measured. These variables are usually measured in the framework of Portuguese NFI. So, applying G7, heights for the remaining trees in the plot can be estimated. These values are presented in Table 10 in Appendix 2 under the column *G_NLS*. From Table 11 in Appendix 2 it is possible to observe that model G7 behaves adequately when estimating individual-tree height. The model presented low bias and the error that in average is committed when using the model (evaluated by MAR), is inferior to the one that was obtained using the equation that was applied in IFN5. As tree height is an important component for the accuracy of volume equations, it is important to estimate *h* the best as possible in order to do wiser evaluations of available volume stocks.

As a remainder note, if stand variables were not measured, the three measured heights could be used to calibrate the local mixed model M1 and use only *d* to recover missing tree heights. Judging for the results from Table 11 in Appendix 2, the results would be very satisfactory also.

Example 2

Models like G7 can be easily integrated in computer simulators that implement a dynamic modeling framework to predict growth and yield of forest stands, such as the simulators GesMO® (González González et al., 2012) or FlorNExT® (Pérez-Rodríguez et al., 2016 *in press*).

In this example we refer the simulator FlorNExT® (Pérez-Rodríguez et al., 2016 *in press*) and its extension FlorNExTPRO® developed for the Northest region of Portugal where our study area is located. These tools have been created in order to solve existing technical constraints that limit sustainable forestry at the stand and at the landscape level and to encourage and support sustainable wood mobilization in the northeastern region of Portugal. The tools are intended for both forest managers and forest owners. FlorNExT® is a free cloud computing application to estimate growth and yield of maritime pine even-aged stands in the Northeast of Portugal (Figure 20).

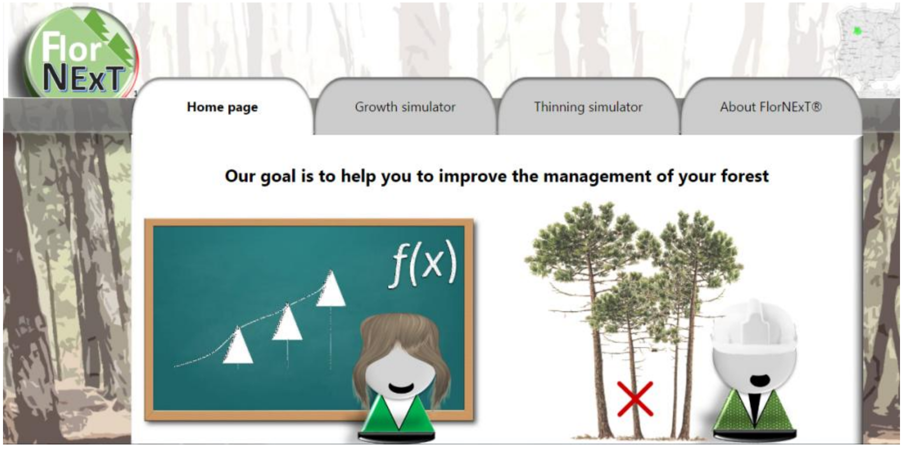


Figure 20. Home-page of FlorNExT®

In the simulator, a set of dynamic growth models that are implemented, project key stand variables over time (projection functions as difference equations). When thinning is applied, the simulator computes the respective diameter distribution using a bi-parametric Weibull probability density function whose parameters are related to the projected stand variables (Pérez-Rodríguez et al., 2016 *in press*).The volume extracted in a particular thinning is computed using an individual-tree volume equation and a generalized height-diameter equation. Here is another situation where our selected model G7 can be very useful. As the simulator uses a probability density function (Weibull) to define the diameter distribution and consequently the diametric structure of the stand for a particular moment in time, then it is necessary to obtain tree heights associated to these diameters which, in turn will feed an individual-tree volume equation in order to then obtain estimates of stand volume.

FlorNExTPRO® extends FlorNExT® and is a desktop tool that provides a spatial framework to forest optimization using a linear programming approach. The tool tests a lot of possible combinations of silvicultural plans defined by the user for a set of management units. In addition, the user can apply constraints to the simulations, such as maximum harvested or thinned area, or minimum volume to be removed per period, among other options. FlorNExTPRO® uses the same set of equations that are implemented in FlorNExT® including a generalized *h-d* equation to support volume calculations.

The programming of an equation like G7 in a simulator such as FlorNExT® or FlorNExTPRO®, is a relatively easy task with the usual computer languages used. The local mixed model M1 is more difficult to implement and in our opinion, is only worthy if gains in precision are substantial and when stand variables are not known.

In a final test, we used equation G7 to estimate measured heights from Portuguese NFI (IFN5, 2005) plots outside Lomba ZIF but inside Nordeste (NE) region. Only plots with more than 5 measured heights were used. Table 12 in Appendix 2 presents the comparative results of our generalized *h-d* equation (G7) and the generalized equation that was used in IFN5. Model G7 presented better performance than IFN5 equation in several of the analyzed plots, showing robustness when applied out of Lomba ZIF area. Despite due caution should be taken because of the size of our sample, this is one more excellent indication for the predictive use of model G7. This same model but using the parameterization of Goméz-García et al. (2015) for Tâmega Valley in North Portugal (column GG in Table 12) seems to produce equally good results for NE region.

7. Final Remarks

The relationship between diameter at breast height and total tree height for maritime pine stands in Lomba ZIF, NE Portugal was modeled to establish provisory *h-d* equations to be available for assisting in the management of these forested areas. A generalized *h-d* equation (G7) derived from Michailoff (1943) equation is recommended for practical applications in the Lomba ZIF:

FI_{adj} = 0.896

RMSE= 0.909

This equation requires information on stand variables that are usually collected in forest inventories in Portugal. The equation presents adequate predictive capacity and, with due caution, regarding the size and range of our sample, it apparently have robust behavior outside the Lomba ZIF area. The generalized mixed model based in this equation did not revealed evident superiority in comparison to G7.

If stand variables are not available, the use of a local mixed model (M1) based on the Meyer (1940), can be used, providing that additional *h-d* measurements (2-5) on quantile (or random) trees are available:

.

To use this equation it is necessary to estimate the random effects (*u* and *v*) in order to localize the model and obtain a calibrated response. In this study we used a program implemented in R which is presented in Appendix 1.

Finally, develop tools implementing the models are needed, because in many cases the final user doesn't have enough knowledge about how to use it. Tools like FlorNExT® or FlorNExTPRO® are the drivers to transfer the models like the presented in this work, making it useful in practice terms. Moreover, providing these tools to forest managers improves the framework of decision making in the forest management, because opens the possibility of get more and better information about how the forest are or could be in the future.

8. References

Abad Viñas, R., Caudullo, G., Oliveira, S., de Rigo, D., 2016. *Pinus pinaster* in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. e012d59+.

Adame, P., I. Cañellas, I., del Río, M., 2005. Modelo de la relación altura-diámetro para rebolares en Castilla y León. IV Congreso Forestal Español, Sociedad Española de Ciencias Forestales. Zaragoza, 26-30 septiembre. 8 pp.

Adame, P., del Río, M., Cañellas, I., 2008. A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). Forest Ecology and Management 256, 88-98.

AFN, 2010. FloreStat. 5th National Forest Inventory Information Retrieval Tool. V1.0 Metacortex, Consultoria e Modelação de Recursos Naturais, S.A.

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19, 716-23.

APA, 2016a. SNIRH: Sistema Nacional de Informação de Recursos Hídricos. Agência Portuguesa do Ambiente, I.P. *http://snirh.pt/snirh*.

APA, 2016b. SNIAmb - Sistema Nacional de Informação de Ambiente. Agência Portuguesa do Ambiente, I.P. *http://sniamb.apambiente.pt/Home/Default.htm*.

Arias-Rodil, M., Castedo-Dorado, F., Cámara-Obregón, A., Diéguez-Aranda, U., 2015. Fitting and Calibrating a Multilevel Mixed-Effects Stem Taper Model for Maritime Pine in NW Spain. PLoS ONE 10 (12), e0143521. doi:10.1371/journal.pone.0143521

Arney, J.D., 1985. A modeling strategy for the growth projection of managed stands. Canadian Journal of Forest Research 15, 511-518.

Avery, T.E., Burkhart, H., 2011. Forest Measurements, 5th edn. McGraw-Hill, New York.

Baptista, F.O., Santos, R.T., 2005. Os Proprietários Florestais. Edições Celta, Oeiras, Portugal. 93 pp. [In Portuguese].

Bates, D.M., Watts, D.G., 1980. Relative curvature measures of nonlinearity. Journal of Royal Statistical Society B 42, 1-16.

Bennett, F., Clutter, J., 1968. Multiple-product yield estimates for unthinned slash pine plantations-pulpwood, sawtimber, gum. USDA For. Serv. Res. Pap., SE-35.

Bi, H., 2000. Trigonometric variable-form taper equations for Australian eucalyptus. Forest Science 46, 397-409.

Blanco E., Casado M., Costa M., Escribano R., Gracía Antón M., Génova M., Gómez Manzaneque G., Gómez Manzanaque F., Moreno J., Morla C., Regato P., Sainz Ollero H., 1997. Los bosques ibéricos, Editorial Planeta, Barcelona, España, 572 pp.

Budhathoki, C.B., Lynch, T.B., Guldin, J.M., 2008. A Mixed-Effects Model for the dbh-Height Relationship of Shortleaf Pine (*Pinus echinata* Mill.). Southern Journal of Applied Forestry 32(1), 5-11.

Burkhart, H., Strub, M., 1974. A model for simulation of planted loblolly pine stands. In: Fries, J. (Ed.), Growth Models for Tree and Stand Simulation. Royal College Forestry, Res. Not. 30, Stockholm, pp. 128-135.

Burkhart, H.E., Tomé, M., 2012. Modeling Height-Diameter Relationships. In: Modeling Forest Trees and Stands, 280-282 p. Springer.

Calama R., Montero G., 2004. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain. Canadian Journal of Forest Research 34, 150-163.

Cañadas Díaz N., Garciá Guemes C., Montero González G., 1999. Relación altura-diámetro para *Pinus pinea* L. en el Sistema Central. In: Actas del Congreso de Ordenación y Gestión Sostenible de Montes, Santiago de Compostela, 4-9 October. Volume I, pp. 139-153.

Carrión, J.S., Navarro, C., Navarro, J., Munuera, M., 2000. The distribution of cluster pine (*Pinus pinaster*) in Spain as derived from palaeocological data: relationships with phytosociological classification. Holocene 10(2), 243-252.

Castedo F., Ruiz A.D., Álvarez González J.G., 2001. Modelización de la relación altura-diámetro para *Pinus pinaster* Ait. en Galicia mediante la función de densidad bivariante SBB, Investigación Agraria, Sistemas y Recursos Forestales 10, 111-126.

Castedo Dorado F., Diéguez-Aranda U., Barrio M., Sánchez M., Gadow Kv., 2006. A generalized height-diameter model including random components for radiate pine plantations in northwestern Spain. Forest Ecology and Management 229, 202-213.

Coelho, I.S., 2003. Propriedade da Terra e Política Florestal em Portugal. Silva Lusitana 11(2), 185-199. [In Portuguese]

Colbert, K.C., Larsen, D.R., Lootens, J.R., 2002. Height–diameter equations for thirteen Midwestern bottomland hardwoods species. Northern Journal of Applied Forestry. 19(4), 171-176.

Corral-Rivas, S., Álvarez-González, J.G., Crecente-Campo, F., Corral-Rivas, J.J., 2014. Local and generalized height-diameter models with random parameters for mixed, uneven-aged forests in Northwestern Durango, Mexico. Forest Ecosystems, 1:6 http://www.forestecosyst.com/1/1/6.

Correia, A.V., Oliveira, A.C., Fabião, A., 2007. Biologia e ecologia do pinheiro-bravo. In Sande Silva Coord.Ed. Árvores e Florestas de Portugal. Pinhais e Eucaliptais. Lisboa, Portugal, 17-35 pp.

Cox F., 1994. Modelos parametrizados de altura, Informe de convenio de investigación interempresas.

Crecente-Campo F., Tomé M., Soares P., Diéguez-Aranda U., 2010. A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. Forest Ecology and Management 259, 943-952.

Curtis, R.O., 1967. Height–diameter and height–diameter–age equations for second-growth Douglas-Fir. Forest Science 13, 365-375.

Curtis, R.O., Clendenen, G.W., DeMars, D.J., 1981. A new stand simulator for coast Douglas-fir: DFSIM user's guide. USDA For. Serv. Gen. Tech. Rep. PNW-128. 79 pp.

Dassot, M., Constant, T., Fournier, M., 2011. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Annals of Forest Science 68, 959-974.

Diéguez-Aranda U., Rojo-Alboreca A., Castedo-Dorado F., Álvarez-González J.G., Barrio-Anta M., Crecente-Campo F., González-González J.M., Pérez-Cruzado C., Rodríguez-Soalleiro R., López-

Sánchez C.A., Balboa-Murias M.A., Gorgoso-Varela J.J., Sánchez-Rodríguez F. 2009. Herramientas selvícolas para la gestión forestal sostenible en Galicia. Consellería do Medio Rural, Xunta de Galicia. 268 pp + CD-Rom.

Eerikäinen, K., 2003. Predicting the height-diameter pattern of planted *Pinus Kesiya* stands in Zambia and Zimbabwe. Forest Ecology and Management 175, 355-366.

ENF, 2006. Estratégia Nacional para as Florestas. Direcção Geral dos Recursos Florestais, Lisboa, Portugal. 163 pp. [In Portuguese]

Fang, Z., Bailey, R.L., 1998. Height–diameter models for tropical forest on Hainan Island in southern China. Forest Ecology and Management 110, 315-327.

Fang Z., Bailey R.L., 2001. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments. Forest Science 47(3), 287-300.

Fang, Z., Bailey, R.L., and Shiver, B.D., 2001. A multivariate simultaneous predictions system for stand growth and yield with fixed and random effects. Forest Science 47(4), 550-562.

Farjon, A., 2010. A handbook of the world's conifers. Brill, e-ISBN : 9789047430629.

Flood, M. 2001. Laser altimetry: from science to commercial lidar mapping Photogramm. Eng. Remote Sens. **67**(11): 1209–1217.

Gaffrey D., 1988. Forstamts-und bestandesindividuelles Sortimentierungsprogramm als Mittel zur Planung, Aushaltung und Simulation. Diplomarbeit Forscliche Fakultät. University of Göttingen.

Gómez-García, E., Fonseca, T.F., Crecente-Campo, F., Almeida, L.R., Diéguez-Aranda, U., Huang, S., Marques, C.P., 2015. Height-diameter models for maritime pine in Portugal: a comparison of basic, generalized and mixed-effects models. iForest 9, 72-78. doi:10.3832/ifor1520-008.

González González, J.M., Castedo-Dorado, F., Diéguez-Aranda, U., Rojo Alboreca, A., Álvarez González, J.G., 2012. GesMO ®, Manual de Usuario. Simulador de crecimiento y producción de rodales forestales. Versión 2.1. Unidade de Xestión Forestal Sostible (UXFS). Escola Politécnica Superior de Lugo Universidade de Santiago de Compostela.

Goyer, G. G.; R. Watson (September 1963). "The Laser and its Application to Meteorology". Bulletin of the American Meteorological Society. **44** (9): 564–575 [568]

Gregoire, T.G., 1987. Generalized error structure for forestry yield models. Forest Science 33, 423-444.

Hanus, M.L., Marshall. D.D., Hann, D.W., 1999. Height-diameter equations for six species in the coastal regions of the Pacific Northwest. Res. Cont. 25, For. Res. Lab. Oregon State Univ., Corvallis, OR. 11 pp.

Harrison, W.C., Burk, T.E., Beck, D.E., 1986. Individual tree basal area increment andtotal height equations for Appalachian mixed hardwoods after thinning. Southern Journal of Applied Forestry 10, 99-104.

Hökkä, H., 1997. Height-diameter curves with random intercepts and slopes for trees growing on drained peatlands. Forest Ecology and Management 97, 63-72.

Huang, S., Titus, S.J., 1993. An index of site productivity for uneven-aged and mixedspecies stands. Canadian Journal of Forest Research 23, 558-562.

Huang, S., Titus, S.J., Wiens, D.P., 1992. Comparison of nonlinear height-diameter functions for major Alberta tree species. Canadian Journal of Forest Research 22, 1297-1304.

Huang, S., Price, D., Titus, S., 2000. Development of ecoregion-based height-diameter models for white spruce in boreal forests. Forest Ecology and Management 129, 125-141.

Hui G., Gadow K.v., 1993. Zur Entwicklung von Einheitshöhenkurven am Beispiel der Baumart Cunninghamia lanceolata, AFJZ 164, 218-220.

Hui G., Gadow K.v., 1999. Modelling Forest Development, Kluwer Academic Publishers.

Hyyppä J. , Hyyppä H. , Leckie D. , Gougeon F. , Yu X., Maltamo M. , 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests ,a Finnish Geodetic Institute, FIN-02431 Masala, Finland

ICNF, 2013. IFN6 - Áreas dos usos do solo e das espécies florestais de Portugal continental. Resultados preliminares. Instituto da Conservação da Natureza e das Florestas, Lisboa. 33 pp. [In Portuguese]

ICNF, 2016. Zonas de Intervenção Florestal. Situação Geral e Informação. Instituto da Conservação da Natureza e das Florestas. *www.icnf.pt/portal/florestas/gf/zif/sit-ger-inf#zif.*

IFFN, 2006. Forest Fires in Portugal in 2005- An Overview, International Forest Fire News No. 34 ISSN 1029-0864

INF5 2005. Inventário Florestal Nacional 5. *http://www.icnf.pt/portal/florestas/ifn*.

IPMA, 2016. Normais Climatológicas - 1981-2010 (provisórias) - Bragança. Instituto Português do Mar e da Atmosfera. *www.ipma.pt/pt/oclima/normais.clima/1981-2010/003*.

Jayaraman, K., Lappi, J., 2001. Estimation of height–diameter curves through multilevel models with special reference to even-aged teak stands. Forest Ecology and Management 142, 155–162.

Kalliovirta J., Laasasenaho J. and Kangas A., 2005. Evaluation of the Laser-relascope. Forest Ecology and Management 204, 2-3, 181-194.

Korf V. 1939. Příspěvek k matematické definici vzrůstového zákona hmot lesních porostů. Lesnická práce, 18, 339-379.

Krumland, B.E., Wensel., L.C., 1978. Generalized height diameter equations for coastal conifers. Co-op Redwood Yield Research Project, Dept. of Forestry and Conservation, Univ. of California, Berkeley, CA Res. Note No.8. 10 pp.

Larjavaara, M., Muller-Landau, H.C., 2013. Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. Methods in Ecology and Evolution 4, 793–801.

Larsen, D.R., Hann, D.W., 1987. Height-Diameter Equations for Seventeen Tree Species in Southwest Oregon. Research Paper 49. Forest Research Laboratory, Oregon State University, Corvallis, OR, 16 pp.

Leduc, D., Goelz, J., 2009. A Height-Diameter Curve for Longleaf Pine Plantations in the Gulf Coastal Plain. Southern Journal of Applied Forestry 33(4), 164-170.

Lefsky, M.A., Harding, D., Cohen, W.B., Parker, G., and Shugart, H.H. 1999. Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA. Remote Sens. Environ. **67**(1): 83–98. doi:10.1016/S0034-4257(98) 00071-6.

Lenhart, J., 1968. Yield of old-field loblolly pine plantations in the Georgia piedmont. Ph.D. Thesis, University of Georgia, Georgia.

Li, R., Stewart, B., Weiskittel, A., 2011. A Bayesian approach for modeling non-linear longitudinal/hierarchical data with random effects in forestry. Forestry 85(1), 17-25.

Lim, K.S., and Treitz, P.M. 2004. Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. Scand. J. For. Res. 19:558–570. doi:10.1080/02827580410019490.

Lindstrom, M.J., Bates, D.M., 1990. Nonlinear mixed effects for repeated measures data. Biometrics 46, 673–687.

Litell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O., 2006. SAS for Mixed Models, 2nd ed. SAS Institute Inc., Cary, NC, 814 pp.

López Sánchez, C.A., Gorgoso Varela, J., Castedo Dorado, F., Rojo Alboreca, A., Rodríguez Soalleiro, R., Alvarez González, J.G., Sánchez Rodríguez, F., 2003. A height–diameter model for Pinus radiata D. Don in Galicia (Northwest Spain). Annals of Forest Science 60, 237-245.

López-Sánchez, C., Rodríguez-Soalleiro, R., Castedo-Dorado, F., Corral-Rivas, S., Álvarez- González, J., 2016. A taper function for *Pseudotsuga menziesii* plantations in Spain. Southern Forests 78(2), 131-135.

Magnussen, S., and Boudewyn, P. 1998. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. Can. J. For. Res. **28**: 1016–1031. doi:10. 1139/cjfr-28-7-1016.

Magnusson, M., Fransson, J.E., Holmgren, J. 2007. Effects on estimation accuracy of forest variables using different pulse density of laser data. For. Sci., 53, 619–626.

Mazza, G., Cutini, A., Manetti, MC. (2014).Influence of tree density on climate-growth relationships in a *Pinus pinaster* Ait. forest in the the northern mountains of Sardinia (Italy) ,iForest 8: 456-463. - doi: 10.3832/ifor1190-007

Mendes, A.S.C., 2002. Financial instruments of forest policy in Portugal in the 80s and 90s. EFI Forest Policy Research Forum: Cross-Sectoral Policy Impacts on Forests, 4-6 April, Sanvonlinna.

Mehtätalo, L., 2004. A longitudinal height-diameter model for Norway spruce in Finland. Canadian Journal of Forest Research 34, 131-140.

Mehtätalo, L., 2005. Height-Diameter Models for Scots Pine and Birch in Finland. Silva Fennica 39(1), 55-66.

Mehtätalo, L., de-Miguel, S., Gregoire, T.G., 2015. Modeling height-diameter curves for prediction. Canadian Journal of Forest Research, 45(7), 826-837, 10.1139/cjfr-2015-0054.

Meyer, H.A., 1940. A mathematical expression for height curves. Journal of Forestry 38, 415-420.

Michailoff, I., 1943. Zahlenmässiges verfahren für die ausführung der bestandeshöhenkurven.Forstwissenschaftliches Centralbatt-Tharandter Forstliches Jahrbuch, 6, 273-279.

Mirkovich J.L., 1958. Normale visinske krive za chrast kitnak i bukvu v NR Srbiji. Zagreb. Glasnik sumarskog fakulteta 13.

Næsset, E. 1997. Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS Journal of Photogrammetry and Remote Sensing 52: 49–56.

Næsset, E., T. Gobakken, J. Holmgren, H. Hyyppä, J. Hyyppä, M. Maltamo, M. Nilsson, H. Olsson, A. Persson and U. Soderman. 2004. Laser scanning of forest resources: The Nordic experience. Scandinavian Journal of Forest Research 19: 482–499.

Newnham, G.J., Armston, J.D., Calders, K., Disney, M., Lovell, J.L., Schaaf, C.B., Strahler, A.H., Danson, F.M., 2015. Terrestrial Laser Scanning for Plot-Scale Forest Measurement. Current Forestry Reports 1(4), 239-251 pp. DOI 10.1007/s40725-015-0025-5.

Newton, P.F., Amponsah, I.G., 2007. Comparative evaluation of five height–diameter models developed for black spruce and jack pine stand-types in terms of goodness-of-fit, lack-of-fit and predictive ability. Forest Ecology and Management 247, 149-166.

Nilsson, M. 1996. Estimation of tree heights and stand volume using an airborne LiDAR system. Remote Sensing of Environment. 56: 1–7.

Oliveira, A.C., Pereira, J.S., Correia, A., 2000. A silvicultura do pinheiro bravo. Centro Pinus, Porto, 112 p. [In Portuguese]

Parresol, B.R., 1992. Baldcypress height-diameter equations and their prediction confidence interval. Canadian Journal of Forest Research 22, 1429-1434.

Paulo, J.A., Tomé, J., Tomé, M., 2011. Nonlinear fixed and random generalized height-diameter models for Portuguese cork oak stands. Annals of Forest Science 68, 295-309.

Pearl, R., and Reed, L.J. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Natl. Acad. Sci. U.S.A 6, 275-288.

Peng, C., 1999. Nonlinear Height-Diameter Models for Nine Boreal Forest Tree Species in Ontario. OFRI-Rep. 155. Ontario Forest Research Institute, Ministry of Natural Resources, 28 pp.

Peng, C., Zhang, L., Huang, S., Zhou, X., Parton, J., Woods, M., 2001. Developing ecoregion-based height–diameter models for jack pine and black spruce in Ontario. Forest Research Report 159. Ministry of Natural Resources. Ontario Forest Research Institute, Ontario, Canada.

Pérez-Rodríguez, F., Nunes, L., Sil, A., Azevedo, J.C., 2016. FlorNExT®, a cloud computing application to estimate growth and yield of maritime pine (*Pinus pinaster* Ait.) stands in northeastern Portugal. Forest Systems (*in press*).

Petráš, R., Bošel'a, M., Mecko, J., Oszlányi, J., Popa, I., 2014. Height-diameter models for mixed-species forests consisting of spruce, fir, and beech. Folia Forestalia Polonica, series A, 56 (2), 93-104.

Pienaar, L.V., Harrison, W.M., Rheney, J.W., 1990. PMRC Yield Prediction System for Slash Pine Plantations in the Atlantic Coast Flatwoods. PMRC Technical Report 1990-3. School of Forest Resources, University of Georgia, Athens, GA, 31 pp.

Pinheiro, J.C., Bates, D.M., 1998. Model Building for Nonlinear Mixed Effects Model. Department of Biostatistics and Department of Statistics, University of Wisconsin, Madison, Wis., 11 pp.

Pretzsch, H. 2009. Forest dynamics, growth and yield: from measurement to model. Berlin, Germany: Springer Verlag; 664 pp.

Prodan, M., 1965. Holzmesslehre. Sauerla¨nders Verlag, Frankfurt.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. The R Foundation for Statistical Computing. Vienna, Austria. R version 3.2.4 (2016-03-10). *https://www.Rproject.org*.

Ratkowsky, D.A., 1990. Handbook of nonlinear regression. Marcel Dekker, Inc., New York.

Ratkowsky, D.A, and Reedy, T.J., 1986. Choosing near-linear parameters in the four-parameter logistic model for radioligand and related assays. Biometrics, 42, 575-582.

Ribeiro, M.M., Plomion, C., Petit, R., Vendramin, G.G., Szmidt, A.E., 2001. Variation in chloroplast single-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). Theoretical and Applied Genetics, 102, 97-103.

Richards, F.J., 1959. A flexible growth function for empirical use. Journal of Experimental Botany 10, 290-300.

Rojo, A., Perales, X., Sanchez-Rodríguez, F., Álvarez-González, J., Von Gadow, K., 2005. Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (northwestern Spain). European Journal of Forest Research 124, 177-186.

Rojo-Alboreca, A., García-Villabrille, J.D., Pérez-Rodríguez, F., 2015. EucaTool®, a cloud computing application for estimating the growth and production of *Eucalyptus globulus* Labill. plantations in Galicia (NW Spain). Forest Systems 24(3), eRC06, 4 pages. dx.doi.org/10.5424/fs/2015243-07865.

Rushforth, K. 1986. Pocket Guide to Trees

Schröder J., Álvarez González J.G., 2001. Developing a generalized diameter-height model for maritime pine in Northwestern Spain, Forstwissenschaftliches Centralblatt 120, 18-23.

Schwarz, G., 1978. Estimating the dimension of a model. The Annals of Statistics 5 (2), 461-464.

Sharma, M., Zhang, S.Y., 2004. Height–diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*. Scandinavian Journal of Forest Research 19, 442-451.

Sharma, M., Parton, J., 2007. Height-diameter equations for boreal tree species in Ontario using a mixed-effects modeling approach. Forest Ecology and Management 249, 187–198.

Sharma, R.P., Breidenbach, J., 2015. Modeling height diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data, Forest Science and Technology, 11(1), 44-53.

Sloboda V.B., Gaffrey D., Matsumura N., 1993. Regionale und locale Systeme von Höhenkurven für gleichaltrige Waldbestände, Allg. Forst. Jagdztg. 164, 225-228.

Soares, P., Tomé, M., 2002. Height–diameter equation for first rotation eucalypt plantations in Portugal. Forest Ecology and Management 166, 99-109.

Spies, T.A., Cohen, W.B., 1992. An Index of Canopy Height Diversity. Coastal Oregon Productivity Enhancement Program Report 5, 5-7. Oregon State University, Corvallis, OR.

Stankova, T.V., Diéguez-Aranda, U., 2013. Height-diameter relationships for Scots pine plantations in Bulgaria: optimal combination of model type and application. Annals of Forest Research 56(1), 149-163.

Staudhammer, C., LeMay, V., 2000. Height prediction equations using diameter and stand density measures. The Forestry Chronicle 76(2), 303-309.

Stoffels, A., van Soest, J., 1953. The main problems in sample plots. Ned. Boschb. Tijdschr. 25, 190–199.

Temesgen H., Gadow Kv., 2004. Generalized height-diameter models - an application for major tree species in complex stands of interior British Columbia. European Journal of Forest Research 123, 45-51.

Temesgen, H., Hann, D.W., Monleon, V.J., 2007. Regional Height-Diameter Equations for Major Tree Species of Southwest Oregon. Westtern Journal of Applied Forestry 22(3), 213-219.

Tomé M., 1989. Modelação do crescimento da árvore individual em povoamentos de *Eucalyptus globulus* Labill. (1ª rotação) na região Centro de Portugal. PhD Thesis, Instituto Superior de Agronomia, Lisboa, Portugal. [In Portuguese]

Tomé M., 2001. Tabela de produção geral para o pinheiro bravo desenvolvida no âmbito do projecto PAMAF 8165. Relatórios GIMREF RT9/2001. CEF, ISA, Lisboa. [In Portuguese]

Trincado, G., VanderSchaaf, C.L., Burkhart, H.E., 2007. Regional mixed-effects height–diameter models for loblolly pine (*Pinus taeda* L.) plantations. European Journal of Forest Research 126, 253-262.

Vanclay, J.K., 1994. Modelling Forest Growth and Yield - Application to Mixed Tropical Forests. CAB International, UK.

Van Laar, A., Akça, A., 2007. Forest Mensuration. Springer. 384 pp.

Vargas-Larreta, B., Castedo-Dorado, F., Álvarez-González, J.G., Barrio-Anta, M., Cruz-Cobos, F., 2009. A generalized height-diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico). Forestry 84(2), 445-462.

Vonesh, E.F., Chinchilli, V.M., 1997. Linear and Nonlinear Models for the Analysis of Repeated Measurements. Marcel Dekker Inc., New York. 560 pp.

Wahid, N., González-Martìnez, S.C., El Hadrami, I., Boulli, A., 2006. Variation of morphological traits in natural populations of maritime pine (*Pinus pinaster* Ait.) in Morocco. Annals of Forest Science 63, 83-92.

Wang, C.H., Hann, D.W., 1988. Height-diameter equations for sixteen tree species in the central Willamette Valley or Oregon. Res. Pap. 51. Oregon State Univ. For. Res. Lab. 7 pp.

Wykoff, W.R., Crookston, N.L., Stage, A.R., 1982. User's guide to the stand prognosis model. USDA For. Serv. Gen. Tech. Rep. INT-133. 122 pp.

Yang, R.C., Kozak, A., Smith, J.H.G., 1978. The potential of Weibull-type functions as a flexible growth curves. Canadian Journal of Forest Research 8, 424-431.

Yang Y., Huang S., 2013. On the statistical and biological behaviors of nonlinear mixed forest models. European Journal of Forest Research. 132(5–6), 727-736.

Yuancai, L., Parresol, B.R., 2001. Remarks on Height–Diameter Modelling. USDA Forest Service, Research Note SRS-10, Southern Research Station, Asheville, NC, 5 pp.

Zeide, B., 1993. Analysis of growth equations. Forest Science 39(3), 594-616.

Zeide, B., Vanderschaaf, C., 2002. The effect of density on the height-diameter relationship. In: Outcalt, K.W. (ed). Proceedings of the 11th Biennial Southern Silvicultural Research Conference. 2001 March 20-22, Knoxville, TN. USDA Forest Service, Gen. Tech. Rep. SRS-48, Asheville, NC, pp. 463-466.

Zhang, L., 1997. Cross-validation of nonlinear growth functions for modeling tree height-diameter distributions. Annals of Botany 79, 251-257.

Zhang, X., Duan, A., Zhang, J., Xiang, C., 2014. Estimating Tree Height-Diameter Models with the Bayesian Method. The Scientific World Journal, Article ID 683691, 9 pages, dx.doi.org/10.1155/2014/683691.

Appendix 1

Example of R program for fitting and calibration of Local nonlinear mixed model

```
#*******************************************************************
#    FITTING NONLINEAR MIXED MODELS with nlme
#*******************************************************************
library(nlme)
library(ggplot2)
#-------------------------------------------------------------------
# 1. LOCAL MODEL:      Reading Data
#-------------------------------------------------------------------
data<-read.csv("C:/ …. Data_loess_calibra.csv", header=TRUE, sep = ";", dec=".")
colnames(data)

data.mixed <- groupedData(h ~ d | id_plot, data=data)

#-------------------------------------------------------------------
#              Fit using nlme package
#-------------------------------------------------------------------
Meyer.func <- function(d, b1, b2) (1.3+b1*(1-exp(-b2*d)))

Meyer.nlme<-nlme(h~Meyer.func(d, b1, b2), data=data.mixed,
            fixed = b1 + b2 ~ 1,
            random = b1 + b2 ~ 1,
            method = "REML",
            start=c(b1=20, b2=0.02))

summary(Meyer.nlme)
intervals(Meyer.nlme)
VarCorr(Meyer.nlme)

# Calibration example for the fitted local model Meyer.nlme,
# for plot p561 with all the observations of trees.
# ***********************************************************
# Adaptation of the program of Manuel Arias-Rodil et al 2015
# PLoS ONE 10 (12): e0143521. doi:10.1371
# ***********************************************************

# Obtain random-effects variance-covariance matrix and residual variance

.G <- VarCorr(Meyer.nlme)
.estimate.plot <- as.numeric(.G[c(1, 2), 1])  # Plot level
.corr.u.v.plot <- as.numeric(.G[2, 3])
.covariance.u.v.plot  <-  sqrt(.estimate.plot[1])  *  sqrt(.estimate.plot[2])  *  .corr.u.v.plot  #  Covariance between random effects
D  <-  matrix(c(.estimate.plot[1],  .covariance.u.v.plot,  .covariance.u.v.plot,  .estimate.plot[2]),  nc  =  2)  # Random-effects variance-covariance matrix
sigma2 <- as.numeric(.G[3, 1])  # Residual variance

# Create some variables that may help to automatize the procedure in a further step

.nameplot <- "p561"
dfplot <- data.mixed[data.mixed$id_plot == .nameplot, ]  # Select all the observations of plot "p561". Note that here we select the set of trees that we will use for calibration, which means that this is the part that should be modified to calibrate with different sets of trees (smallest, medium-size, biggest...)
```

```
randparms <- names(ranef(Meyer.nlme))
fparms <- fixef(Meyer.nlme)
tolerance <- 1e-4

# Calibration procedure
# EstimateRandomEffects.Meyer <- function(dfplot, randparms, fparms, D, sigma2, tolerance = 1e-4){
   nrand <- length(randparms)
     nobs <- nrow(dfplot)
# Ri matrix
       Mi <- diag(1, nrow = nobs, ncol = nobs)
        Ri <- sigma2 * Mi
# Additional trees used for calibration
        yi <- dfplot$h  # Tree heights
# Initial values for iterative procedure to estimate random effects
      b.0 <- rep(0, nrand)
      tol <- rep(1, nrand)
# Iterative procedure ("while" loop)
          while (sum(tol > tolerance) > 0){ # Stop if all tol values are bigger than 1e-4
# Z matrix and f(x_i, B, b)
 Zi <- attr(numericDeriv(quote(Meyer.func(d = dfplot$d, b1 = fparms[["b1"]] + b.0[1], b2 = fparms[["b2"]]
+ b.0[2])), theta = "b.0"), "gradient")
  fxiBb <- Meyer.func(d = dfplot$d, b1 = fparms[["b1"]] + b.0[1], b2 = fparms[["b2"]] + b.0[2])
  if(is.null(nrow(Zi))) Zi <- matrix(Zi, nrow = 1)
  # Random-effects estimation (Lindstrom & Bates, 1990; First Order Conditional Expectation method,
FOCE)
b <- D %*% t(Zi) %*% solve (Ri + Zi %*% D %*% t(Zi)) %*% ((yi - fxiBb) + Zi %*% b.0)
        if (all(b.0 == 0)) b.prev <- rep(1, nrand) else b.prev <- b.0
 tol <- abs((b - b.prev) / b.prev)  # Compute relative difference with random effects of previous iterations
b.0 <- b
}
bi <- t(b)
colnames(bi) <- randparms
# return(bi)
# }

# Checking whether we obtain the same random effects estimates
# by calibrating with information of all trees of a plot and computations
# from mixed-model fitting (they should be approximately equal)

# Random-effects from calibration with all trees of a plot
bi
# Random-effects from mixed model fitting
ranef(Meyer.nlme)[.nameplot, ]

# Note that L73 and L98-99 are as comments.
# If we uncomment those lines, we would have a function that would
# estimate random effects from model estimates
# (randparms, names of parameters with random effects; fparms,
# fixed-effects parameter estimates; D, random-effects
# variance-covariance matrix; sigma2, residual variance)
#and a set of trees use in calibration (dfplot)
```

Appendix 2

Plot inside Lomba ZIF area measured in the framework of IFN5 (2005)
Area=500 m^2; ***Nv***=520; ***hdom***=8.2; ***dg***=18.1; ***ddom***=21.2; ***t***=17

Table 10.Estimation of heights in an NFI plot located in Lomba ZIF

id_tree	*d*	*h*	*L_mixed*	*L_Fixed*	*L_NLS*	*G_mixed*	*G_Fixed*	*G_NLS*	*IFN5*
22	**10.55**		6.6	9.2	8.7	6.3	6.3	6.3	7.1
25	**11.2**		6.8	9.5	9.0	6.5	6.5	6.4	7.3
18	**12.2**		7.0	9.9	9.5	6.7	6.7	6.7	7.5
24	**12.25**		7.1	9.9	9.5	6.7	6.8	6.7	7.5
17	**13.2**		7.3	10.2	9.9	7.0	7.0	7.0	7.7
3	**14.1**		7.4	10.5	10.3	7.2	7.2	7.2	7.9
23	**14.25**		7.5	10.6	10.4	7.2	7.2	7.2	7.9
21	**14.9**		7.6	10.8	10.7	7.3	7.3	7.3	8.0
12	**15.1**		7.6	10.8	10.8	7.4	7.4	7.4	8.1
5	**15.1**		7.6	10.8	10.8	7.4	7.4	7.4	8.1
16	**15.5**		7.7	10.9	10.9	7.4	7.4	7.4	8.1
26	**16.25**		7.8	11.1	11.2	7.6	7.6	7.6	8.2
19	**18.5**	**7.6**	**8.1**	**11.6**	**12.0**	**7.9**	**7.9**	**7.9**	**8.5**
14	**18.7**		8.1	11.7	12.1	7.9	7.9	7.9	8.5
11	**19.1**		8.2	11.7	12.2	8.0	8.0	8.0	8.5
13	**19.95**		8.3	11.9	12.5	8.1	8.1	8.1	8.6
2	**20.3**		8.3	11.9	12.6	8.1	8.1	8.1	8.6
10	**20.55**		8.3	12.0	12.7	8.1	8.1	8.1	8.6
15	**21.1**		8.4	12.1	12.9	8.2	8.2	8.2	8.7
4	**21.25**		8.4	12.1	12.9	8.2	8.2	8.2	8.7
20	**21.65**	**8.2**	**8.4**	**12.2**	**13.0**	**8.3**	**8.3**	**8.3**	**8.7**
9	**21.7**		8.4	12.2	13.0	8.3	8.3	8.3	8.7
8	**22.55**		8.5	12.3	13.3	8.4	8.4	8.4	8.8
6	**22.75**		8.5	12.3	13.3	8.4	8.4	8.4	8.8
1	**23.05**	**8.85**	**8.6**	**12.3**	**13.4**	**8.4**	**8.4**	**8.4**	**8.8**
7	24		8.6	12.5	13.7	8.5	8.5	8.5	8.8

Note: In bold are measured values; trees measured for height are identified with gray background

Prediction Results

Table 11. Mean residuals and mean absolute residuals of height estimation in a NFI plot from Lomba ZIF

Model	*L_mixed*	*L_Fixed*	*L_NLS*	*G_mixed*	*G_Fixed*	*G_NLS*	*IFN5*
Bias (m)	-0.152	-3.819	-4.604	0.030	0.030	0.031	-0.439
MAR (m)	0.351	3.819	4.604	0.270	0.271	0.267	0.482

Bias – mean of residuals
MAR – mean of absolute residuals

L_mixed - Local mixed model with random effects
L_Fixed - Local mixed model with fixed effects only
L_NLS - Local model from nonlinear least squares fit
G_mixed - Generalized mixed model with random effects
G_Fixed - Generalized mixed model with fixed effects only
G_NLS - Generalized model from nonlinear least squares fit

IFN5 – Generalized model used in IFN5

Table 12. Mean of prediction residuals (Bias) and mean of absolute prediction residuals (MAR) of tree height estimates in inventory plots from Nordeste region (outside Lomba ZIF).

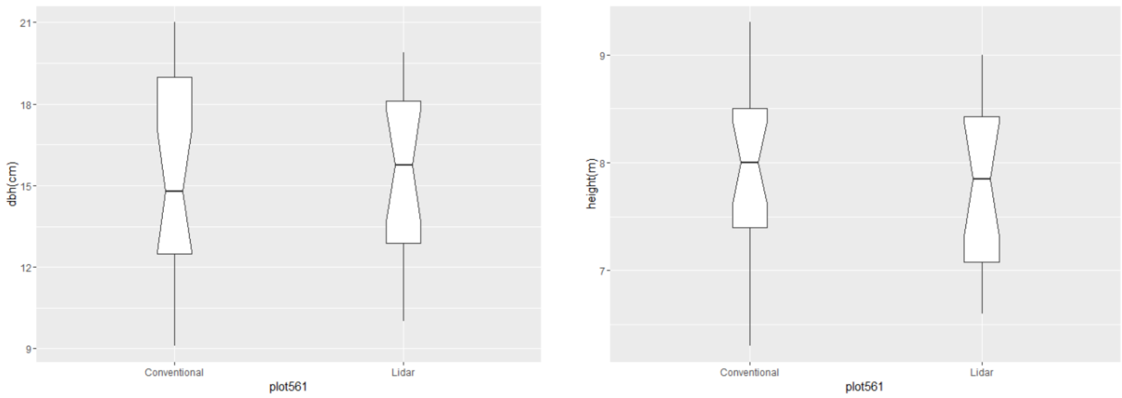
Stand structure					Bias (m)	Bias (m)	Bias (m)	MAR (m)	MAR (m)	MAR (m)
Nv (trees/ha)	***hdom (m)***	***dg (cm)***	***ddom (cm)***	***t (years)***	***(G_NLS)***	***(IFN5)***	***GG***	***(G_NLS)***	***(IFN5)***	***GG***
840	12.8	20.5	30.51	30	0.80	0.03	0.68	1.19	1.12	1.11
300	12.9	23.2	29.2	31	0.48	-0.05	0.37	1.44	1.31	1.37
900	6.29	10.40	13.60	15	0.06	-0.14	0.04	0.24	0.34	0.24
160	11.03	24.57	27.00	30	0.19	-0.36	0.14	0.53	0.48	0.47
1199	8.90	13.43	16.90	17	-0.13	-0.21	-0.16	0.38	0.40	0.41
3658	11.10	11.42	18.38	30	-0.15	-0.23	-0.03	0.37	0.35	0.24
260	8.20	14.20	19.80	18	-0.01	-0.25	-0.03	0.69	0.72	0.69
200	6.70	15.70	17.80	15	-0.12	-0.54	-0.18	0.90	1.18	0.98
840	5.90	11.10	11.60	14	0.18	0.19	0.18	0.50	0.52	0.50
840	15.50	23.60	27.30	27	0.85	1.00	0.76	1.23	1.32	1.19

Nv - number of living trees per hectare; hdom - dominant height; dg - quadratic mean diameter; ddom - dominant diameter; t – Age; G_NLS refers to the generalized *h-d* model G7; IFN5 refers to the generalized h-d equation used in IFN5; GG refers to the generalized *h-d* equation (eqn 5 in the original work) of Gómez-García et al. (2015)

Appendix 3

Results of comparing conventional and LiDAR measurements of *h* and *d*

Plot 561 notched box-plots for *d* (left) and *h* (right):

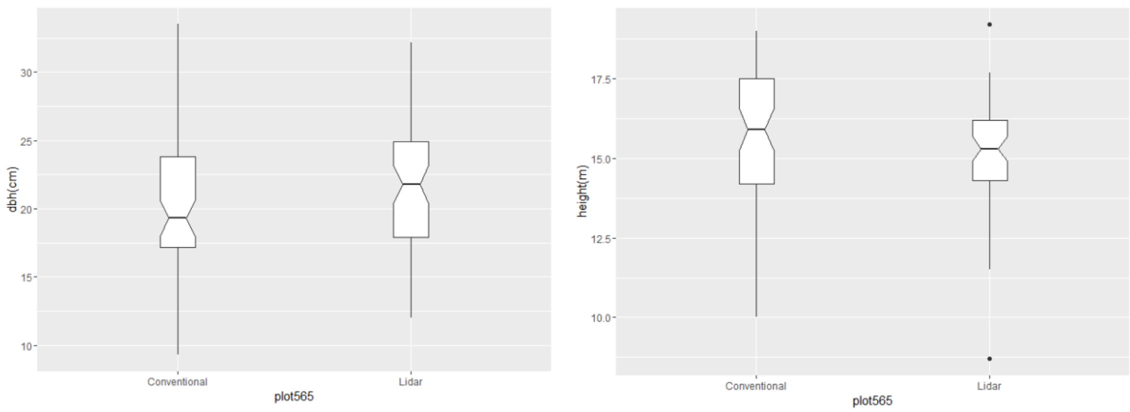


Plot 561 statistics for *d* and *h*:

Statistics	d (cm)		h (m)	
	Conventional	*LiDAR*	*Conventional*	*LiDAR*
Mean	15.1 **a**	15.3 **a**	7.9 **a**	7.8 **a**
Max.	21.0	19.9	9.3	9.0
Min.	9.1	10.0	6.3	6.6
Sd.	3.8	3.2	0.9	0.8
n	21	16	21	16

a – *t-test* results (same letters between Conventional and LiDAR measurements mean no-signifficant differences at 5% level). This applies also to the other plots.

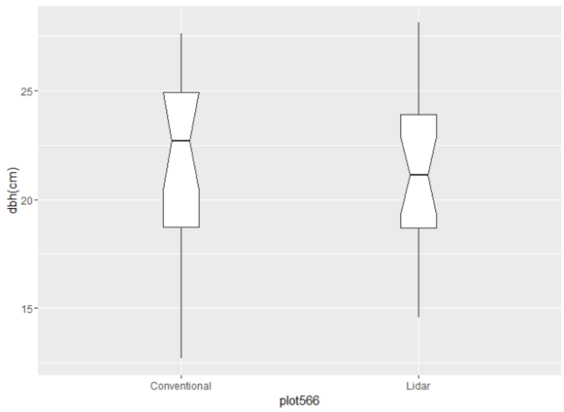
Plot 565 notched box-plots for *d* (left) and *h* (right):

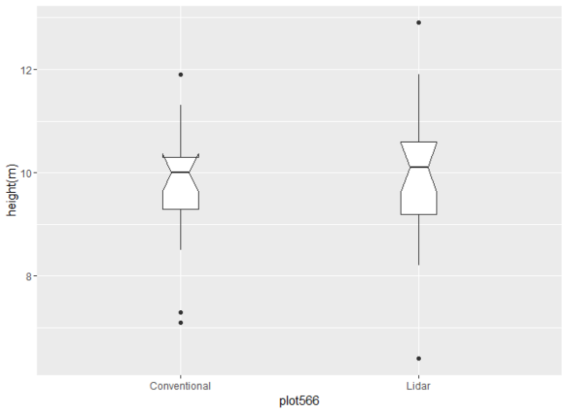


Plot 565 statistics for *d* and *h*:

Statistics	d (cm)		h (m)	
	Conventional	*LiDAR*	*Conventional*	*LiDAR*
Mean	20.4 **a**	21.8 **a**	15.6 **a**	15.1 **a**
Max.	33.5	32.2	19.0	19.2
Min.	9.3	12.0	10.0	8.7
Sd.	5.0	5.0	2.1	1.8
n	61	61	61	61

Plot 566 notched box-plots for *d* (left) and *h* (right)

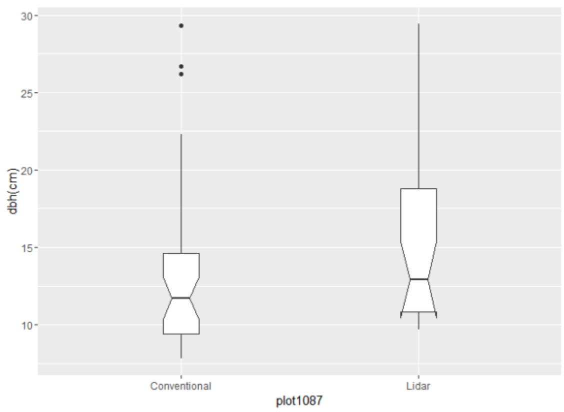


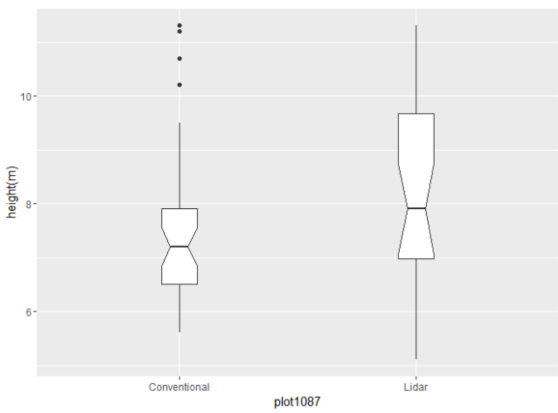


Plot 566 statistics for *d* and *h*:

Statistics	d (cm)		h (m)	
	Conventional	*LiDAR*	*Conventional*	*LiDAR*
Mean	21.7 **a**	21.6 **a**	9.8 **a**	9.9 **a**
Max.	27.6	28.1	11.9	12.9
Min.	12.7	14.6	7.1	6.4
Sd.	4.1	3.9	1.2	1.4
n	19	21	19	21

Plot 1087 notched box-plots for *d* (left) and *h* (right)

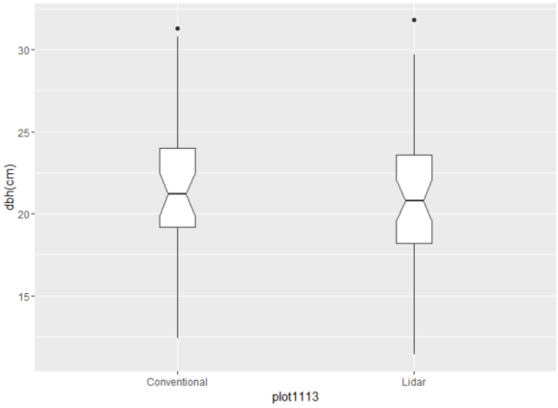


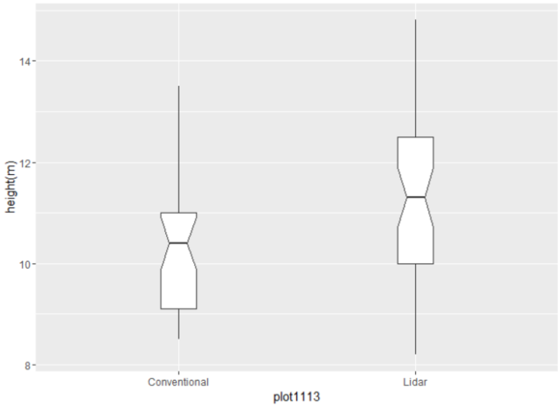


Plot 1087 statistics for *d* and *h*:

Statistics	d (cm)		h (m)	
	Conventional	*LiDAR*	*Conventional*	*LiDAR*
Mean	13.4 **a**	15.4 **a**	7.6 **a**	8.2 **a**
Max.	29.3	29.4	11.3	11.3
Min.	7.8	9.7	5.6	5.1
Sd.	5.6	5.8	1.6	1.6
n	37	26	37	26

Plot 1113 notched box-plots for *d* (left) and *h* (right)





Plot 1113 statistics for *d* and *h*:

Statistics	d (cm)		h (m)	
	Conventional	*LiDAR*	*Conventional*	*LiDAR*
Mean	21.7 **a**	20.5 **a**	10.4 **a**	11.3 **b ***
Max.	31.3	31.8	13.5	14.8
Min.	12.4	11.4	8.5	8.2
Sd.	4.9	4.6	1.4	1.6
n	33	45	33	45

** no differences are detected if considering the 1% significance level*