Available online at sjuoz.uoz.edu.krd

journals.uoz.edu.krd

Science Journal of University of Zakho

Vol. 11, No. 1, pp. 84– 90, January-March 2023

p-ISSN: 2663-628X
e-ISSN: 2663-6298

QUANTIFICATION OF THE RELATIONSHIP BETWEEN THE HEIGHT – DIAMETER OF (Pinus brutia Ten), CALABRIAN PINE TREES IN FOUR DIFFERENT MICROSITES IN DUHOK GOVERNORATE. KURDISTAN REGION, IRAQ.

Tariq K. Salih[a, *], Hayfa M. Saeed Abdulaziz [b]

[a] College of Agricultural Engineering Science, University of Duhok, Duhok, Kurdistan Region, Iraq - (tariq.salih@uod.ac)
[b] Directorate of Forests and Rangelands, Duhok, Kurdistan Region, Iraq - (haifawafa2@gmail.com)

Received: *13 Sep., 2022 /* ***Accepted:*** *18 Nov., 2022 /* ***Published:*** *30 Jan., 2023* ***https://doi.org/10.25271/sjuoz.2022.11.1.1015***

ABSTRACT

A sample of 120 trees was purposely selected from four different microsites (Behere, Swaratoka, Zawita, and Semel), (30 trees from each microsite) in the Duhok governorate. 100 of which were used for model calibration and the rest were used for validation of the selected regression equation. This study consisted of two main parts. In the first part ,In the beginning, a scatter diagram was conducted for each microsite to detect the type of relationship between height and diameter at breast, which in turn will decrease the number of regression models that will be tested. Accordingly, 4 regression equations were developed for each of the four microsites separately. These models can be used to see how the ratio of height diameter in each of the studied sites is inter-correlated and which of them is the most appropriate for producing Calabrian pine trees. In the second part, all microsites were then treated as one sample for estimating the parameters of 25 regression models using Excel and Statographic packageThe developed regression models underwent a screening process in order to find the most appropriate one to be used for the prediction of the height of Calabrian pine grown in the four mentioned microsites in Duhok governorate. Many measures of precision among them coefficient of determination, Bias%, Mean absolute error, Ohtomo's unbiased test, Furnival Index, and AIC criterion were used for testing the performance of the developed equations in the prediction of the height. At last, the equation $\sqrt{H} = 1.3 + 0.3825\sqrt{D}$ was selected, as the best regression model. This equation shows that there is a linear relationship between $\sqrt{H}$ *and* $\sqrt{D}$ and a curvilinear relationship between D and H. The study showed that the height–diameter ratio was highest in Swaratoka and was the most appropriate microsite among the rest, followed by Behere, then Zawita and Semel came in last place.

KEYWORDS: Calabrian Pine Trees, Height – Diameter models, Height – Diameter relationship, Modeling Validation.

1. INTRODUCTION

Calabrian pine is the most important coniferous forest tree that is grown naturally in Duhok governorate, Kurdistan region, Iraq (Shahbaz, 2010). It has been widely used in afforestation and reforestation in different parts of Iraq and specially in Kurdistan region, such as Zawita, Atrosh disrricts and Azmer mountains. It is well adapted to the regions site and climate conditions and therefore is even used in afforestation of parks, city roads and the roads between cities. In addition to wood production, it can provide a huge number of services to habitation, ecosystem and wildlife. It decreases the rate of runoff water and thereby decrease soil erosion, purification of air and water, supply shade to both human being and animals in hot summers.

The height of a tree is a very important tree attribute that comes in the next rank in importance after diameter at breast height. It can be used in many concepts, among them in models describing the relationship between total tree height and diameter at breast height, which is an extremely valuable tool for forest management planning, in site index determination, and it is used with breast height diameter for construction of standard volume table. Such a relationship can be used to convert a standard volume equation to a local volume equation.

The data used for conducting the height diameter studies are either constructed from stem analysis (Dyer and Bailey, 1987; Kariuki, 2002; Sumida et al, 2013) or direct measurement of pairs of diameter at breast height and heights of trees (Carron, 1968; Larsen and Haan, 1987; Avery and Burkhart 2015; Amin, 2016; Kershaw, 2016). Height–diameter models provide predictions of tree heights based on the measured tree stem diameters (usually stem diameter at breast height). They are needed for quantifying the growing stock in conducting forest inventory (Ng'andwe, et al 2021), and in the determination of biomass, and carbon sequestration (Mensah et al, 2018). Many investigations have been conducted on different types of relationships between the height and diameter of trees and many of them have focused on the nonlinear relationship (Philip 1994; Huang et al, 2000; Huang et al, 2009; El mamoun et al, 2013). The height-diameter relationship differs from one stand to another due to differences in site quality, stand age, and silvicultural treatments, and even within the same stand due to differing competitive situations among the trees (e.g., Calama and Montero, 2004; Sharma and Parton, 2007; Schmidt et al, 2011). The height-diameter relationship is thus highly site-specific and stands density-specific, and varies over time even within the same stand (e.g., Zeide and Vanderschaaf, 2002; Adame et al. 2008; Pretzsch 2009). The height-diameter curve increases faster for small DBH than for larger DBH (e.g., Lappi, 1997; Pretzsch 2009; Salih 2019). Chai et al (2018) proposed using nonlinear regression models with 2-4 parametric forms as the best one to describe the relationship between the height and diameter.

This study is aimed to:

1. Developing height diameter relationship in four different microsites in Duhok governorate, of which one contains naturally grown trees, two are plantations established in the mountainous region, and the last one is a plantation in a hilly area.

* Corresponding author

2. To see if there is a significant difference in the height–diameter relationship between these regions or not.
3. Analysis of the parameters of the H/D curves can be used to determine how well a species of trees are adapted to the site.

2. MATERIALS AND METHODS

2.1 study area and geographical description

The data used in this study were collected from four different microsites Table 1 and Figure 1.

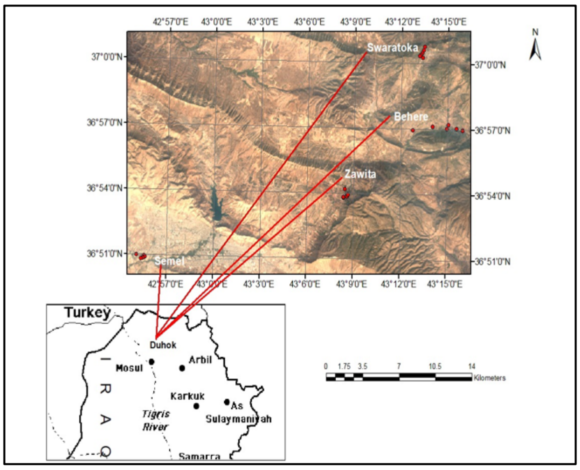


Figure 1 the location of the study area.
Source of picture: (Source: Google map. 2022. Duhok. 1:5. Google Maps (online) Available through: Google maps https://www.google.com/maps/@36.982324,43.1321247,33644m/data=!3m1!1e3 (Accessed 13 September. 2022), generated using Arc Map 10.3

Table 1: Geographical information about studied area

Micro-sites	Type	Altitude	Latitude	Longitude	Precipitation
1-Zawita	Natural forest	862	36°53'76.3"	43°08'16.58"	770.62
2-Behere	Plantation/ Mountainous region	1000	36°53'54.9"	43°14'59.93"	874.36
3-Swaratoka	Plantation/ Mountainous region	1200	37°00'13.5"	43°13'10.22"	1074.2
4-Semel	Plantation/Hilly region	490	36°50'52.3"	42°55'17.75"	616.29

2.2 Data collection

About 30 trees were purposely selected from each of the four mentioned microsites, of which 25 trees were used for calibration of regression equations and 5 trees (Vanderchaaf 2008) for validation. Hence a total of 100 trees and 20 trees were used for calibration and validation respectively. The following data were collected for each tree:

1. Breast height diameter in cm by diameter tape.
2. Total height in m up to one decimal using Haga Altimeter.

These pairs of data constitute the main basic data for such a study.

2.3 Model Development

This study consisted of two parts:

In the first part, the parameters of 4 regression equations were estimated using the data collected from each of the microsites separately, Table 2. The purpose of this part is to see if there is a significant difference between the slope of the developed simple linear equations between the height of the trees and different transformed forms of the diameter at breast height or not. These values reflect the rate of height increase for each unit of increase of diameter at breast height. Such data determine the suitability of the studied microsites for the planting of Pinus brutia Ten trees.

In the second part, all microsites were taken as one sample, and the parameters of 25 regressions were estimated in Table 3.

2.4 Criteria used for screening the developed equations

The criteria used to examine the performance of the regression models in the prediction of the response variable/(s) can be classified into two different types;

1. **The first type** is used, when the regression models under test have the same form of the dependent variable. The coefficient of determination is the most important criterion that can be used in such conditions.

a) Coefficient of determination (R^2)

Here, two types of coefficients of determination can be used (R^2) and adjusted coefficient of determination(R'^2). Both of them are used only if the dependent variable of the tested equations appeared in the same form. Furthermore the (R^2) is used when the tested equations have the same numbers of the independent variables, unlike the (R'^2) which can be used when the number of independent variables is not the same (Furnival, 1961; Neter et al, 1996; Studenmund and Johnson, 2006; Younis 2019; Salih 2019). The following formulas are used for calculating them:

$$\mathbf{R^2 = (1 - \frac{residual\ ss}{total\ ss}} \text{------------------------------ } \mathbf{(1)}$$

$$\mathbf{R'^2 = (1 - \frac{(n-1)(Residual\ SS)}{(n-p) Total\ SS})} \text{ ---------------------}\mathbf{(2)}$$

The value of (R^2/R'^2) ranges between (0 and 1). There is a direct relationship between the values of (R^2/R'^2) and the precision of the equation in predicting the value of the dependent variable.

b) Durbin Watson

This criterion was proposed by Durbin and Watson (1950 and 1951). The main function of this criteria is to see if there is autocorrelation between the independent variables. The value of this criterion ranges between (0 and 4).

The following formula can be used to calculate the value of Durbin Watson (DW):

DW value = 2(1-p)

As it can be seen that the value of DW depends on the value of p and as follow:

If p= -1 then DW = 2(1-(-1)) = 4 = negative autocorrelation.

If p = 0 then DW = 2(1- 0)) = 2 = no autocorrelation.

If p= 1 then DW = 2(1- 1) = 0 = positive autocorrelation.

As a rule of thumb, the value of DW which lies between 1.5 to 2.5 is acceptable.

Mathematically the acceptance region of DW can be expressed as follow: ($1.5 \leq DW \leq 2.5$)

2. **The second group** of criteria can be used even if the dependent variables appeared in different transformed forms. The researchers have proposed different criteria, among them are:

1) Ohtomo's unbiased test

As mentioned earlier, it is not possible to compare the precision of equations in the estimation of the dependent variable unless their dependent variable appeared in the same form ((Furnival, 1961; Neter et al, 1996; Studenmund 2006; Salih 2020. Ohtomo (1956) proposed a method to overcome such a problem. He proposed regressing the predicted values of the dependent variable with the actual (observed) values in a simple linear regression; $\hat{y} = k + my$. According to this equation, the most accurate regression models are those in which the estimated values of the dependent variable ($\hat{y}$) are close to the original values (y), and this is achieved when the values of k and m approach zero and one respectively. However, the value of R^2 is also very important to be taken into consideration. Instead of taking these three statistics separately Salih (2019) proposed a new index, which is a modification to Ohtomo's unbiased test as follows:

Proposed Index = $|k - 0| + |1 - m| + |1 - R^2|$ ----- (3)

The first term calculates the deviation of the k value from zero, while the second and third terms calculate the deviation of both m and R^2 from one. Based on this criterion the most accurate equation is the one having the lowest value of the mentioned index.

2) Mean absolute error (MAE)

This measure has been proposed and used by many researchers as a measure of precision for testing the performance ability of equations even if their response variable appears in a different form.

MAE= $\frac{\sum|y_i - \hat{y}_i|}{n}$ ------------------------------ (4)

The lowest value of this criteria the higher the precision of the regression model (Salih, 2021)

3) Bias%

This statistic is calculated as follows:

Bias% = $\frac{\sum(y_i - \hat{y}_i)^2}{\sum y_i} * 100$ -------------------- (5)

Based on this criterion the lowest value of this statistics reveals the more precision of the regression model.

4) Akaike Information Criterion (AIC)

This criterion deals with the trade-off between the goodness of fit of a model and its complexity. It offers a relative estimate of the information lost when a given model is used to represent the process that generates the data. The general form of this criteria is:

AIC= $n * \ln(\frac{RSS}{n}) + 2k$ ---------------------------(6)

A correction factor ($\frac{2k(k+1)}{(n-k-1)}$) has to be added to this formula if the ratio of if $\frac{n}{k} < 40$, so it will take the following form:

AIC= $n * \ln(\frac{RSS}{n}) + 2k + \frac{2k(k+1)}{(n-k-1)}$ --------------(7)

Where RSS = residual sum of squares, p = number of the independent variables, and, (K= p+2). The precision of an equation increases as the value of AIC decreases.

5) Furnival Index (FI)

This index can be calculated as follow:

$$FI = \sqrt{mse} * \frac{1}{Geomtric\ mean\ of\ Y^{-1}}$$

2.5 Validation of selected regression equation

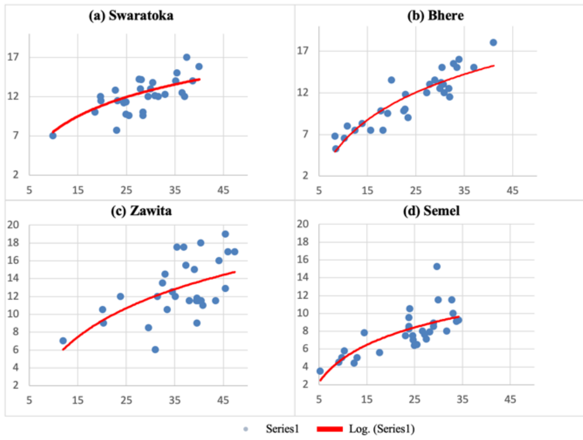
The collected data (120 pairs of diameters at breast height and total heights) were split into two parts, the first part consisted of 100 pairs of height and diameter, which were used for estimation of the parameters of 25 regression models, and the rest (20 pairs) of the mentioned variables for validation. Vanderschaaf (2008) used only 6 trees for validation of taper equations. It is desirable to partition the collected data into two parts one for estimation of the parameter and the other for validation of the developed model (Ajit, 2010). Geisser (1975) suggested that out of a single data set, a random sample (without replacement) of about 80% of data points should be used for model estimation and the remaining 20% of data points may be kept for validation of the selected model.

3. RESULTS AND DISSCUTION

3.1 generation of regression models

a) Generating of regression equations separately for each microsite

In the beginning, a scatter diagram was conducted for each microsite separately to detect the type of the model to be tested, Figure 2. The scatter diagrams show that there is a curvilinear relationship between the total height and the diameter at breast



height of the tree.

Figure 2. The scatter diagrams for the data set of total height and diameter at breast height, (a) Swaratoka, (b) Beher, (c) Semel, (d) Zawita.

After studying the scatter diagrams, the Stratigraphic package was used for estimating the parameters of 16 regression models (four regression models for each microsite) Table 2a and Table 2b).

Table 2a linear regression equations developed for Swaratoka and Behere.

Eq. no.	Swaratoka	R'^2	Behere	
1	H = 1.3 +0.419666*D	0.982	H = 0.418067*D	0.974
2	H = 2.10588*(D)^0.5	0.977	H = 1.96942*(D)^0.5	0.977
3	H = 3.28324*ln(D)	0.963	H = 2.96225*ln(D)	0.967
4	H = 0.0161454*(D)^2	0.912	H = 0.0127356*(D)^2	0.887

Table 2b linear regression equations developed for Zawita and Semel.

Eq. no.	Zawita	R'^2	Semel	
1	H = 0.31511*D	0.946	H = 0.266899*D	0.951
2	H = 1.93713*(D)^0.5	0.948	H = 1.27466*(D)^0.5	0.963
3	H = 3.25259*ln(D)	0.942	H = 1.93219*ln(D)	0.958
4	H = 0.0105389*(D)^2	0.912	H = 0.00784257*(D)^2	0.878

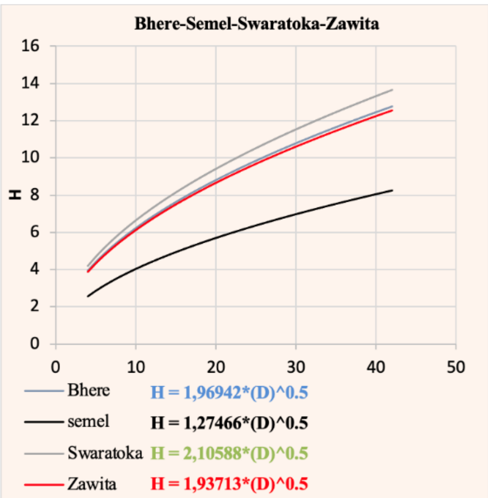
The tables above show that the $\sqrt{D}$ is the most important explanatory variable that explains the variation in the response variable, $\sum(H_i - \bar{H})^2$. The regression coefficient of this equation is highest in Swaratoka, followed by Behere, then by Zawita, and in last place came Semel. The same conclusion can be drawn by taking the first derivative of this form of a model for all microsites Table 3.

Table 3 the first derivative of the second regression model for all regions

Regions	Regression Model	First derivative of the model
Swaratoka	$H = 2.10588\sqrt{D}$	$H' = 1.0529D^{-0.5}$
Behere	$H = 1.9694\sqrt{D}$	$H' = 0.9847D^{-0.5}$
Zawita	$H = 1.9371\sqrt{D}$	$H' = 0.9686D^{-0.5}$
Semel	$H = 1.2767\sqrt{D}$	$H' = 0.6373D^{-0.5}$

It shows that the first derivative of Swaratoka has the highest coefficient compared with the others. This can be due to two reasons; 1) the first one may be due to the density, as we know that the height growth increases as the density increases, (unlike the diameter increment which decreases as the stand density increases) (Calama and Montero, 2004; Sharma and Parton, 2007; Schmidt et al, 2011). This means the selected trees in Swaratoka may be closer to each other than those selected from other microsites. 2) The second reason may hang with the suitability of Swaratoka for the growth of Calabrian pine over the other microsites. This can be due to the high rate of precipitation in Swaratoka 1000mm) which is much more than the other microsites,Table1.

The superiority of Swaratoka can be extracted from the graphs shown in Figure 3. The line representing Swaratoka is the highest one. The lines representing Behere and Zawita almost coincide with each other because of having almost the same slope (1.9694 and 1.9371). The slope of the equation belonging to Semel was the lowest.



b) Generating of regression equations for all microsites together

The whole range of datasets that were collected from all microsites was used for estimating the parameters of 25 regression equations Table 4

Table 4. The developed regression models for all four microsites.

Eq. no.	Equation	R^2
Group 1: equations with original form of H		
1	$H = 1.3 + 0.3408D$	93.73
2	$H = 1.3 + 1.9023\sqrt{D}$	93.93
3	$H = 1.3 + 3.0462\ln(D)$	92.91
4	$H = 1.3 + 148.01/D$	50.17
5	$H = 1.3 + 0.0096D^2$	86.15
Group 2 Equations with square root of H		
6	$\sqrt{H} = 1.3 + 0.0679D$	95.07
7	$\sqrt{H} = 1.3 + 0.3825\sqrt{D}$	96.79
8	$\sqrt{H} = 1.3 + 0.6145\ln(D)$	96.36
9	$\sqrt{H} = 1.3 + 30.91/D$	55.69
10	$\sqrt{H} = 1.3 + 0.00227D^2$	85.25
Group 3 Equations with logarithmic forms of H		
11	$\ln H = 1.3 + 0.0364D$	94.74
12	$\ln H = 1.3 + 0.2041\sqrt{D}$	95.43
13	$\ln H = 1.3 + 0.3271\ln(D)$	94.59
14	$\ln H = 1.3 + 15.79/D$	50.38
15	$\ln H = 1.3 + 0.0010D^2$	85.96
Group 4 Equations with H reciprocal		
16	$H^{-1} = 1.3 - 0.0388D$	91.20
17	$H^{-1} = 1.3 - 0.2239\sqrt{D}$	97.45
18	$H^{-1} = 1.3 - 0.3633\ln(D)$	98.90
19	$H^{-1} = 1.3 - 20.58/D$	72.50
20	$H^{-1} = 1.3 - 0.2239D^2$	97.45
Group 5 equations with H-square form		
21	$H^2 = 1.3 + 4.863D$	84.30
22	$H^2 = 1.3 + 26.61\sqrt{D}$	81.21
23	$H^2 = 1.3 + 42.27\ln(D)$	79.01
24	$H^2 = 1.3 + 1884.5/D$	35.87
25	$H^2 = 1.3 + 0.1408D^2$	82.48

c) Selection procedure

1) For equations with the same form of the dependent variable

It can be seen that 25 regression equations were developed and out of which (5) equations have the original form of the dependent variable (H). The dependent variable in equations number (6 to 10) appeared in root square form, while equations (11 to 15) have the logarithmic. The equations (16 to 20) have a reciprocal form of (H). The last group appeared in the square form of the dependent variable (H^2). Many criteria have been used by researchers for testing the efficiency of the developed equations in predicting the dependent variables. The most important criterion is the coefficient of determination (R^2). However this criterion can't be used unless the regression equations have the same form of the dependent variable (Ohtomo, 1956; Furnival, 1961; Amaro et al, 1998; Amin, 2016; Younis, 2019; Salih et al, 2019; Salih, 2021). Therefore, the precision of equations belonging to the same group can be tested using R'^2/R^2.

The precision of a regression equation is directly proportional to Ŕ², therefore the equation number (2), (7), (12), (18), and (21) were selected from group 1, group 2, group 3, group 4 and group 5 respectively because of having the highest values of R'^2/R^2. These equations underwent further analysis in order to select the best one Table 5.

Table 5. The selected regression equations after the first stage of screening using R^2.

Group and eq.	Equation	R^2
G (1,2)	$H = 1.3 + 1.9023\sqrt{D}$	93.93
G (2,7)	$\sqrt{H} = 1.3 + 0.3825\sqrt{D}$	96.79
G (3,12)	$\ln H = 1.3 + 0.2041\sqrt{D}$	95.43
G (4,18)	$H^{-1} = 1.3 - 0.3633\ln(D)$	98.90
G (5,21)	$H^2 = 1.3 + 4.863D$	84.30

2) Testing of performance ability of heterogeneous models

Before conducting such tests, it is worth making mathematical and biological analyses of the developed equations to determine their limitations. One of the most important limitations in quantitative variables is having a negative estimation. The only equation listed in Table 4 has such type of limitations because of obtaining a negative term, which is equation G (4.18). Keeping the left side of the mentioned equation positive the term $0.3633\ln(D)$ should be less than 1.3 ($0.3633\ln(D) \leq 1.3$). Solving this inequality leads to D ≤ 35.8. This entails that this regression equation does not apply for large trees, an equation with such characters is not accepted, and therefore is eliminated from the competition list. The rest of the equations were subjected to the following tests of criteria:

a) Ohtomo's unbiased test

The regression models listed in Table 5 were subjected to the proposed modified test of Ohtomo that was proposed by Salih (2021) Table 6.

Table 6: shows the proposed index of Salih's unbiased test for the selected equation.

Sub Gr and eq.	Equation	R^2	$\lvert k-0\rvert + \lvert 1-m\rvert + \lvert 1-R^2\rvert$
G (1,2)	$\hat{H} = 7.08 + 0.37H$	0.47	8.22
G (2,7)	$\hat{H} = 5.73 + 0.47H$	0.51	6.77
G (3,12)	$\hat{H} = 6.35 + 0.40H$	0.51	7.46
G (5,21)	$\hat{H} = 6.78 + 0.43H$	0.51	7.86

Although it can be concluded that the equation G (2,7), seems to be superior to the rest of the equations in the competition list, because of having the lowest value of this criterion, all equations listed in Table 6 were subjected to other tests, including bias%, (MAE), Furnival index and AIC Table 7. The formula for calculating these criteria is already given under the topic of Material and Methods.

b) Bias%

Based on this criterion, the precision of the equations in the competition list is very close to each other, therefore they were subjected to another criterion called Furnival Index, which can be used instead of SEE when the dependent variable appeared in a different transform form.

c) Mean absolute error (MAE)

The precision of an equation in the prediction of the response variable is inversely proportional to the value of this criterion.

d) Akaike Information Criterion

Based on this criterion the most precise model is the one having the lowest value Table 7.

Table 7: Calculation of bias, MAE, and AIC statistics for the competed models

Group and eq.	Equation	$\sum(y_{ij} - \hat{y})^2$	$\sum y_{ij}$	Bias %	MAE	AIC
G (1,2)	$H = 1.3 + 1.9023\sqrt{D}$	743.84	1328.27	56	2.039	6.86
G (2,7)	$\sqrt{H} = 1.3 + 0.3825\sqrt{D}$	736.9	1328.27	55	1.971	6.83
G (3,12)	$\ln H = 1.3 + 0.2041\sqrt{D}$	771.08	1328.27	58	2.032	6.86
G (5,21)	$H^2 = 1.3 + 4.863D$	773.54	1328.27	58	2.024	6.86

The performance of the competed regression equations in the prediction of the dependent variable is very close to each other based on biased %, MAE, and AIC, therefore they were subjected to Furnival Index criterion, Table 8.

Table 8: shows the Furnival Index test for the selected regression models.

Group and eq.	Equation	FI
G (1,2)	$H = 1.3 + 1.9023\sqrt{D}$	1.60
G (2,7)	$\sqrt{H} = 1.3 + 0.3825\sqrt{D}$	0.1127
G (3,12)	$\ln H = 1.3 + 0.2041\sqrt{D}$	0.2089
G (5,21)	$H^2 = 1.3 + 4.863D$	1.025

It seems that after applying all mentioned criteria, one can certainly confirm that the regression equation $\sqrt{H} = 1.3 + 0.3825\sqrt{D}$ is the most appropriate model that describes the relationship between the height and breast height diameter of Calabrian pine grown in the four mentioned microsites

3.2 Test of independence of residuals

The selected regression model must have a consistent precision for the whole range of data. This can be tested by plotting the residuals ($Y_i - \hat{Y}_i$) against the independent variable. The plotted points should be normally and independently distributed with the mean of zero and a standard deviation of σ, Figure 4. This statement can be expressed as:

$$\textbf{Error } (\varepsilon) \textbf{ or residuals} \sim \textbf{NID}(0, \sigma).$$

The figure shows that there is no special trend for the plotted points, even if the variance is more for a large tree. This means that the selected equation has a consistent accuracy for the whole range of data.

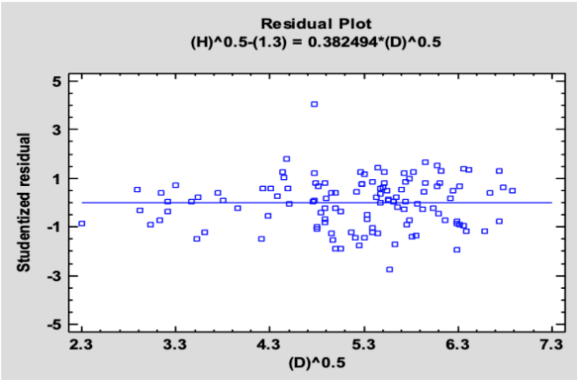


Figure. 4 plotting of the residuals $(\hat{H} - H)$ against the values of $\sqrt{D}$

The plotted points of the above-mentioned figure show that they have no special trend but they are scattered which entails that the selected regression equation has a consistent precision for the whole range of data

3.3 Validation of the selected regression equation

The selected regression equation has undergone a test of validation. For conducting such a test, a sample of 20 trees was selected randomly from the collected data (five trees per each microsite) to see how well the selected equation is suited to independent data (the data that were not used in the calibration of regression models). The resulting equation was $\sqrt{H}$ =1.3+0.4211 $\sqrt{D}$ with an R^2 of 0.892. The previously selected regression equation was $\sqrt{H} = 1.3 + 0.3825\sqrt{D}$ with an R^2 of 0.968. This means that the precision of the developed equation is well suited for the independent dataset.

4. CONCLUSION

It can be seen that the height/ diameter ratio was different in different locations depending on the microsite of the location. The highest ratio was found in Swaratoka followed by Behere, then followed by Zawita and Semel comes in the last place. This entails that either the stand density of trees in Swaratoka is more than in other locations or Calabrian pine trees are better adapted to Swaratoka compared to other studied locations. The most precise regression models were gained when no Y-intercept regression was used. It was forced to introduce a Y-intercept of 1.3 in the developed regression equations by making a modification in the expressions of the dependent variable, such as using Log (H) – (1.3) instead of Log (H). It can be concluded that there is a curvilinear relationship between tree height and its diameter at breast height, and this result agreed with what was found by (Philip 1994; Huang et al, 2000; Huang et al, 2009; El mamoun et al, 2013; and Chai et al, 2018).

Height – diameter at breast height table

As it is well known that the main purpose of developing regression equations is to select the most appropriate equation that can be used for the prediction of the dependent variable (H) corresponding to different values of the independent variable (D), Table 9.

Table 9: shows the estimated values of the height of Calabrian pine trees in Zawita, Behere, Swaratoka, and Semel against the breast height diameter.

DBH (cm)	Expected Height(m)	DBH (cm)	Expected Height(m)	DBH (cm)	Expected Height(m)
5	4.65	17	8.28	29	11.29
6	5.00	18	8.54	30	11.53
7	5.35	19	8.80	31	11.76
8	5.67	20	9.06	32	12.00
9	5.99	21	9.32	33	12.23
10	6.30	22	9.57	34	12.46
11	6.60	23	9.82	35	12.69
12	6.89	24	10.07	36	12.92
13	7.18	25	10.32	37	13.15
14	7.46	26	10.56	38	13.38
15	7.74	27	10.81	39	13.61
16	8.01	28	11.05	40	13.83

Estimated from regression equation:

$$\sqrt{\hat{H}} = 1.3 + 0.3825\sqrt{D}$$

R^2 =0.9679 MAE= 1.971 Bias=55% AIC= 6.83
Furnival Index = 0.1127
Modified Ohtomos Index (proposed by Salih) = 6.77
Date July 2022

REFERENCES

1. 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*(1-2), 88-98. https://www.sciencedirect.com/science/article/pii/S0378112708003253
2. Ajit, S. (2010). Estimation and validation methods in tree volume and biomass modelling: statistical concept. *National Research Centre for Agroforestry, Jhansi, India 18p*.
3. Amaro, A., Reed, D., Tomé, M., & Themido, I. (1998). Modeling dominant height growth: Eucalyptus plantations in Portugal. *Forest Science*, *44*(1), 37-46. https://doi.org/10.1093/forestscience/44.1.37
4. Amin, H.M. (2016). quantification of diementional properties of Quercus infectoria. Oliv grown naturally in Chamanke area. Duhok. MSc Thesis. College of Agriculture, Duhok University.
5. Avery, T. E., & Burkhart, H. E. (2015). *Forest measurements*. Waveland Press. https://books.google.iq/books?id=IWx1CQAAQBAJ
6. Calama, R., & Montero, G. (2004). Interregional nonlinear height diameter model with random coefficients for stone pine in Spain. *Canadian Journal of Forest Research*, *34*(1), 150-163. https://cdnsciencepub.com/doi/abs/10.1139/x03-199
7. Carron, L. T. (1968). *An outline of forest mensuration with special reference to Australia*. Australian National University Press. https://openresearch-repository.anu.edu.au/handle/1885/114864
8. Chai, Z., Tan, W., Li, Y., Yan, L., Yuan, H., & Li, Z. (2018). Generalized nonlinear height–diameter models for a Cryptomeria fortunei plantation in the Pingba region of Guizhou Province, China. *Web Ecology*, *18*(1), 29-35. https://doi.org/10.5194/we-18-29-2018
9. Dyer, M. E., & Bailey, R. L. (1987). A test of six methods for estimating true heights from stem analysis data. *Forest Science*, *33*(1), 3-13. https://doi.org/10.1093/forestscience/33.1.3
10. Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, *37*(3/4), 409-428. https://www.jstor.org/stable/2332391
11. Durbin, J. and Watson, G.S. (1951). Testing for serial correlation in least squares regression. II. Biometrika, 38(1/2), pp.159-177.

12. El Mamoun, H. O., El Zein, A. I., & El Mugira, M. I. (2013). Modelling height-diameter relationships of selected economically important natural forests species. *Journal of forest products & industries*, *2*, 34-42. https://www.researchgate.net/publication/269095707_Modelling_Height-Diameter_Relationships_of_Selected_Economically_Important_Natural_Forests_Species
13. Furnival, G. M. (1961). An index for comparing equations used in constructing volume tables. *Forest Science*, *7*(4), 337-341. https://doi.org/10.1093/forestscience/7.4.337
14. Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American statistical Association*, *70*(350), 320-328. https://www.tandfonline.com/doi/abs/10.1080/01621459.1975.10479865
15. Huang, S., Price, D., & Titus, S. J. (2000). Development of ecoregion-based height–diameter models for white spruce in boreal forests. *Forest ecology and management*, *129*(1-3), 125-141. https://doi.org/10.1016/S0378-1127(99)00151-6
16. Huang, S., Meng, S. X., & Yang, Y. (2009). Using nonlinear mixed model technique to determine the optimal tree height prediction model for black spruce. *Modern Applied Science*, *3*(4), 3-18. https://doi.org/10.5539/mas.v3n4p3
17. Kariuki, M. (2002). Height estimation in complete stem analysis using annual radial growth measurements. *Forestry*, *75*(1), 63-74. https://doi.org/10.1093/forestry/75.1.63
18. Kershaw Jr, J. A., Ducey, M. J., Beers, T. W., & Husch, B. (2016). *Forest mensuration*. John Wiley & Sons. https://books.google.iq/books?id=SGVJDQAAQBAJ
19. Lappi, J. (1997). A longitudinal analysis of height/diameter curves. *Forest science*, *43*(4), 555-570.https://doi.org/10.1093/forestscience/43.4.555
20. Larsen, D. R., & Hann, D. W. (1987). Height-diameter equations for seventeen tree species in southwest Oregon. http://hdl.handle.net/1957/8245
21. Mensah, S., Pienaar, O. L., Kunneke, A., du Toit, B., Seydack, A., Uhl, E., ... & Seifert, T. (2018). Height–Diameter allometry in South Africa's indigenous high forests: Assessing generic models performance and function forms. *Forest Ecology and Management*, *410*, 1-11. https://doi.org/10.1016/j.foreco.2017.12.030
22. Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*.
23. Ng'andwe, P., Chungu, D., & Tailoka, F. (2021). Stand characteristics and climate modulate height to diameter relationship in Pinus merkusii and P. michoacana in Zambia. *Agricultural and Forest Meteorology*, *307*, 108510. https://doi.org/10.1016/j.agrformet.2021.108510 .
24. Ohtomo, E. (1956). A Study on Preparation of volume Table. (I). *Journal Of the Japanese Forestry Society*, *38*(5), 165-177. https://www.jstage.jst.go.jp/article/jjfs1953/38/5/38_5_165/_article/-char/en
25. Philip, M.S., 1994. Measuring trees and forests. CAB international. https://www.cabdirect.org/cabdirect/abstract/19936791705
26. Pretzsch, H. (2009). Forest dynamics, growth, and yield. *Forest dynamics, growth and yield*, 1-39. https://doi.org/10.1007/978-3-540-88307-4_1
27. Salih, T.K., Younis, M.S. and Wali, S.T. (2019). Dendroclimatological Analysis of Pinus brutia Ten. Grown in Swaratoka, Kurdistan Region—Iraq. In *Recent Researches in Earth and Environmental Sciences* (pp. 9-19). Springer, Cham.
28. Salih, T.K., Younis, M.S. and Wali, S.T. (2021) Allometric Regression Equations Between Diameter Growth and Age of Valonia Oak Trees Grown in Duhok Province, Iraq. International Hasankefy and Innovation Congress 06 – 07 november 2021 Batman.
29. Schmidt, M., Kiviste, A., & von Gadow, K. (2011). A spatially explicit height–diameter model for Scots pine in Estonia. *European Journal of Forest Research*, *130*(2), 303-315. https://link.springer.com/article/10.1007/s10342-010-0434-8
30. Shahbaz, S. E. (2010). Trees and Shrubs, A field guide to the trees and shrubs of Kurdistan region of Iraq. *Journal of university of Duhok*.
31. 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*(3), 187-198. https://doi.org/10.1016/j.foreco.2007.05.006
32. Studenmund, A. H., & Johnson, K. B. (2006). Using econometrics: A practical guide 5th edition.
33. Sumida, A., Miyaura, T., & Torii, H. (2013). Relationships of tree height and diameter at breast height revisited: analyses of stem growth using 20-year data of an even-aged Chamaecyparis obtusa stand. *Tree physiology*, *33*(1), 106-118. https://doi.org/10.1093/treephys/tps127 .
34. VanderSchaaf, C. (2008). Compatible stem taper and total tree volume equations for loblolly pine plantations in southeastern Arkansas. *Journal of the Arkansas Academy of Science*, *62*(1), 103-106. https://scholarworks.uark.edu/jaas/vol62/iss1/16/
35. Younis A.J. 2019.Stand volume equations for *Quercus aegilops* L. and *Quercus infectoria* Oliv. In Duhok governorate. MSc thesis
36. Zeide, B., & Vanderschaaf, C. (2002, March). The effect of density on the height-diameter relationship. In *Proceedings of the 11th Biennial Southern Silvicultural Research Conference* (pp. 463-466). General Technical Report SRS-48. USDA, Forest Service. Southern Research Station, Asheville, NC. https://www.fs.usda.gov/treesearch/pubs/3103