ESTIMATING TREE BOLE HEIGHT WITH BAYESIAN ANALYSIS

E. Milios 1, K. Kitikidou 2, E. Pipinis 3, A. Stamoulidis 4, M. Gotsi 5

1,2,4,5 Dep. of For. & Manag. of the Envi. & Nat. Res., Demokritus University of Thrace, Orestiada, Greece
3 Fac. of Agric., For. & Nat. Envi. School of For., The Aristotle University of Thessaloniki, Thessaloniki, Greece

ABSTRACT. The classical method for estimating a height-diameter model is based on the Least Squares Method (LSM) and the fit of a regression line. The Bayesian method has an exclusive advantage, compared with the classical method, in that the parameters to be estimated are considered as random variables. In this study, the Simple Linear Regression (SLR) model and the Bayesian model were used to estimate bole height from breast height diameter. We used data of the forest stands of Rhodope (north-eastern Greece), variables that we used were the tree bole height and the diameter at breast height. The results showed that there is an improvement in prediction accuracy with the Bayesian model; however, this didn’t lead to narrower confidence intervals of the predicted value, compared to SLR. Narrower confidence intervals are not necessarily achieved with Bayesian methods; confidence intervals’ width is related to both statistical analyses and nature of data (in this case, species ecology and structure - composition of the stands where the sampled trees belong).

Keywords: Bayes, regression; tree height; tree bole.

1 INTRODUCTION

Forests play an important role in the wood production, recreation, carbon sinks and climate change (FAO 2006). Height and diameter are the most common variables measured to estimate tree volume, site quality and other important variables in forest growth and yield, species’ succession and carbon estimation models (Peng et al. 2001). Breast height diameter of individual trees is easy to measure with high accuracy in the field, and minimum cost. On the contrary, tree height is not easy to estimate on standing trees; it is a time-consuming procedure, subjected to observation error, and affected by optical obstacles (Colber et al. 2002). Tree volume estimation and site quality, as well as the description of a stand’s dynamics and species’ succession over time, are greatly dependent by the accuracy of height-diameter models (Curtis 1967). There is a series of height-diameter models, developed for several forest species (Peng et al. 2004, Temesgen et al. 2007, Van Laar & Akca 2007). These models can be used to estimate a missing height for a tree, while having the diameter at breast height measured (Shongming et al. 1992, Hann 2006), to estimate indirectly the height increment (Larsen & Hann 1987), and to estimate tree biomass applying biomass models (Peng et al. 1997). Chave et al. (2005) found that the most important parameters in biomass estimation of tropical forests, in decreased order, are the breast height diameter, the wood density, the height and the forest’s moisture (dry forest, forest with medium or heavy moisture). Comprising height as a variable, in biomass models, reduces significantly the standard error of biomass estimation. Therefore, an accurate tree height estimation should be imposed in forest inventories, simulation models, forest management and decision support systems (Peng et al. 2001, Colber et al. 2002, Curtis 1967).

Curtis (1967) compared the fit of linear height-diameter models, in data of Douglas-fir trees (Pseudotsuga menziesii (Mirb.) Franco). Since then, with the relatively easy fit of nonlinear models, many nonlinear models have been developed for estimating height (Fekedulegn et al. 1999). However, because the tree shape and allometry are affected by competition and environmental factors (Batzliou et al. 2016, King 1991, Kitikidou et al. 2016), any changes in these conditions over time will probably affect the height-diameter relationship. This can cause uncertainty in height estimation from the diameter. An important limitation in these models is that they produce very different results, when applied in different stands from those where
they originally were developed (Calama & Montero 2004, Nogueira et al. 2008). Also, the height-diameter relationship is not stable over time, even in the same stand (Flewelling & Jong 1994, Lappi 1991). These differences could have significant effects in biomass and dynamic carbon sinks. Uncertainty in height estimation, due to time changes, must be taken into account, when height-diameter relationships in natural stands are examined, and there are no available methods to resolve this problem.

From all these points previously described, one can assess that, even a small improvement in height-diameter models, is worthy of investigation, especially regarding bole height, which is referring to merchantable volume. Merchantable volume is the main interest of forest inventories and the most contributing variable in forest biomass and carbon sinks. Bayesian analysis is an alternative method of statistical inference, frequently used in ecological models’ evaluation (Anholt et al. 2000, Toivonen et al. 2001, Shen et al. 2003). In forestry, Bayesian methods were adopted in several applications, like the estimation of aboveground biomass (Zapata-Cuartas et al. 2012), in diameter distribution (Bullock & Boone 2007) and basal area distribution (Nyström & Stål 2001), in tree growth estimation (Clark et al. 2007), mortality estimation (Wyckoff & Clark 2000, Metcalf et al. 2009), stands’ height and volume estimation (Stewart & Weiskittel 2012), and in individual tree height estimation (Zhang et al. 2014).

_Fagus sylvatica_ L. expands into a wide area of Europe and up to southern Scandinavia, but in the Mediterranean area appears only in the mountains (Korakis 2015). _Fagus sylvatica_ is a very important tree species for Greece, since beech forests compose 5.17% of Greek forests, while beech stands provide 20.05% of the merchantable volume of forests in the country (Ministry of Agriculture 1992).

In this study, we have developed Bayesian height-diameter models for beech trees, examining numerous trials from one wide study area. Since nonlinearity in a height-diameter relationship is not always detectable within an even-aged stand, due to sample sizes that cannot display lack of fit of the linear model, or excessive random variability of heights within certain diameter classes (Van Laar & Akca 2007), Simple Linear Regression (SLR) models were developed. We compared the Bayesian height-diameter models to their corresponding SLR models (i.e. regression models developed with Least Squares Method (LSM)), in order to see if there is any improvement in height estimation.

2 MATERIALS AND METHODS

2.1 Study area

The study was conducted in the central part of the Rhodope mountains, in the north part of the Xanthi region, in north-eastern Greece (Fig. 1). The climate can be characterized as humid with harsh winters. The soils in the areas where data was collected are mainly acid brown forest soils (Dystric Cambisols) (Milios 2000). The data were taken from areas having an elevation that ranges from approximately 580 to 1700 m.

![Figure 1: Study area (wide view).](image1)

The elevation of the western part of the study area (Fig. 2, red polygon points) ranges from 1100 to 1725 m, while that of the eastern part (Fig. 2, green polygon points) ranges from 580 to 750 m.

![Figure 2: Study area (large display).](image2)

In the total study area, 2809 _F. sylvatica_ trees were measured in plots of 500 m$^2$ (25 m x 20 m). The plots had been established randomly, in the context of other...
Table 1: Descriptive statistics of measured variables

<table>
<thead>
<tr>
<th>Trial No. (Productivity)</th>
<th>Trees No. (n)</th>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2809</td>
<td>D (cm)</td>
<td>14.78</td>
<td>10.95</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.65</td>
<td>3.63</td>
<td>0</td>
<td>23.5</td>
</tr>
<tr>
<td>2</td>
<td>1495</td>
<td>D (cm)</td>
<td>12.4</td>
<td>8.75</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.02</td>
<td>2.43</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>1398</td>
<td>D (cm)</td>
<td>13.4</td>
<td>10.32</td>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.26</td>
<td>3.13</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>2157</td>
<td>D (cm)</td>
<td>13.48</td>
<td>9.99</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3</td>
<td>2.95</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>753</td>
<td>D (cm)</td>
<td>13.48</td>
<td>9.28</td>
<td>4</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>2.39</td>
<td>2.42</td>
<td>0</td>
<td>14.5</td>
</tr>
<tr>
<td>6</td>
<td>210</td>
<td>D (cm)</td>
<td>27.85</td>
<td>13.72</td>
<td>4</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>8.98</td>
<td>6.31</td>
<td>1</td>
<td>23.5</td>
</tr>
<tr>
<td>7</td>
<td>1526</td>
<td>D (cm)</td>
<td>14.24</td>
<td>11</td>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.53</td>
<td>3.45</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>2367</td>
<td>D (cm)</td>
<td>14.75</td>
<td>11.15</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.8</td>
<td>3.75</td>
<td>0</td>
<td>23.5</td>
</tr>
<tr>
<td>9</td>
<td>754</td>
<td>D (cm)</td>
<td>16.69</td>
<td>11.56</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>4.7</td>
<td>4.24</td>
<td>0.5</td>
<td>23.5</td>
</tr>
<tr>
<td>10</td>
<td>1968</td>
<td>D (cm)</td>
<td>14.39</td>
<td>10.75</td>
<td>4</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hb (m)</td>
<td>3.38</td>
<td>3.33</td>
<td>0</td>
<td>20</td>
</tr>
</tbody>
</table>

studies, for the analysis of stand structure and dynamics in pure and mixed stands in areas of various productivities. In each tree, the breast height diameter and the bole height were measured. In particular, the measured trees came from pure beech stands in good and medium productivity sites, from *Pinus sylvestris* – *F. sylvatica* stands growing in good, medium and poor productivity sites, from *Fagus sylvatica* – *Abies borsii-regis* stands growing in medium productivity sites, and from *Quercus petraea* – *Fagus sylvatica* found in medium productivity sites (Milios 2000). In *P. sylvestris* – *F. sylvatica* stands, *P. sylvestris* was the dominant species of the overstorey in almost all cases (Milios 2000). From different mixtures and productivity sites, ten different combinations (trials) were created for the analysis. Descriptive statistics of measured variables, in each trial, are given in Table 1.

2.2 Trial No (Productivity)

1. All

2. Good for *Pinus sylvestris* - *Fagus sylvatica*, medium for *Pinus sylvestris* - *Fagus sylvatica*, and poor for *Pinus sylvestris* - *Fagus sylvatica*


5. Medium for *Quercus petraea* - *Fagus sylvatica* (western part), and medium for *Quercus petraea* - *Fagus sylvatica* (eastern part)

6. Good for *Fagus sylvatica*, and medium for *Fagus sylvatica*

7. Medium for *Pinus sylvestris* - *Fagus sylvatica*, medium for *Fagus sylvatica* - *Abies borsii-regis*, medium for *Quercus petraea* - *Fagus sylvatica* (western part), and medium for *Fagus sylvatica*

2.3 SLR models developed with LSM

In a scatterplot with an independent (X) variable and a dependent (Y) variable, the goal of a linear regression model is to fit a line through the points. Specifically, with LSM, the squared deviations of the observed points from that line are minimized.

2.4 Bayesian model

Suppose \( y = (y_1, y_2, y_3, \ldots) \) is a vector of data and \( \theta = (\theta_1, \theta_2, \theta_3, \ldots) \) is a vector of parameters, which will be estimated. The Bayes rule is expressed as follows:

\[
p(\theta | y) = \frac{p(y | \theta) p(\theta)}{p(y)}
\]

where \( p \) is a density probability function. While the values of \( \theta \) are estimated with LSM or Maximum Likelihood Estimation in the classical approach, in the Bayesian approach we use probability distributions to describe the uncertainty in the parameters that are about to be estimated. The \( \theta \) have a probability distribution, which is calculated as another form of (1):

\[
p(\theta | y) = \frac{p(y | \theta) p(\theta)}{p(y)}
\]

where \( p(y) = \int p(y | \theta) p(\theta) d\theta \) for continuous \( \theta \). The integration of the acceptable values \( \theta \), \( p(y) \) is not dependent on \( \theta \) and can be considered as constant for constant \( y \), a fact that leads to the relationship (3):

\[
p(\theta | y) \propto p(y | \theta) p(\theta)
\]

What we are interested in the Bayesian analysis is the estimation of the conditional probability, i.e., the posterior probability distribution. The \( p(y | \theta) \) is giving us the distribution of \( y \) assuming that the \( \theta \) is known, i.e., it is the function of maximum likelihood when it is considered as a function of the \( \theta \) parameters (Edwards 1992). The \( p(\theta) \) is the prior probability distribution of the \( \theta \) parameters, and it represents all available information regarding \( y \). So, the relation (3) is suggesting that the posterior distribution of \( \theta \) is analogous to the likelihood of \( y \), given the \( \theta \) and the prior distribution of \( \theta \).

The important characteristic of the Bayesian analysis is that the models’ parameters are considered to be random variables (Stewart & Weiskittel 2012), while in the classical method of Least Squares, the models’ parameters are considered fixed values (Edwards 1992, De Valpine & Hastings 2002).

The selection of the prior distribution is essential in the Bayesian analysis (Gelman et al. 2004). If there is no available information regarding parameters’ distribution, we can accept ignorance of the prior distribution, i.e., accept an uninformative, Gaussian prior distribution.

Bayesian parameters are estimated by applying the SPSS AMOS (Analysis of Moment Structures) software, v.21.0 (Arbuckle 2012). We defined to run 100000 iterations, from which the initial 500 were considered as initial stages of the chain before it converges (burn-in period).

2.5 Comparison of the classical and the Bayesian method

Three statistical criteria were used for the classical (SLR - LSM) and the Bayesian method comparison (Kikitidou 2005):

- Absolute mean error \( \text{Bias} = \frac{\sum_{i=1}^{n} |H_{bi} - \hat{H}_{bi}|}{n} \) (optimum value = 0).
- Standard error of the estimate of theoretical values \( se = \sqrt{\frac{\sum_{i=1}^{n} (H_{bi} - \hat{H}_{bi})^2}{n-p}} \) (optimum value = min).
- Coefficient of determination \( R^2 = 1 - \frac{\sum_{i=1}^{n} (H_{bi} - \hat{H}_{bi})^2}{\sum_{i=1}^{n} (H_{bi} - \overline{H}_b)^2} \) (optimum value = 1).

where:

- \( H_{bi} \) = bole height of the \( i \)-th tree
- \( \hat{H}_{bi} \) = estimated bole height of the \( i \)-th tree
- \( \overline{H}_b \) = mean bole height of the sampled trees
- \( n \) = sample size (number of trees in each trial)
- \( p \) = number of the model’s parameters = 2.

3 Results

From the ten trials of Table 1, bole height estimation was improved by applying the Bayesian method, in all of them. Comparison statistics for these ten trials, the estimated mean bole height, its confidence interval, and variance, are shown in Table 2, while SLR are shown in Table 3. We observe that in general (i.e. except for trials 2, 6, and 9), the Bayesian models gave bigger values for the estimated mean bole height, wider confidence intervals, and bigger variance.

In the trace plots of Figures 3a-3j, the convergence that was succeeded by running the algorithm, is assessed. The fast, up-and-down interchange, without trends (patterns), shows that the convergence was succeeded in a short time.
Figure 3: Trace plot of the variance of the breast height diameter \( D \) (a: Trial 1, b: Trial 2, c: Trial 3, d: Trial 4, e: Trial 5, f: Trial 6, g: Trial 7, h: Trial 8, i: Trial 9, j: Trial 10).

Table 2: Statistics for the trials in which bole height estimation was improved with the Bayesian method (SE – Standard Error; LB – Lower Bound; UB – Upper Bound; CI – Confidence Interval; EM – Estimated Mean; No – Number; ht – height).

<table>
<thead>
<tr>
<th>Trial No</th>
<th>No of trees (n)</th>
<th>Classical method (SLR, LSM)</th>
<th>Bayesian method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>No trees</td>
<td>Bias</td>
</tr>
<tr>
<td>1</td>
<td>2809</td>
<td>2.103</td>
<td>3.009</td>
</tr>
<tr>
<td>2</td>
<td>1495</td>
<td>1.593</td>
<td>2.1891</td>
</tr>
<tr>
<td>3</td>
<td>1398</td>
<td>1.756</td>
<td>2.5304</td>
</tr>
<tr>
<td>4</td>
<td>2157</td>
<td>1.798</td>
<td>2.2673</td>
</tr>
<tr>
<td>5</td>
<td>753</td>
<td>1.558</td>
<td>5.2663</td>
</tr>
<tr>
<td>6</td>
<td>210</td>
<td>4.286</td>
<td>5.2663</td>
</tr>
<tr>
<td>7</td>
<td>1526</td>
<td>1.898</td>
<td>2.7464</td>
</tr>
<tr>
<td>8</td>
<td>2367</td>
<td>2.96</td>
<td>3.0081</td>
</tr>
<tr>
<td>9</td>
<td>754</td>
<td>2.589</td>
<td>3.5377</td>
</tr>
<tr>
<td>10</td>
<td>1968</td>
<td>1.953</td>
<td>2.7852</td>
</tr>
</tbody>
</table>

In Figures 4a-4j, the autocorrelation of the values of the variance of \( D \), during iterations, is illustrated. The lag across the horizontal axis is referred to the interval in which the autocorrelation is estimated. In common situations, we expect that the autocorrelation coefficient is reducing till zero, and stays close to zero in all the following lag intervals. As illustrated on Figures 4a-4j, the initial 500 iterations (burn-in period) were more than enough, to ensure convergence.
Figure 4: Autocorrelation of the variance of the breast height diameter $D$ (a: Trial 1, b: Trial 2, c: Trial 3, d: Trial 4, e: Trial 5, f: Trial 6, g: Trial 7, h: Trial 8, i: Trial 9, j: Trial 10).

Table 3: SLR models for each trial.

<table>
<thead>
<tr>
<th>Trial No</th>
<th>Classical method (SLR, LSM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\hat{H}_b = 0.911 + 0.186D$</td>
</tr>
<tr>
<td>2</td>
<td>$\hat{H}_b = 1.512 + 0.122D$</td>
</tr>
<tr>
<td>3</td>
<td>$\hat{H}_b = 0.868 + 0.179D$</td>
</tr>
<tr>
<td>4</td>
<td>$\hat{H}_b = 1.207 + 0.155D$</td>
</tr>
<tr>
<td>5</td>
<td>$\hat{H}_b = 1.140 + 0.093D$</td>
</tr>
<tr>
<td>6</td>
<td>$\hat{H}_b = 1.892 + 0.254D$</td>
</tr>
<tr>
<td>7</td>
<td>$\hat{H}_b = 0.819 + 0.190D$</td>
</tr>
<tr>
<td>8</td>
<td>$\hat{H}_b = 0.841 + 0.201D$</td>
</tr>
<tr>
<td>9</td>
<td>$\hat{H}_b = 1.331 + 0.202D$</td>
</tr>
<tr>
<td>10</td>
<td>$\hat{H}_b = 0.942 + 0.169D$</td>
</tr>
</tbody>
</table>

In Figures 5a-5j, the classical (SLR, LSM) vs. the Bayesian model is illustrated, for each trial, in which the Bayesian model improved the bole height estimation.

4 Discussion and Conclusions

Estimating bole height from breast height diameter, applying the Bayesian method to the classical regression method (SLR, LSM) lead to an improvement of estimation accuracy (an increase of the coefficient of determination $R^2$), in all our trials, conducted in the same wide study area. However, the confidence intervals of the estimated mean bole height were larger, a fact that is in contradiction to the study of Zhang et al. (2014). This is the result of the combination of the species ecology and the structure – composition of the stands where the measured trees came from. Beech is a shade tolerant species having growth plasticity (Assman 1970). As a result, in mixed stands where beech grows together with less shade tolerant species that dominate in the overstorey, beech trees in the understory and middle story have enough light to grow a long live crown. Shade intolerant species do not have as dense crowns as shade tolerant species, and they generate less shade, since, in contrast
Figure 5: Height-diameter model (a: Trial 1, b: Trial 2, c: Trial 3, d: Trial 4, e: Trial 5, f: Trial 6, g: Trial 7, h: Trial 8, i: Trial 9, j: Trial 10).
to shade tolerant species, after their youth, they cannot form shade leaves (see Dafis 1986, Oliver & Larson 1996, Lacointe et al. 2004). Thus, when shade-tolerant trees with large diameters of the overstorey, grow with shade intolerant species, they grow longer crowns compared to those growing in pure stands, or mixed stands having other shade tolerant species since more light reaches the lower part of boles. In the present study, this is obvious in all the trials where beech grows with P. sylvestris or Q. petraea. Pinus sylvestris is shade intolerant, while Q. petraea is moderately shade-intolerant (Korakis 2015). In these trials, the Bayesian method overestimates the bole height of trees having large diameters, compared to the classical regression method, leading to larger confidence intervals (Figs. 5a-5c, Figs. 5g-5j). Only in trial 6, which is the only trial with pure beech stands (and crowns of the trees with large diameters are short) the Bayesian method does not overestimate the bole height (Fig. 5f). So, the Bayesian method does not give a priori narrower confidence intervals; confidence intervals’ width is related to the characteristics of the data.

Another finding of the present study is that in both methods (classical and Bayesian), in trials from mixed stands of P. sylvestris - F. sylvatica and/or Q. petraea - F. sylvatica (trials 2 and 5) the R^2 values are lower than in the other trials. One should note that, in trials with pure beech stands (1, 6, 7, 8, 9 and 10) the R^2 values in the Bayesian method are rather high, while in some of these trials (1, 8, 9 and 10) the increase of the R^2, compared to the classical method, is very high. The lowest R^2 value of these trials is found in trial 6, where trees only from pure stands participate. This pattern indicates that bole height prediction is more accurate where trees from all growth environments are present. However, even in the case where only trees from pure beech stands are present, the Bayesian method provides a rather high R^2 value. The behaviour of these models is affected by the growth characteristics and ecology of beech, which have been described previously. Probably the existence of trees from pure beech stands provides an adequate number of large trees having high bole height in the sample (small crown length is a result of competition among shade-tolerant trees), improving the model’s performance.

In trial 9 (good productivity sites) the R^2 value of 0.56 is an indication that growth conditions in different site productivities may influence the model’s performance. However, in trial 10 (medium productivity sites) the R^2 value was lower than that of trial 1.

To sum up, the Bayesian method is an important tool, used more and more by ecologists (Hui et al. 2006, Hui et al. 2011, McCarthy et al. 2007). For this specific application in bole height estimation, more independent variables could be added in the future, such as site quality index, age, or stand density (Tenesgen & Gadow 2004, Newton & Amponsah 2007).

ACKNOWLEDGEMENTS

Thanks are due to the anonymous reviewers and the MCFNS editors for their helpful comments improving the manuscript.

REFERENCES


Ministry of Agriculture. 1992. Results of the rst national forest inventory (in Greek), Publication of Department of Forest Mapping, 134 p, Athens, Greece.


