Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Quantifying the Impact of Climate Change on Pinus Hartwegii Lindl Forests: A Novel Approach Using AI-Powered Allometric Models

Moisés Arreguín Sámano[1], José Abelardo Paucar Camacho[2], Eduardo Patricio Salazar Castañeda[3], María Transito Vallejo Illijama[4], Miguel Ángel Guallpa Calva[5], Ángel Leyva Ovalle[6]

Abstract

In Mexico, it is essential to generate sufficiently simple biometric models with reliable projections of the current and future condition of the forest. The objective of this study was to build a dynamic stochastic model using AI tools that estimates TTV allometric equations to evaluate carbon capture and describe the temporal evolutionary displacement in forest Pinus hartwegii Lindl in the states of Mexico and Puebla facing climate change scenarios. Numerical bases were used INFyS, NASA Power data, OLS mathematical models, random forest bookstore, Ridge Model with its regression, R algorithms, residual validation graphs, residuals vs. fitted, normal Q-Q, scale location model, residuals vs. Laverage, Newton's volumetric estimation equations, excurrent dendrometrics (theoretical or logs) and traditional allometric equations by Federal Entity. The allometric equation estimated by mathematical models with the best estimate of the TTV cc is theoretical models for excurrent dendrometric types in the states of Mexico with an evaluation of 90.5% and Puebla with 95.0%, respectively. The allometric equations of "commercial volumetric dimension" were estimated for Pinus hartwegii Lindl with significant climatic variables. There is a better volumetric approach to climate change scenarios using Newton's mathematical equations and theoretical models for excurrent dendrometric types.

Keywords: *Biocene Ensemble, Abiocene Ensemble, CO2, Allometric Equation, Climate Change, Dynamic Stochastic Model.*

Introduction

"Climate change is one of the most pressing crises to humanity. The rise of greenhouse gases (GEI) in Earth's atmosphere is causing severe anomalies in our planet's climate system due to increasing temperatures around the globe (IPCC, 2021; NASA, 2022). It is crucial to curb the anthropogenic emissions of GEI and to reduce the carbon concentration in the atmosphere (UNFCCC, 2020). Carbon bio-sequestration is accepted by scientists, practitioners, and politicians as a promising way to reduce atmospheric GEI concentrations and therefore mitigate climate change (Smith et al., 2020; Lal, 2019). In this context, it is important to learn about the capacity that different plant species have to sequester carbon and store it in their biomass (Pan et al., 2018). Biometric models have become a tool to estimate the amount of carbon that trees and other plants can store (Brown, 2021)."

In Mexico there is still a need to generate simple biometric models with reliable projections of the current and future condition of the forest[1]. Therefore, developing the theoretical-practical mathematical modeling capacity based on the dialectical relationship between the logics of mathematics and engineering responds to the interest of, according to [2], resolve situations specific to the profession; for example, a stochastic mathematical model defined as a mathematical abstraction of an empirical process governed by probabilistic laws and as a quantitative model that calculates conditional results of the decision alternatives

[1] Ingeniería Forestal, División de Ciencias Forestales, Universidad Autónoma Chapingo, Chapingo estado de México, Email: m_arreguin230@yahoo.com.mx
[2] Ingeniería en Gestión del Riesgo, Facultad de Ciencias de la Salud y del Ser Humano, Universidad Estatal de Bolívar, Provincia Bolívar, Ecuador, Email: apaucar@ueb.edu.ec.
[3] Ingeniería Forestal, Facultad de Recursos Naturales, Escuela Superior Politécnica de Chimborazo, Provincia Chimborazo, Ecuador, Email: eduardo.salazar@espoch.edu.ec.
[4] Ingeniería en Gestión del Riesgo, Facultad de Ciencias de la Salud y del Ser Humano, Universidad Estatal de Bolívar, Provincia Bolívar, Ecuador, Email: mvallejo@ueb.edu.ec.
[5] Ingeniería Forestal, Facultad de Recursos Naturales, Escuela Superior Politécnica de Chimborazo, Provincia Chimborazo, Ecuador, Email: miguel.guallpa@espoch.edu.ec
[6] Ingeniería Forestal, División de Ciencias Forestales, Universidad Autónoma Chapingo, Chapingo estado de México; Tel.: (+52 55 6629 1056), Email: aleyvao@chapingo.mx

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

in each state [3]. An example is the importance of forest masses as CO_2 reservoirs based on their sensitivity to climate change at a global level [4].

Each species or plant population requires specific conditions of temperature, humidity and light for its physiological processes. If it exceeds its degree of tolerance, it will be reflected in the alteration of its distribution, being more severe in individuals from mountainous ecosystems. Even though there is extensive research with climate change criteria, xylotechnological and molecular variation, quantification of organic carbon stored in soil by ha^{-1}, physical-chemical properties of wood, dendro-epidometric measurements, development, validation, simulation with mathematical modeling and that the theory of Stochastic Processes includes random vectors of infinite dimension or arbitrary infinite collections of arbitrary variables with dynamic model analysis, that collects its evolution over time [5], there are included in few forestry investigations and are currently scarce in Mexico.

The objective of this research was to build a dynamic stochastic model using AI tools that estimates TTV (TOTAL TREE VOLUME) allometric equations to evaluate carbon capture and describes the temporal evolutionary displacement in forest ***Pinus hartwegii*** Lindl in the states of Mexico and Puebla facing climate change scenarios. The allometric equation estimated by mathematical models with the best estimate of the TTV cc are the theoretical models for excurrent dendrometric types in the states of Mexico with 90.5% and Puebla with 95.0%. The allometric equations of "commercial volumetric dimension" were estimated for ***Pinus hartwegii*** Lindl with significant climatic variables. There is a better volumetric approach to climate change scenarios with Newton's mathematical equations and theoretical models for excurrent dendrometric types.

Materials and Methods

The Comisión Nacional Forestal implemented the second cycle of the forest inventory INFyS in 2009, concluding in 2013. The third cycle of remeasurements was in the period 2015-2020 [6]. The sampling design was systematically stratified by clusters, they were located every 5x5 km in forests and the sampling units were located in all the forest ecosystems of the country, including forests. The cluster or primary sampling unit (UMP) is a circular plot of one hectare (56.42 m radius) in which four secondary sampling units (MSU) or sites are evaluated, arranged in an inverted "Y" with respect to the north and are circular. in the case of forests.

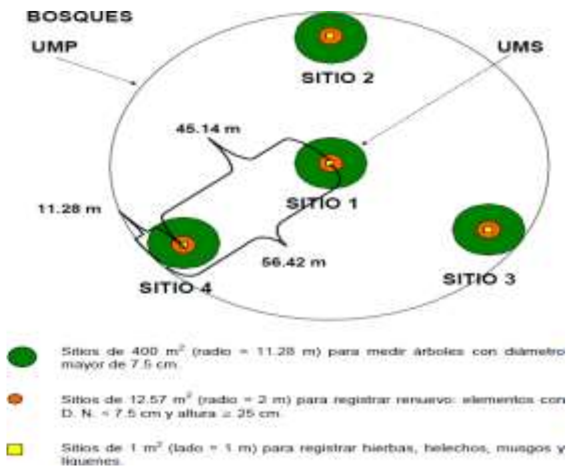

Figure 1. Primary Sampling Units (UMP) And Secondary Sampling Units (UMS) For the National Forestry and Soil Inventory in Forests and Vegetation in Arid Areas.

The registration of trees with a normal diameter greater than 7.5 cm is done in circular UMS of 400 m^2 (radius 11.28 m) and to record the reforestation within each UMS of 400 m^2 a concentric circle of 12.56 m^2 (2 m de radio) is drawn in the case of forests. The main variables that are recorded in the tree elements of forests are scientific name, normal diameter, crown and total height, presence/absence of damage (biotic

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

and abiotic); In the case of conifers, the age of the three trees closest to the center of the 400 m^2 UMS is recorded [7,8].

According to the methodology proposed by Leyva-Ovalle [9], the INFyS database was cleaned to avoid redundancy and inconsistency in the dasometric data. SQL statements were used to select only data useful for information processing and thus only consider live trees, normal diameter ≥ 7.5 cm and total height greater than 0.

In order to carry out the total volumetric estimation using Newton's adjusted allometric equations, excurrent dendrometrics (theoretical or logs) and Federal Entities, using mathematical models of Ordinary Least Squares, random Forest package and R algorithms, by ***Pinus hartwegii*** Lindl in the states of México and Puebla considering the climate series for the period 2022-2023, the volumetric estimates of the Hubert equations were compared, LOG((3.1416/4)*((normal_diameter)^2)*total_height(m)), Smalian, LOG(((((3.1416*((smaller_log_diameter_dn/2)^2))+(3.1416*((normal_diameter/2)^2))+(3.1416*((0/2)^2)))/2)*total_height(m)), Newton, LOG((1/6)*((3.1416*((Log_smaller_diameter_dn/2)^2))+(3.1416*((Log_larger_diameter/2)^2))+(4*(3.1416*((normal_diameter/2)^2))))).

The theoretical model was included by adding volumetric calculations of the geometric figures of a tree, Neloid ((3.1416/16)*((normal_diameter)^2))*Length_log_dn), Cylinder(((3.1416/4)*(normal_diameter)^2)*Log_length_dn)), Paraboloid ((3.1416/8)*(normal_diameter)^2)*Log_length), Cone ((3.1416/12)*(normal_diameter)^2)*Log_length) and the allometric equation traditionally used by the government institutions of each state, such as Protectora de Bosques (PROBOSQUE) in the state of Mexico ((2.7183^(-10.024))*(normal_diameter ^(2.0632))*(total_height^(0.8640))) and Government of the State of Puebla (-0.0292)+((0.0006)*(normal_diameter^2))+((0.00001)*(normal_diameter^2)*(total_height(m)).

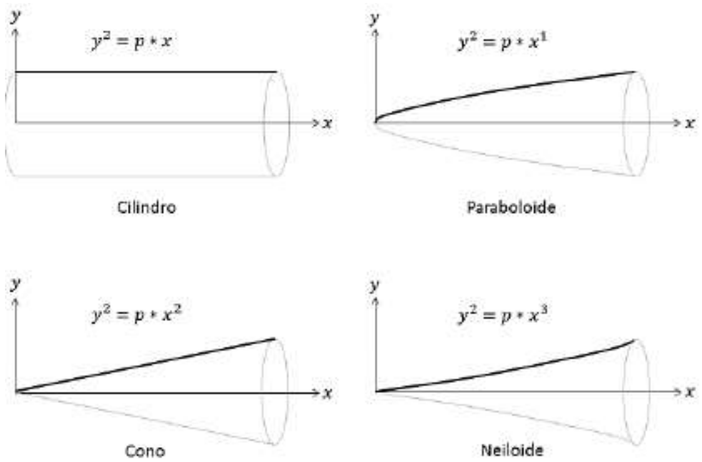

Figure 2. The Figures Show the Profile of The Following Basic Geometric Solids: (A) Cylinder; (B) Paraboloid; (C) Cone; (D) Neloid.

Thus, generally, the lower part of the tree conforms to a *neiloid*, the lower middle section to a *cylinder* or a *paraboloid*, the upper middle section to a paraboloid and the upper end to a *cone* or a paraboloid. However, the turning points between these forms do not present a defined pattern.

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

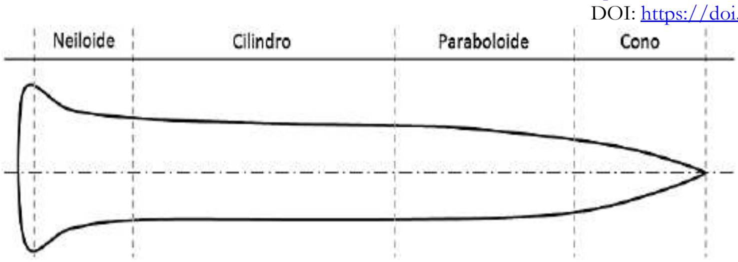

Figure 3. Decomposition of a Tree into Logs by The Following Basic Geometric Solids: (A) Neloid; (B) Cylinder; (C) Paraboloid; (D) Cone.

Different types of volume can be identified in an individual tree: the total tree volume (TTV), which considers all components (including branches) and the total tree roll volume (vrta).

The Newton-adjusted allometric equations, excurrent dendrometric equations (theoretical or logs) and Federative Entities were mathematically validated by residual graph, residual vs. graph. Fitted, normal Q-Q plot, Scale Location Model plot and Residuals vs. Laverage. In order to carry out the total volumetric estimation in the face of climate change using allometric equations of ***Pinus hartwegii*** Lindl In the states of Mexico and Puebla considering the climate series for the period 2022-2023, allometric equations were estimated based on volumetric increase (endogenous variable), Ridge Model [10] with its regression, Ordinary Least Squares, R algorithms [11], POWER climate databases [12,13] of 60 NASA variables [14] (POWER CERES/MERRA2).

With the objective of correcting the model and its estimation, two common methods were used to eliminate heteroskedasticity: Transformation of variables, the most used transformations being Logarithmic $\left(Y' = ln(Y)\right)$, Quadratic $(Y' = Y^2)$ and inverse $\left(Y' = \left(\frac{1}{Y}\right)\right)$ and estimation using weighted least squares (WLS). The residual variance could be estimated in $V\left(u \middle| X_1, X_2, \dots, X_p\right) \quad = \quad f\left(X_1, X_2, \dots, X_p\right)$ as:

$$V\left(u_j \middle| X_1, X_2, \dots, X_p\right) \quad = \quad \sigma^2 h_j, 1 \leq j \leq n \quad (1)$$

where h_j is a function of exogenous and endogenous variables that can be calculated for each observation in the sample. Assuming that the model meets all the assumptions of the classical linear model, except for homoscedasticity. In certain cases, it is possible to relax the assumption of normality of the residuals, thus allowing hypothesis testing to be carried out. The only requirement is that the sample be large. To check the normality of the model residuals, the simplest way was used: normal probability graphs. For example:

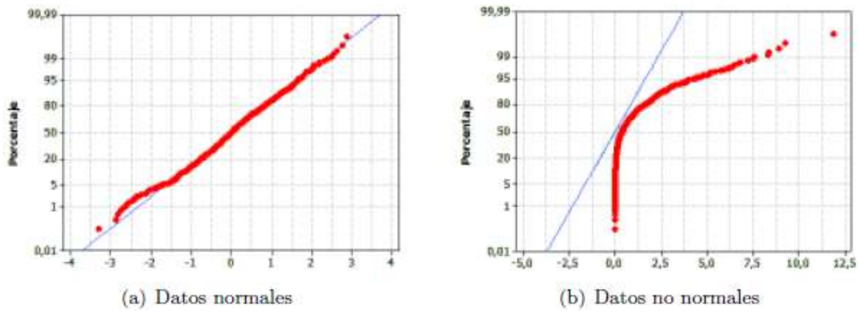

Figure 4. Normal Probability Plot for Normal and Non-Normal Data.

The allometric equations estimated based on the volumetric increase against climate change scenarios were evaluated using the R^2 criterion. To select the best model, three criteria were considered: numerical analysis, residual graphs and three statistics. The latter were the estimated Adjusted Coefficient of Determination for non-linear regression (R^2), the Root Mean Square of the Error (REMC) and the average bias (e).

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

$$R^2 = r^2_{Y_i \hat{Y}_i} \quad (2)$$

$$REMC = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n - p}} \quad (3)$$

$$e = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n} \quad (4)$$

Where: $r_{Y_i \hat{Y}_i}$ = It is the correlation coefficient between the observed value (yi) and the predicted value (yi) of the dependent variable, y_i = observed value of the dependent variable, $\hat{y}_i$ = value predicted by the model, n = number of data used in model fitting, p = number of model parameters.

According to [15], the Ridge estimation procedure provided a way to address the problem of multicollinearity. The model matrix is altered to avoid the implications derived from its poor conditioning. [16] proposed to optimize the following expression, where their solution is in effect the Ridge estimator:

$$f(\beta_*) \quad = \quad \|y - X\beta\|_2^2 + \lambda\|\beta\|_2^2 \quad (5)$$

Note that on the left side of the sum there is the expression $f(\beta_*) \quad = \quad \|y - X\beta_*\|_2^2 = (y - X\beta_*)^T(y - X\beta_*)$ which is optimize to obtain the least squares estimator $\hat{\beta}$. Now to this same expression the term $\lambda\|\beta\|_2^2$ is added, which penalizes the squared norm of the parameter vector through the hyperparameter $\lambda < 0$ which will be responsible for generating the regularization. It should be noted that if $\lambda = 0$; then, the penalty has no effect and it would return to $f(\beta_*) \quad = \quad \|y - X\beta_*\|_2^2 = (y - X\beta_*)^T(y - X\beta_*)$. However, as $\lambda \to \infty$, the impact of the penalty increases and the estimates of the Ridge regression coefficient will approach zero.

Since some forest sites have differences that may be due to soil (water availability, etc.), topography (altitude, slope, exposure, etc.), or climate (temperature, precipitation), these can influence the composition of species and their growth patterns[17]; so the, INFyS numerical bases were used because they contain topographic information, physiographic factors, such as slope, exposure and altitude, which have been used as variables to measure the productivity of georeferenced trees [18].

Based on the fact that the NASA POWER project covered the ESS component, the agroclimatology data set with agricultural applicability was added; MERRA-2 data covers the period from 1981 to several real-time months: NASA/POWER CERES/MERRA2 60-variable POWER climate databases were used, such as TS, T2M, CLOUD_AMT, TOA_SW_DWN, PS, QV2M, RH2M, GWETROOT, etc.

Random Forest was used because it is an ensemble-based Machine Learning algorithm that combines multiple decision trees to improve accuracy and reduce variance in predictions (Breiman, 2001) cited by [19], presents several advantages compared to other supervised learning algorithms (Liaw & Wiener, 2002; Cutler et al., 2007) cited by [19], one of its main advantages is its ability to improve accuracy and offers better generalization on non-data. seen.

Results

Allometric Equations Estimated by OLS With R Algorithm In 2023 For ***Pinus Hartwegii*** Lindl.

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Estado de México

Table 1. Summary Of Allometric Equations Estimated by OLS With R Algorithm In 2023 For ***Pinus Hartwegii*** Lindl in Estado De México (See Appendix A for Allometric Equations Estimated by Region).

Timber volume classification by OLS criteria			**Statistical criterion**	**Statistical result ANOVA by method**		**Allometric equation estimated by method**		
						OLS	**SiBiFor**	
State	**Volumetric method**	**Volumetric dimension**		**OLS**	**SiBiFor (R^2)**		**UMAFOR**	**Equation**
México	Newton	Commercial	Multiple R-squared	0.9827	0.9618	log(Vol) = −0.12 + 0 log_h + 0.33 log_d	1508	0.00005 *Potencia(Diam,1.97258) *Potencia(Alt,0.92797)
			Adjusted R-squared	0.9826				
			F-statistic	8037				
			p-value	< 2.2e-16				
	Theoretical (wood logs)	Commercial	Multiple R-squared	0.9748		log(Vol) = 0.11 + 0.12 log_h + 0.25 log_d		
			Adjusted R-squared	0.9746				
			F-statistic	5463	0.9585		1510	0.00004 *Potencia(Diam,2.15224) *Potencia(Alt,0.78917)
			p-value	< 2.2e-16				
	Federal entity	Commercial	Multiple R-squared	0.9971		log(Vol) = −9.94 + 0.85 log_h + 2.05 log_d		
			Adjusted R-squared	0.997				
			F-statistic	4.373e+04				
			p-value	< 2.2e-16				

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Puebla

Table 2. Summary Of Allometric Equations Estimated by OLS With R Algorithm In 2023 For ***Pinus Hartwegii*** Lindl in Estado De Puebla (See Appendix A For Allometric Equations Estimated by Region).

Timber volume classification by OLS criteria			**Statistical criterion**	**Statistical result ANOVA by method**		**Allometric equation estimated by method**		
						OLS	**SiBiFor**	
State	**Volumetric method**	**Volumetric dimension**		**OLS**	**SiBiFor (R^2)**		**UMAFOR**	**Equation**
Puebla	Newton	Commercial	Multiple R-squared	0.9627	0.9721	$\log(Vol) = -0.07 + 0.05 \log\_h + 0.27 \log\_di$		0.00006 *Potencia(Diam,1.99376) *Potencia(Alt,0.90054)
			Adjusted R-squared	0.9624				
			F-statistic	3831				
			p-value	< 2.2e-16	0.9772		2105	0.00004 *Potencia(Diam,2.20234) *Potencia(Alt,0.66536)
	Theoretical (wood logs)	Commercial	Multiple R-squared	0.9642		$\log(Vol) = 0.12 + 0.15 \log\_h + 0.22 \log\_di$		
			Adjusted R-squared	0.964				
			F-statistic	4000	0.9817		2108	0.00011 *Potencia(Diam,1.75875) *Potencia(Alt,0.95559)
			p-value	< 2.2e-16				
	Federal entity	Commercial	Multiple R-squared	0.9709		$\log(Vol) = -8.18 + 0.61 \log\_h + 1.79 \log\_di$		
			Adjusted R-squared	0.9707	0.9939		-	EXP(-9.63495649+1.86670523 *LN(diam)+0.99551381 *LN(alt))
			F-statistic	4889				
			p-value	< 2.2e-16				

Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

The most used criteria to select the best model are based on the numerical analysis of the estimated Adjusted Coefficient of Determination for non-linear regression (R^2 the Root Mean Square of the Error (REMC), the average bias (e) and graph of waste. Based on the use of Newton's mathematical equation to estimate timber volumes, theoretical models for excurrent dendrometric types (volume per log), allometric equations used in the field by government institutions of the Federal Entities, MCO and R algorithms, the equations were estimated. allometric "commercial volumetric dimension" of ***Pinus hartwegii*** Lindl using mathematical models in the states of Mexico and Puebla. According to the validation plots, the residuals appear random, they are randomly dispersed around the horizontal line near zero, there is no fan, cone, U-shape or curve pattern, they have a constant dispersion along the independent variable or fitted values, are centered around zero, minimal outliers exist, do not show autocorrelation, points on the graph do not curve up or down at the extremes, most points are close to the diagonal line in the center of the graph and do not cross the horizontal line on the graph that indicates a reference threshold for the residuals.

Field Cross-Validation of Allometric Equations Estimated by OLS For ***Pinus Hartwegii*** Lindl

Table 3. Field Cross-Validation of Allometric Equations Estimated By OLS, Mathematical Modeling, R Algorithm and Infys In 2023 For ***Pinus Hartwegii*** Lindl.

Evaluation of allometric equations estimated by mathematical methods that estimate TTV with bark in *Pinus hartwegii* Lindl			
Traditional allometric equations, estimated by fitting with mathematical models and estimated by INFyS		**Evaluation (%) by Federal Entity**	
Name	**Equation**	**México**	**Puebla**
Traditional mathematical equations used to estimate TTV volume with cortex (m³)			
Huber	LOG((3.1416/4)*((Alt_Tot)^2)*DN)	75.0	80.0
Smalian	LOG(((((3.1416*((Diameter_minor_log at 1.30 m /2)^2))+(3.1416*((DN)/2)^2))+(3.1416*((0/2)^2)))/2)*Alt_Tot)	85.0	90.0.
Newton	LOG((1/6)*((3.1416*((Diameter_minor_log at 1.30 m/2)^2))+(3.1416*((Largest_diameter_last_log /2)^2))+(4*3.1416*((DN/2)^2))))	95.0	98.0
Theoretical (wood logs) It is the LOG of the sum of the volumes of the following geometric figures of the tree.	**Neloid:** The geometric figure of the neloid in the bole was estimated with the equation of ((3.1416/16)*((normal_diameter at 1.30 m)^2))* Log_length at 1.30 m)	80.0	85.0
	Cylinder: Sum of logs that make up approx. the geometric figure of the cylinder in the bole with equation per log of ((3.1416/4)*((AVERAGE(Larger_diameter_log + Smaller_diameter_log))^2)* Length_log)	60.0	60.0
	Paraboloid: Sum of logs that make up approx. the geometric figure of the paravoloid in the stem with equation per log of ((3.1416/8)*((AVERAGE(Larger_diameter_log + Smaller_diameter_log))^2)*Length_log)	90.0	95.0
	Cone: The geometric figure of the cone in the stem was estimated with the equation of ((3.1416/12)*((Largest_diameter_last_log)^2)*Length_last_log_up_to _tip_tree)	70.0	75.0

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Evaluation of allometric equations estimated by mathematical methods that estimate TTV with bark in *Pinus hartwegii* Lindl			
Traditional allometric equations, estimated by fitting with mathematical models and estimated by INFyS		**Evaluation (%) by Federal Entity**	
Name	**Equation**	**México**	**Puebla**
Allometric equation used by government institutions by Federal Entity (state)	**Estado de México:** $TTV = e^{(-10.024)} * (DN)^{(2.06319)} * (Alt\_Tot)^{(0.86404)}$ This equation was adapted, by comparison to: $LOG(TTV) = LOG\left(e^{(-10.024)} * (DN)^{(2.06319)} * (Alt\_Tot)^{(0.86404)}\right)$ **Puebla:** $VOL = (-0.02920284) + ((0.00060663) * (DN^2)) + ((0.00001296) * (DN^2) * (Alt\_Tot))$ This equation was adapted, by comparison to: $LOG(VOL) = LOG\left((-0.02920284) + ((0.00060663) * (DN^2)) + ((0.00001296) * (DN^2) * (Alt\_Tot))\right)$	90.0	95.0
Traditional mathematical equations adjusted by mathematical models to estimate TTV volume with cortex (m^3)			
Newton	**Estado de México:** $\log(Vol) = -0.12 + 0 \log\_h + 0.33 \log\_diam$	90.0	0.0
	Puebla: $\log(Vol) = -0.07 + 0.05 \log\_h + 0.27 \log\_diam$	0.0	90.0
Theoretical (wood logs)	**Estado de México:** $\log(Vol) = 0.11 + 0.12 \log\_h + 0.25 \log\_diam$	95.0	0.0
	Puebla: $\log(Vol) = 0.12 + 0.15 \log\_h + 0.22 \log\_diam$	0.0	95.0
Allometric equation used by government institutions by Federal Entity (state)	**Estado de México:** $\log(Vol) = -9.94 + 0.85 \log\_h + 2.05 \log\_diam$	80.0	0.0
	Puebla: $\log(Vol) = -8.18 + 0.61 \log\_h + 1.79 \log\_diam$	0.0	80.0-85.0
Allometric equations estimated by SiBiFor [20]			
México	**T3_U1508_Mex_UMAFOR:1508:** $0.00005 * POTENCIA(DN, 1.97258) * POTENCIA(Alt\_Tot, 0.92797)$ **T4_U1510_Mex_UMAFOR:1510:** $0.00004 * POTENCIA(DN, 2.15224) * POTENCIA(Alt\_Tot, 0.78917)$	90.0	0.0
Puebla	**T5_U2101_Pue_UMAFOR:2101:** $0.00006 * POTENCIA(DN, 1.99376) * POTENCIA(Alt\_Tot, 0.90054)$ **T4_U2105_Pue_UMAFOR:2105:** $0.00004 * POTENCIA(DN, 2.20234) * POTENCIA(Alt\_Tot, 0.66536)$ **T8_U2108_Pue_UMAFOR:2108:** $0.00011 * POTENCIA(DN, 1.75875) * POTENCIA(Alt\_Tot, 0.95559)$ **TBz2y3_Pue_UMAFOR_-:** $EXP(-9.63495649 + 1.86670523 * LN(DN) + 0.99551381 * LN(Alt\_Tot))$	0.0	90.0

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

According to field cross-validation [21], the allometric equation estimated by mathematical models with the best estimate of the TTV cc are the theoretical models for excurrent dendrometric types with an evaluation of 90.5% and 95.0%, respectively. According to [22], The mathematical foundations of these models are based on if the response variable is denoted by Y and the explanatory variables by $X_1, X_2, \dots, X_k$, then a general model relating these variables is $E[Y\,|X_1 = x_1, X_2 = x_2, \dots, X_k = x_k] = \phi(x_1, x_2, \dots, x_k)$, although, we direct our attention to $\phi(x_1, x_2, \dots, x_k) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$ which is linear in the parameters β_j. This restriction to linearity is not as restrictive, they may be made linear by a suitable transformation using logarithms, we get the straight line $\log F = \log\alpha - \beta\log d$. For the linear model, the x_i could be functions of other variables z, w, etc.; We can also have $x_i = x^i$ which leads to a polynomial model; the linearity refers to the parameters, not the variables. With random X-variables, we carry out the regression conditionally on their observed values, provided that they are measured exactly (or at least with sufficient accuracy).

Suppose that after having fitted the regression model $E[Y] = X\beta, Var[Y] = \sigma^2 I_n$. We decide to introduce additional x_j´s into the model so that the model is now enlarged to $G: E[Y] = X\beta + Z\gamma = (X, Z)\begin{pmatrix}\beta\\ \gamma\end{pmatrix} = W\delta$. Say, where X is $n \times p$ of rank p, Z is $n \times t$ of rank t, and the columns of Z are linearly independent of the columns of X; that is, W is $n \times (p + t)$ of rank $p + t$. Then to find the least squares estimate $\hat{\delta}_G$ of δ there are two possible approaches. We can either compute $\hat{\delta}_G$ and its dispersion matrix directly from $\hat{\delta}_G = (WW)^{-1}W´Y$ and $Var[\hat{\delta}_G] = \sigma^2(W'W)^{-1}$, or to reduce the amount of computation, we can utilize the calculations already carried out in fitting the original model, as in Theorem below.

Lemma 1. *If* $R = I_n - P = I_n - X(X'X)^{-1}X'$, *then* $Z'RZ$ *is positive definite.*

Proof of Lemma 1. Let $Z'RZa = 0$; then, by Theorem 3.1(i), suppose that X is $n \times p$ of rank p, so that $P = X(X'X)^{-1}X'$. Then the following hold, (i) P and $I_n - P$ are symmetric and idempotent, $a'Z'R'RZa = a'Z'RZa = 0$, Hence $Za = X(X'X)^{-1}X'Za = Xb$, say, which implies that $a = 0$, as the columns of Z are linearly independent of the columns of X. Because $Z'RZa = 0$ implies that $a = 0, Z'RZ$ has linearly independent columns and is therefore nonsingular. Also, $a'Z'RZa = (RZa)'(RZa) \geq 0$.

Theorem 1. *Let* $R_G = I_n - W(W'W)^{-1}W', L = (X'X)^{-1}X'Z, M = (Z'RZ)^{-1}$, *and* $\hat{\delta}_G = \begin{pmatrix}\hat{\beta}_G\\ \hat{\gamma}_G\end{pmatrix}$. *Then: (i)* $\hat{\gamma}_G = (Z'RZ)^{-1}Z'RY$, *(ii)* $\hat{\beta}_G = (X'X)^{-1}X'(Y - Z\hat{\gamma}_G) = \hat{\beta} - L\hat{\gamma}_G$, *(iii)* $Y'R_GY = (Y - Z\hat{\gamma}_G)'R(Y - Z\hat{\gamma}_G) = Y'RY - \hat{\gamma}_G'Z'RY$, *(iv)* $Var[\hat{\delta}_G] = \sigma^2\begin{pmatrix}(X'X)^{-1} + LML' & -LM\\ -ML' & M\end{pmatrix}$.

Proof of Theorem 1. (i) We first "orthogonalize" the model. Since $\mathcal{C}(PZ) \subset \mathcal{C}(X)$, $X\beta + Z\gamma = X\beta + PZ\gamma + (I_n - P)Z\gamma = X\alpha + RZ\gamma = (X, RZ)\begin{pmatrix}\alpha\\ \gamma\end{pmatrix} = V\lambda$, say, where $\alpha = \beta + (X'X)^{-1}X'Z\gamma = \beta + L\gamma$ is unique. We note that $\mathcal{C}(X) \perp \mathcal{C}(RZ)$. Also, by $rank(A) \approx rank(A') \approx rank(A'A) \approx rank(AA')$ and the previous lemma, $rank(RZ) = rank(Z'R'RZ) = rank(Z'RZ) = t$, so that V has full rank $p + t$. Since $XR = 0$, the least squares estimate of λ is $(V'V)^{-1}V'Y = \begin{pmatrix}X'X & X'RZ\\ Z'RX & Z'R'RZ\end{pmatrix}^{-1}\begin{pmatrix}X'\\ Z'R\end{pmatrix}Y = \begin{pmatrix}X'X & 0\\ 0 & Z'R'RZ\end{pmatrix}^{-1}\begin{pmatrix}X'\\ Z'R\end{pmatrix}Y = \begin{pmatrix}(X'X)^{-1}X'Y\\ (Z'RZ)^{-1}Z'RY\end{pmatrix} = \begin{pmatrix}\hat{\alpha}\\ \hat{\gamma}\end{pmatrix}$. Now the relationship between (β, γ) and (α, γ) is one-to-one, so that the same relationships exist between their least square estimates. Hence $\hat{\gamma}_G = \breve{\gamma} = (Z'RZ)^{-1}Z'RY$. (ii) $\hat{\beta}_G = \breve{\alpha} - L\breve{\gamma} = \hat{\beta} - L\hat{\gamma}_G = (X'X)^{-1}X'(Y - Z\hat{\gamma}_G)$. (iii) Using (ii) gives $R_GY = Y - X\hat{\beta}_G - Z\hat{\gamma}_G = Y - (X'X)^{-1}X'(Y - Z\hat{\gamma}_G) - Z\hat{\gamma}_G = (I_n - (X'X)^{-1}X')(Y - Z\hat{\gamma}_G) = R(Y - Z\hat{\gamma}_G) = RY - RZ(Z'RZ)^{-1}Z'RY$ so that by $R(Y - Z\hat{\gamma}_G)$, $Y'R_GY = (Y - W\hat{\delta}_G)'(Y - W\hat{\delta}_G) = (Y - X\hat{\beta}_G - Z\hat{\gamma}_G)'(Y - X\hat{\beta}_G - Z\hat{\gamma}_G) = (Y - Z\hat{\gamma}_G)'R'R(Y - Z\hat{\gamma}_G) = (Y - Z\hat{\gamma}_G)'R(Y - Z\hat{\gamma}_G)$, since R is symmetric

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

and idempotent, Theorem 3.1(i), *suppose that* X *is* $n\ x\ p$ *of rank* p*, so that* $P = X(X'X)^{-1}X'$. Then the following hold, (i) P and $\mathrm{I_n - P}$ are symmetric and idempotent. Expanding equation $(\mathrm{Y} - Z\hat{\gamma}_G)'\mathrm{R}(\mathrm{Y} - Z\hat{\gamma}_G)$ gives us $\mathrm{Y'R_GY = Y'RY} - 2\hat{\gamma}_G'\mathrm{Z'RY} + \hat{\gamma}_G'\mathrm{Z'RZ}\hat{\gamma}_G = \mathrm{Y'RY} - \hat{\gamma}_G'\mathrm{Z'RY} - \hat{\gamma}_G'(\mathrm{Z'RY - Z'RZ}\hat{\gamma}_G) = \mathrm{Y'RY} - \hat{\gamma}_G'\mathrm{Z'RY}$ by $\hat{\gamma}_G = \check{\gamma} = (\mathrm{Z'RZ})^{-1}\mathrm{Z'RY}$. (iv) $Var[\hat{\gamma}_G] = (\mathrm{Z'RZ})^{-1}\mathrm{Z'R\,Var[Y]RZ(Z'RZ)^{-1}} = \sigma^2(\mathrm{Z'RZ})^{-1}(\mathrm{Z'RZ})(\mathrm{Z'RZ})^{-1} = \sigma^2(\mathrm{Z'RZ})^{-1} = \sigma^2\mathrm{M}$. Now, by Theorem, *if* X *and* Y *Y are* $m \times 1$ *and* $n \times 1$ *vectors of random variables, and* A *and* B *are* $l \times m$ *and* $p \times n$ *matrices of constants, respectively, then* $Cov[AX, BY] = A\,Cov[X,Y]B'$, $\mathrm{Cov}[\hat{\beta}, \hat{\gamma}_G] = \mathrm{Cov[(X'X)^{-1}X'Y, (Z'RZ)^{-1}Z'RY]} = \sigma^2\mathrm{(X'X)^{-1}X'RZ(Z'RZ)^{-1}} = 0$, since $\mathrm{X'R} = 0$. Hence using (i) above, we have, from previous Theorem, $\mathrm{Cov}[\hat{\beta}, \hat{\gamma}_G] = \mathrm{Cov}[\hat{\beta} - \mathrm{L}\hat{\gamma}_G, \hat{\gamma}_G] = \mathrm{Cov}[\hat{\beta}, \hat{\gamma}_G] - \mathrm{LVar}[\hat{\gamma}_G] = -\sigma^2\mathrm{LM}$ by $\mathrm{Cov}[\hat{\beta}, \hat{\gamma}_G] = \mathrm{Cov[(X'X)^{-1}X'Y, (Z'RZ)^{-1}Z'RY]} = \sigma^2\mathrm{(X'X)^{-1}X'RZ(Z'RZ)^{-1}} = 0$ and $\mathrm{Var}[\hat{\beta}_G] = \mathrm{Var}[\hat{\beta} - \mathrm{L}\hat{\gamma}_G] = \mathrm{Var}[\hat{\beta}] - \mathrm{Cov}[\hat{\beta}, \mathrm{L}\hat{\gamma}_G] - \mathrm{Cov}[\mathrm{L}\hat{\gamma}_G, \hat{\beta}] + \mathrm{Var}[\mathrm{L}\hat{\gamma}_G] = \mathrm{Var}[\hat{\beta}] + \mathrm{LVar}[\hat{\gamma}_G]\mathrm{L}'$, by $= -\sigma^2\mathrm{LM}$, $= \sigma^2[\mathrm{(X'X)^{-1} + LML'}]$. From **Theorem 1** we see that once $\mathrm{X'X}$ has been inverted, we can find $\hat{\delta}_G$ and its variance-covariance matrix simply by inverting the $t\ x\ t$ matrix $\mathrm{Z'RZ}$ we need not invert the $(\mathrm{p + t})\ \mathrm{x}\ (\mathrm{p + t})$ matrix $\mathrm{W'W}$.

*Allometric Equations Estimated by OLS With Ridge Model for **Pinus Hartwegii** Lindl*

Table 4. Summary Of Allometric Equations Estimated by OLS With Ridge Model for ***Pinus Hartwegii*** Lindl with Significant Climate Variables (See Appendix A For Allometric Equations Estimated by Region).

Concept	***Estado de México***			***Estado de Puebla***		
	Method for commercial volumetric dimension			***Method for commercial volumetric dimension***		
	Newton	***Theoretical (wood logs)***	***Federal entity***	***Newton***	***Theoretical (wood logs)***	***Federal entity***
Percentage of variance explained (R^2)	*0.951499*	*0.939338*	*0.928746*	*0.930383*	*0.934537*	*0.935750*
(Intercept)	*39.859663*	*50.647257*	*-64.304510*	*105.415300*	*154.616600*	*-131.921100*
Total height (h)	*0.015720*	*0.045025*	*0.031397*	*0.008089*	*0.027984*	*0.013225*
Normal diameter (diam)	*0.025014*	*0.026166*	*0.028203*	*0.014932*	*0.018439*	*0.047649*
Earth Skin Temperature (C) (TS)	*-0.010908*	*-0.012784*	*0.001895*	*-0.003008*	*-0.005506*	*0.007905*
Temperature at 2 Meters (C) (T2M)	*-0.010975*	*-0.013165*	*0.007274*	*-0.012261*	*-0.020297*	*0.021197*
Top-Of-Atmosphere Shortwave Downward	*-1.924237*	*-2.505947*	*2.253705*	*-3.461111*	*-4.817098*	*3.986277*

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Concept	***Estado de México***			***Estado de Puebla***		
	Method for commercial volumetric dimension			***Method for commercial volumetric dimension***		
	Newton	***Theoretical (wood logs)***	***Federal entity***	***Newton***	***Theoretical (wood logs)***	***Federal entity***
Irradiance (TOA_SW_DWN)						
All Sky Surface PAR Total (ALLSKY_SFC_PAR_TOT)	*0.002949*	*0.004038*	*-0.003183*	*0.004371*	*0.005090*	*-0.003492*
Clear Sky Surface PAR Total (CLRSKY_SFC_PAR_TOT)	*0.006476*	*0.008882*	*-0.004877*	*0.005972*	*0.005833*	*-0.003405*
Surface Pressure (PS)	*0.338027*	*0.454903*	*-0.163521*	*0.178731*	*0.146928*	*-0.074175*
Specific Humidity at 2 Meters (QV2M)	*0.045728*	*0.059961*	*-0.013913*	*0.013604*	*0.009449*	*-0.004173*
Relative Humidity at 2 Meters (RH2M)	*0.004126*	*0.005259*	*-0.000788*	*0.000691*	*0.000407*	*-0.000158*
Root Zone Soil Wetness (GWETROOT)	*0.229238*	*0.283766*	*-0.027523*	*0.021713*	*0.010888*	*-0.003766*
Precipitation Corrected (PRECTOTCORR)	*-0.029698*	*-0.035689*	*0.002254*	*-0.001604*	*-0.000687*	*0.000212*
Age_no_rings	*0.006924*	*0.006797*	*-0.000682*	*0.005656*	*0.007813*	*0.015973*

Based on the “commercial volumetric dimension” category of the INFyS numerical bases, the application of the 'randomForest' package to estimate the variable “Age_no_rings” in UMAFOR 1508 of the state of Mexico, 2105 of the state of Puebla, the use of POWER data, the relevant climatic variables in the volumetric growth of a tree ([23–25]), Paterson C. P. V. Productivity Index, the allometric equations of “commercial volumetric dimension” were estimated for ***Pinus hartwegii*** Lindl with significant climatic variables TS, T2M, TOA_SW_DWN, ALLSKY_SFC_PAR_TOT, PS, QV2M, RH2M, GWETROOT, PRECTOTCORR and Age_no_rings. According to [40], the advantage of Ridge regression over least squares has its origin in bias-variance compensation and, according to [26], its mathematical foundations are supported by the standardized form of the regression model is given as $\widehat{\mathrm{Y}} = \theta_1 \widehat{\mathrm{X}}_1 + \theta_2 \widehat{\mathrm{X}}_2 + \cdots + \theta_p \widehat{\mathrm{X}}_p + \epsilon'$. The estimating equations for the ridge regression coefficients are

$$\begin{array}{ccccccc} (1+k)\theta_1 & + & r_{12}\theta_2 & +\cdots+ & r_{1p}\theta_p & = & r_{1y,} \\ r_{21}\theta_1 & + & (1+k)\theta_2 & +\cdots+ & r_{2p}\theta_p & = & r_{2y,} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ r_{p1}\theta_1 & + & r_{p2}\theta_1 & +\cdots+ & (1+k)\theta_p & = & r_{py,} \end{array} \quad (6)$$

where r_{1j} is the correlation between the ith and jth predictor variables and $r_{iy,}$ is the correlation between the ith predictor variable and the response variable $\widehat{\mathrm{Y}}$. The solution to (6), $\hat{\theta}_1, \ldots, \hat{\theta}_p$, is the set of estimated ridge regression coefficients. The essential parameter that distinguishes ridge regression from OLS is k. On the other hand, the total variance (the sum of the variances of the estimated regression coefficients), is

$$\text{Total Variance}(\mathrm{k}) = \sum_{\mathrm{j}=1}^{\mathrm{p}} \mathrm{Var}\left(\hat{\theta}_{\mathrm{j}}(\mathrm{k})\right) = \sigma^2 \sum_{\mathrm{j}=1}^{\mathrm{p}} \left(\frac{\lambda_{\mathrm{j}}}{(\lambda_{\mathrm{j}}+\mathrm{k})^2}\right) \quad (7)$$

Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

which is a decreasing function of k. Substituting $k = 0$ in (7), we obtain $\text{Total Variance}(0) = \sigma^2 \sum_{j=1}^{p}\left(\frac{1}{\lambda_j}\right)$ which shows the effect of small eigenvalue on the total variance of the OLS estimates of the regression coefficients. In practice, a value of k is chosen by computing $\hat{\theta}_1, \dots, \hat{\theta}_p$ for a range of k values between 0 and 1 and plotting the results against k. The estimator is (for a given value of k) $\hat{\theta}(k) = (Z^TZ + kI)^{-1}Z^TY = (Z^TZ + kI)^{-1}Z^TZ\hat{\theta}$. The expected value of $\hat{\theta}(k)$ is $E[\hat{\theta}(k)] = (Z^TZ + kI)^{-1}Z^TZ\theta$ and the variance-covariance matrix is $\text{Var}[\hat{\theta}(k)] = (Z^TZ + kI)^{-1}Z^TZ(Z^TZ + kI)^{-1}\sigma^2$. The variance inflation factor, $\text{VIF}_j(k)$, as a function of k is the jth diagonal element of the matrix $(Z^TZ + kI)^{-1}Z^TZ(Z^TZ + kI)^{-1}$. The residual sum of squares can be written as $\text{SSE}(k) = \left(Y - Z\hat{\theta}(k)\right)^T\left(Y - Z\hat{\theta}(k)\right) = (Y - Z\hat{\theta})^T(Y - Z\hat{\theta}) + (\hat{\theta}(k) - \hat{\theta})^T Z^TZ(\hat{\theta}(k) - \hat{\theta})$. The total mean square error is $\text{SSE}(k) = E\left[(\hat{\theta}(k) - \hat{\theta})^T(\hat{\theta}(k) - \hat{\theta})\right] = \sigma^2\text{trace}\left[(Z^TZ + kI)^{-1}Z^TZ(Z^TZ + kI)^{-1}\right] + k^2\theta^T(Z^TZ + kI)^{-2}\theta = \sigma^2\sum_{j=1}^{p}\lambda_j(\lambda_j + k)^{-2} + k^2\theta^T(Z^TZ + kI)^{-2}\theta$. Note that the first term on the right-hand side of last equation is the sum of the variances of the components of $\hat{\theta}(k)$ (total variance) and the second term is the square of the bias.

There exists a value of $k > 0$ such that $E\left[(\hat{\theta}(k) - \theta)^T(\hat{\theta}(k) - \theta)\right] < E\left[(\hat{\theta} - \theta)^T(\hat{\theta} - \theta)\right]$ that is, the mean square error of the ridge estimator, $\hat{\theta}(k)$, is less than the mean square error of the OLS estimator, $\hat{\theta}$. To see one possible generalization, consider the regression model restated in terms of the principal components, $C = (C_1, \dots, C_p)$. The general model takes the form $Y = C\alpha + \varepsilon$, where

$$C = ZV, \qquad \alpha = V^T\theta,$$
$$V^TZ^TZV = A, \quad V^TV = VV^T = I,$$

And

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_{p-1} & 0 \\ 0 & 0 & 0 & & 0 & \lambda_p \end{pmatrix}$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$ is a diagonal matrix consisting of the ordered eigenvalues of Z^TZ. The total mean square error in $\sigma^2\sum_{j=1}^{p}\lambda_j(\lambda_j + k)^{-2} + k^2\theta^T(Z^TZ + kI)^{-2}\theta$ becomes $\text{TMSE}(k) = E\left[(\hat{\theta}(k) - \theta)^T(\hat{\theta}(k) - \theta)\right] = \sigma^2\sum_{j=1}^{p}\left(\frac{\lambda_j}{(\lambda_j + k)^2}\right) + \sum_{j=1}^{p}\left(\frac{k^2\alpha_j^2}{(\lambda_j + k)^2}\right)$ where $\alpha^T = (\alpha_1, \alpha_2, \dots, \alpha_p)$. Instead of taking a single value for k, we can consider several different values k, say $k_1, k_2, \dots, k_p$. The quantity k, instead of being a scalar, is now a vector and denoted by k. The total mean square error given in $\text{TMSE}(k)$ now becomes $\text{TMSE}(k) = E\left[(\hat{\theta}(k) - \theta)^T(\hat{\theta}(k) - \theta)\right] = \sigma^2\sum_{j=1}^{p}\left(\frac{\lambda_j}{(\lambda_j + k)^2}\right) + \sum_{j=1}^{p}\left(\frac{k^2\alpha_j^2}{(\lambda_j + k)^2}\right)$. The total mean square error given in last $\text{TMSE}(k)$ is minimized by taking $k_j = \sigma^2/\alpha_j^2$.

Estimates of the Linear Coefficients of the Significant Climate Variables in the Period 2022-2024 for ***Pinus Hartwegii*** Lindl

Table 5. Estimates of the Linear Coefficients of the Significant Climate Variables in the Period 2022-2024 for ***Pinus Hartwegii*** Lindl in the States of México Y Puebla.

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Climate variable	Estimation of linear coefficients by Mexican Federal Entity					
	México			Puebla		
	2022	2023	2024	2022	2023	2024
Earth Skin Temperature (C) (TS)	17.3631605	17.2562204	17.2075609	15.0913378	15.0938670	15.0937800
Temperature at 2 Meters (C) (T2M)	16.0688180	15.9727000	15.9824568	13.9756634	13.9538095	13.9546881
Top-Of-Atmosphere Shortwave Downward Irradiance (TOA_SW_DWN)	34.0226381	34.0161901	34.0216112	34.2100035	34.2068739	34.2110193
All Sky Surface PAR Total (ALLSKY_SFC_PAR_TOT)	102.2113483	101.4544898	101.8056317	107.4957497	107.3481481	107.4148577
Clear Sky Surface PAR Total (CLRSKY_SFC_PAR_TOT)	122.7126980	122.6059464	122.5449116	125.1707133	125.4006844	125.3800238
Surface Pressure (PS)	77.2082785	77.2019272	77.2015360	76.6635899	76.6581183	76.6585532
Specific Humidity at 2 Meters (QV2M)	8.4555264	8.5050894	8.4706968	9.2473988	9.2468102	9.2464988
Relative Humidity at 2 Meters (RH2M)	61.9231341	62.2058595	62.3261776	72.2132697	72.4343616	72.0744410
Root Zone Soil Wetness (GWETROOT)	0.4857650	0.4844374	0.4903292	0.4207275	0.4195527	0.4167194
Precipitation Corrected (PRECTOTCORR)	1.8141824	1.7827351	1.8955158	2.6239403	2.6892061	2.6296263

Based on the evaluation of the mathematical model of the timber volume based on the allometric equations estimated by OLS with the Ridge Model, the simulations with the linear coefficients of diameter-height per tree located in the UMAFOR and the significant climate variables for the time period 2022 -2024 in the states of Puebla and México, there is a better volumetric approximation of Newton's mathematical equations and theoretical models for excurrent dendrometric types. According to [27], “if a coefficient of 50.81 % is applied to the biomass calculated with the allometric equations for basal diameter or total height, it is sufficient to know the carbon content in a tree, a stand or a sapling plantation of ***Pinus hartwegii*** Lindl; therefore, its use is reliable.” Although, [28] affirm that the distribution of biomass in the trees was 65.3% in stem, 23.8% in branches, 10.9% in foliage, the equation for estimating biomass was $B = 0.0635\text{DN}^{2.4725}$ and for biomass content carbon $C = 0.0309\text{DN}^{2.4722}$, with R^2 de 0.98. The adjusted model is of the form $Y = bX^k$where the dependent variable (Y) is biomass or carbon and the independent variable is the normal diameter.

Discussion

Based on the application of OLS, R algorithms, Theorems, t_{n-p-1} *distribution*, p-value with values $\ll 0.001$, interpretive criteria of frequentist statistics, residual validation graphs, residuals vs. fitted, normal Q-Q, scale location model, residuals vs. laverage and $\alpha \ll 0.001$, there is randomness, homoscedasticity, normal distribution, independence, absence of non-linear patterns, close to zero, concentration in the center, symmetry, there is no fan pattern, cone, U shape or a curve, they have a constant dispersion along the independent variable or adjusted values, they are centered around zero, there are minimal outliers, they do not show autocorrelation and there is uniformity in the dispersion in the residuals, H_o and it is accepted that the allometric equations of “commercial volumetric dimension” estimated, with dasometric

Journal of Ecohumanism
2024
Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

measurements of dominant and co-dominant trees, by mathematical models for ***Pinus hartwegii*** Lindl are significant and explain the variability of Y, with the theoretical models for excurrent dendrometric types (volume per log) being the best estimate of the TTV cc in the states of Mexico and Puebla, with an evaluation of 90.5% and 95.0%, respectively.

On the one hand, although the dasometric measurements in many individuals are atypical, the number of observations is few and the information on the variable "Age_no_rings" is non-existent in UMAFOR 1508, state of Mexico and 2105, state of Puebla, since The trees are located in the ANPs "Parque Nacional Nevado de Toluca" and "Parque Nacional Iztaccíhuatl-Popocatépetl", respectively, the estimated allometric equations allow a more precise estimation of the TTV cc than the Newton equations, with an evaluation of 90.0 % and 90.0%, or those traditionally used in the Federal Entities, with an evaluation of 80.0% and between 80.0-85.0% respectively, of the states of Mexico and Puebla. On the other hand, the simulation of Newton's mathematical equations and theoretical models for excurrent dendrometric types (theoretical or volume per log) estimated by OLS with the Ridge Model and the significant climatic variables, in the period 2022-2024, offers a volumetric approximation significant timber yield of ***Pinus hartwegii*** Lindl, in the face of climate change scenarios.

Therefore, according to [5], this research is a particular case in Mexico, since there is no previous research that is similar, it includes estimation of allometric equations by mathematical models based on sets of biocene and abiocene parts *in situ*, Random Forest and Ridge algorithms, climate bases, simulation with mathematical modeling, mathematical validation, cross-validation in the field and mathematical demonstration with the objective of obtaining significant allometric equations, practical in the field, that describe the temporal evolutionary displacement and estimate the timber volumetric increase of the ***Pinus hartwegii*** Lindl forest in the states of Mexico and Puebla in the face of climate change scenarios. Finally, it creates the theoretical-practical mathematical research bases for future investigations of bio-timber forest species sensitive to climate change, such as ***Abies religiosa*** (Kunth Schltdl. et Cham.) and ***Pinus montezumae*** in the different states of the Mexican Republic.

Conclusions

- Ho is rejected and it is accepted that the linear regression models are significant and explain the variability of Y; The null hypotheses are rejected and the alternatives are accepted that the variables are significant in the models; The null hypotheses are rejected and the alternatives are accepted; there is randomness, homoscedasticity, normal distribution, independence, absence of non-linear patterns, close to zero, concentration in the center, symmetry and uniformity in the dispersion in the residues in allometric equations of "commercial volumetric dimension" estimated by mathematical models for Pinus hartwegii Lindl in the states of México and Puebla.

- The allometric equation estimated by mathematical models with the best estimate of the TTV cc is theoretical models for excurrent dendrometric types (volume per log) in the states of México and Puebla with an evaluation of 90.5% and 95.0%, respectively, while the second equation with the best estimate of the TTV cc is Newton in the states of Mexico and Puebla with an evaluation of 90.0% and 90.0%, respectively and, finally, the allometric equation estimated with the lowest estimate of the TTV cc is the Federal Entity with an evaluation of 80.0% and between 80.0% and 85.0%, respectively.

- "Commercial volumetric dimension" allometric equations were estimated for ***Pinus hartwegii*** Lindl with significant climate variables Earth Skin Temperature (C) (TS), Temperature at 2 Meters (C) (T2M), Top-Of-Atmosphere Shortwave Downward Irradiance (TOA_SW_DWN), All Sky Surface PAR Total (ALLSKY_SFC_PAR_TOT), Surface Pressure (PS), Specific Humidity at 2 Meters (QV2M), Relative Humidity at 2 Meters (RH2M), Root Zone Soil Wetness (GWETROOT), Precipitation Corrected (PRECTOTCORR) and Age_no_rings.

ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

- In the period 2022-2024, which approximate the timber volume of ***Pinus hartwegii*** Lindl, there is a better volumetric approximation in the face of climate change by Newton's mathematical equations and theoretical models for excurrent dendrometric types (theoretical or volume per wood logs).

Author Contributions: The contribution of the co-authors was Moisés Arreguín Sámano conceptualization and numerical databases; José Abelardo Paucar Camacho methodology, investigation, writing—review, risk analysis and editing; Eduardo Patricio Salazar Castañeda organization of databases, data curation, visualization, formal, validation, resources, risk analysis and writing—original draft preparation; María Transito Vallejo Illijama methodology, investigation, writing—review, risk analysis and editing; Miguel Ángel Guallpa Calva organization of databases, data curation, formal, validation, resources, risk analysis, visualization and writing—original draft preparation; Ángel Leyva Ovalle methodology, software, R algorithms, resources, data curation, supervision, visualization , writing—review and editing

Funding: "This research was funded jointly by the co-authors".

Data Availability Statement: The set of databases are available in the following links: https://snmf.cnf.gob.mx/datos-del-inventario/ (Inventory data-SNMF CNF, INFyS), https://power.larc.nasa.gov/data-access-viewer/ (Climate data POWER), https://drive.google.com/drive/folders/1I12k6NeRfrNU9Qz70LPcMTerKv1Qh0oR?usp=drive_link (R Algorithm).

Acknowledgments: We wish to express our infinite gratitude to Ing. Mat. Andrés Mauricio Acuña Torres for his invaluable contribution, admirable scientific mathematical and programming knowledge, Dr. De los Santos Posadas Héctor Manuel from Colegio de Postgraduados, Montecillos, estado de México for his guidance, recommendations and guidance in the databases, Dr. Guillermo Carrillo Espinosa from the División de Ciencias Forestales, Universidad Autónoma Chapingo, Government of México, specifically to the Comisión Nacional Forestal (CONAFOR) for the numerical bases of the Inventory data-SNMF CNF, INFyS.

Conflicts of Interest: "The authors declare no conflicts of interest."

References

Santiago G., W., Quiñonez B., G., De los Santos P., H., & Rodríguez O., G. (n.d.). Avances y perspectivas en la modelación aplicada a la planeación forestal en México. Retrieved February 3, 2024, from https://www.redalyc.org/journal/617/61763413015/html/

Torrecilla Díaz, R., Diéguez Batista, R., & Hall Aguilar, L. (2015). La formación de la capacidad modeladora matemática en el ingeniero. Rafael María de Mendive, 1–6.

Vega, G. C. (2022). Clasificación de modelos matemáticos. Teoría de modelos. IESIP: Caracas, Venezuela.

Sandoval, M. M. I. (2020). Análisis de las iniciativas del Ecuador en las negociaciones internacionales sobre cambio climático COP21 y COP23. Área de Estudios Sociales y Globales. Sede Ecuador. Universidad Andina Simón Bolívar: Quito, Ecuador.

Valderrama Bonnet, M. J. (2009). Modelos estocásticos dinámicos. Granada, España.

Comisión Nacional Forestal. (s.f.). Datos del inventario. Sistema Nacional de Monitoreo Forestal. Recuperado el 26 de junio de 2024, de https://snmf.cnf.gob.mx/datos-del-inventario/

Velasco-Bautista, E. (2015). El desarrollo de técnicas estadísticas para la evaluación y monitoreo de ecosistemas forestales en México (Tesis doctoral). Colegio de Postgraduados.

Leyva Ovalle, A. (2016). Sistema de cómputo para el procesamiento y análisis de datos del inventario nacional forestal y de suelos (Tesis doctoral). Colegio de Postgraduados. https://doi.org/10.13140/RG.2.2.18209.07522

Leyva-Ovalle, A., Valdez-Lazalde, J. R., De Los Santos-Posadas, H. M., Martínez-Trinidad, T., Herrera, A., Lugo, O., & García-Nava, J. R. (2017). Monitoreo de la degradación forestal en México con base en el Inventario Nacional Forestal y de Suelos (INFyS). Madera y Bosques, 23(2), 69-83. https://doi.org/10.21829/myb.2017.2321431

Piña-Monarrez, M., Rodríguez-Medina, M. A., & Aguirre-Solís, J. (2007). Regresión Ridge y la distribución central t. Ciencia Ergo Sum, 14(1).

Mohamed, N. F., Govindasamy, P., Isa, N. J. M., Noor, A. M., Ma, L., Olmos, A., & Green, K. (2024). A systematic review of exploratory factor analysis packages in R software. WIREs Computational Statistics, 16(1), e1630. https://doi.org/10.1002/wics.1630

Volume: 3, No: 8, pp. 6936 – 6952
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v3i8.5292

Al-Kilani, M. R., Rahbeh, M., Al-Bakri, J., Tadesse, T., & Knutson, C. (2021). Evaluation of remotely sensed precipitation estimates from the NASA POWER project for drought detection over Jordan. Earth Systems and Environment, 5(3), 561-573. https://doi.org/10.1007/s41748-021-00245-2

NASA. (s.f.). POWER Prediction of Worldwide Energy Resources (POWER). Retrieved February 3, 2024, from https://power.larc.nasa.gov/docs/methodology/meteorology/assessment/

NASA. (s.f.). NASA Power Data Access. Retrieved February 3, 2024, de https://power.larc.nasa.gov/data-access-viewer/|

Moreno Quintero, E. S., & Cristiano Muñoz, G. N. (2021). Técnicas de regularización en el marco del aprendizaje de máquina: regresiones ridge y lasso.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67. https://doi.org/10.2307/1267351

Vanclay, J. (1994). Modelling forest growth and yield: Applications to mixed tropical forests. CAB International. ISBN 978-0-85198-913-6.

Coble, D. W., Milner, K. S., & Marshall, J. D. (2001). Above- and below-ground production of trees and other vegetation on contrasting aspects in western Montana: A case study. Forest Ecology and Management, 142(1-3), 231-241. https://doi.org/10.1016/S0378-1127(00)00353-4.

Alaminos, F. A. F. (2023). Árboles de decisión en R con Random Forest. OBETS Ciencia Abierta, Instituto de Desarrollo Social y Paz, Universidad de Alicante.

Larreta, B., Aguirre Calderón, O. A., Aguirre-Calderón, C., Zamudio Sánchez, F., Corral-Rivas, J. J., Garza, E., De Los Santos-Posadas, H. M., Martinez-Salvador, M., & López-Martínez, J. (2018). Manual del Sistema Biométrico Forestal (SIBIFOR): Herramientas para el manejo de los bosques templados y tropicales de México. Retrieved from https://biblioteca.semarnat.gob.mx/janium/Documentos/Ciga/libros2018/CD006165.pdf;

Carrillo-Espinoza, G. (2024). Validación cruzada de campo.

Seber, G. A. F., & Lee, A. J. (2012). Linear regression analysis. John Wiley & Sons. ISBN 978-1-118-27442-2.

Walter, H. (1977). Zonas de vegetación y clima. Ediciones Omega.

Valverde-Valdés, Y., & Cano-Santana, Z. (2005). Ecología y medio ambiente. Pearson Educación. ISBN 978-970-26-0536-2.

Carvajal, A. F., & Pabón, J. D. (2014). Temperatura de la superficie terrestre en diferentes tipos de cobertura de la Región Andina Colombiana. Sociedade & Natureza, 26, 95–112. https://doi.org/10.1590/1982-451320140107

Chatterjee, S., & Hadi, A. S. (2006). Regression analysis by example. John Wiley & Sons. ISBN 978-0-470-05545-8.

Martínez-Luna, J. E., Carrillo-Anzures, F., & Romero Sánchez, M. E. (2020). Ecuaciones alométricas para estimar carbono en brinzales de Pinus hartwegii Lindl. Revista Mexicana de Ciencias Forestales. Retrieved from https://cienciasforestales.inifap.gob.mx/index.php/forestales/article/view/726

Carrillo Anzures, F., Acosta Mireles, M., Jiménez Cruz, C. del R., González Molina, L., & Etchevers Barra, J. D. (2016). Ecuaciones alométricas para estimar la biomasa y el carbono de la parte aérea de Pinus hartwegii en el Parque Nacional Ixta-Popo, México. Revista Mexicana de Ciencias Agrícolas, 7, 681–691. Retrieved from https://dialnet.unirioja.es/servlet/articulo?codigo=5699354.