



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Postgrado
Maestría en Estadística Aplicada

**ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO
MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES
DE TECA (*TECTONA GRANDIS L.F.*), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA
VERAPAZ, GUATEMALA**

Ing. Eric Abel Ortega Orellana

Asesorado por el Msc. Edwin Estuardo Vaides López

Guatemala, marzo de 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO
MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES
DE TECA (*TECTONA GRANDIS L.F.*), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA
VERAPAZ, GUATEMALA.**

TRABAJO DE GRADUACIÓN

PRESENTADO A LA JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA
POR

ING. ERIC ABEL ORTEGA ORELLANA
ASESORADO POR EL MSC. EDWIN ESTUARDO VAIDES LÓPEZ

AL CONFERÍRSELE EL TÍTULO DE

MAESTRO EN ESTADÍSTICA APLICADA

GUATEMALA, MARZO 2023

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Vladimir Cruz Lorente
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN DE DEFENSA DE TÉSIS

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADOR	Mtro. Ing. William Eduardo Fagiani Cruz
EXAMINADOR	Mtro. Ing. Edwin Adalberto Bracamonte Orozco
EXAMINADORA	Dra. Aura Marina Rodriguez Pérez
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

**ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO
MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES
DE TECA (*TECTONA GRANDIS L.F.*), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA
VERAPAZ, GUATEMALA**

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Postgrado, con fecha 6 de agosto de 2022.



Ing. Eric Abel Ortega Orellana



Decanato
Facultad de Ingeniería
24189101-24189102
secretariadecanato@ingenieria.usac.edu.gt

LNG.DECANATO.OI.259.2023

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Estudios de Posgrado, al Trabajo de Graduación titulado: **ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES DE TECA (TECTONA GRANDIS L.F), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA VERAPAZ, GUATEMALA**, presentado por: Ing. Eric Abel Ortega Orellana, que pertenece al programa de Maestría en artes en Estadística aplicada después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:


Inga. Aurelia Anabela Cordova Estrada



Decana

Guatemala, marzo de 2023

AACE/gaoc

Escuelas: Ingeniería Civil, Ingeniería Mecánica Industrial, Ingeniería Química, Ingeniería Mecánica Eléctrica, - Escuela de Ciencias, Regional de Ingeniería Sanitaria y Recursos Hídricos (RSR), Post-Grado Maestría en Sistemas Mención Ingeniería Civil, Carreras: Ingeniería Mecánica, Ingeniería Eléctrica, Ingeniería en Ciencias y Sistemas, Licenciatura en Matemática, Licenciatura en Física, Centro de Estudios Superiores de Energía y Minas (CESEM), Guatemala, Ciudad Universitaria, Zona 12, Guatemala, Centroamérica.



Guatemala, marzo de 2023

LNG.EEP.OL259.2023

En mi calidad de Director de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor, verificar la aprobación del Coordinador de Maestría y la aprobación del Área de Lingüística al trabajo de graduación titulado:

**"ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y
SELECCIONAR EL MODELO MATEMÁTICO DE MEJOR AJUSTE, QUE
PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES DE TECA (TECTONA
GRANDIS L.F), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA
VERAPAZ, GUATEMALA"**

presentado por Ing. Eric Abel Ortega Orellana
correspondiente al programa de Maestría en artes en Estadística
aplicada ; apruebo y autorizo el mismo.

Atentamente

"Id y Enseñad a Todos"



Mtro. Ing. Edgar Dario Álvarez Cotí
Director

Escuela de Estudios de Postgrado
Facultad de Ingeniería





Guatemala, 29 de septiembre de 2022

M.A. Ing. Edgar Dario Alvarez Cotí
Director
Escuela de Estudios de Postgrado
Presente

Estimado M.A. Ing. Alvarez Cotí

Por este medio informo a usted, que he revisado y aprobado el **INFORME FINAL Y ARTÍCULO CIENTÍFICO** titulado: **ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES DE TECA (TECTONA GRANDIS L.F), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA VERAPAZ, GUATEMALA** del estudiante **Eric Abel Ortega Orellana** quien se identifica con número de carnet 009517275 del programa de Maestría En Estadística Aplicada.

Con base en la evaluación realizada hago constar que he evaluado la calidad, validez, pertinencia y coherencia de los resultados obtenidos en el trabajo presentado y según lo establecido en el **Normativo de Tesis y Trabajos de Graduación aprobado por Junta Directiva de la Facultad de Ingeniería Punto Sexto Inciso 6.10 del Acta 04-2014 de sesión celebrada el 04 de febrero de 2014**. Por lo cual el trabajo evaluado cuenta con mi aprobación.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.



Mtro. Ing. Edwin Adalberto Bracamonte Orozco
Coordinador
Maestría En Estadística Aplicada
Escuela de Estudios de Postgrado

Guatemala, 29 de septiembre de 2022

M.A. Ing. Edgar Dario Álvarez Coti
Director
Escuela de Estudios de Postgrados
Presente

Estimado M.A. Ing. Álvarez Coti

Por este medio informo a usted, que he revisado y aprobado el Trabajo de Graduación y el Artículo Científico: "ANÁLISIS DE REGRESIÓN MÚLTIPLE PARA EVALUAR Y SELECCIONAR EL MODELO MATEMÁTICO DE MEJOR AJUSTE, QUE PERMITA ESTIMAR EL VOLUMEN DE ÁRBOLES DE TECA (*TECTONA GRANDIS L.F.*), EN PLANTACIONES DEL DEPARTAMENTO DE ALTA VERAPAZ, GUATEMALA" del estudiante Eric Abel Ortega Orellana del programa de Maestría En Estadística Aplicada identificado(a) con número de carnet 009517275.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.


Edwin Estuardo Vaides López
Ingeniero Agrónomo M.Sc.
Colegiado No. 2,556

Msc. Ing. Edwin Estuardo Vaides López

Colegiado No. 2556

Asesor de Tesis

ACTO QUE DEDICO A:

Dios	Por ser el creador y protector de la vida.
Mis padres	Daniel Ortega y Olga Orellana por su esfuerzo, apoyo y creer siempre en mí.
Mis hermanas	Olga, Klaudia y Melvy Ortega, por estar siempre pendientes de mí.
Mis hermanos	Aníbal, Adelzo y Otto Ortega, por sus consejos y muestras de cariño.
Mis abuelos	Abel Ortega (q. e. p. d.), Cristina Morales (q. e. p. d.), y Natalia Sánchez (q. e. p. d.), quienes desde pequeño me instruyeron y brindaron amor.
Mi esposa e hijos	Sara Berreondo, Daniel y Victoria Ortega, por ser mi inspiración diaria para luchar y alcanzar mis metas.

AGRADECIMIENTOS A:

Universidad de San Carlos de Guatemala	Mi <i>alma mater</i> , que le debo mi formación académica y don de servicio.
Facultad de Ingeniería	Por brindarme la oportunidad de formarme en el campo que me apasiona.
Co asesor	Msc. Ing. Carlos Ardón, por la orientación en los análisis estadísticos y por compartir sus conocimientos amplios en el tema.
Docente	Dra. Aura Marina Rodríguez, docente del curso de Seminario III, por su paciencia y apoyo en conformar el informe final de tesis.
Colaborador	Ing. Huber López, por apoyarme en la obtención de datos en campo.

ÍNDICE GENERAL

2.1.5.2.	Coeficiente de determinación múltiple....	13
2.1.5.3.	Coeficiente ajustado de determinación...	13
2.1.5.4.	Bondad de ajuste de ANDEVA y coeficientes.....	14
2.1.5.5.	Índice de condición.....	15
2.1.5.6.	Error estándar de estimación múltiple....	16
2.2.	Muestreo	17
2.2.1.	Métodos de muestreo.....	17
2.2.1.1.	Muestreo aleatorio simple.....	18
2.2.1.2	Muestreo aleatorio estratificado.....	18
2.2.1.3.	Muestreo sistemático.....	19
2.2.1.4.	Muestreo aleatorio por conglomerados..	20
2.2.2.	Tamaño de la muestra.....	20
2.2.3.	Selección de la muestra	20
2.3.	Descripción e importancia de la especie <i>Tectona grandis</i> L.f ...	21
2.4.	Tablas de volumen	22
2.4.1.	Tipos de tablas de volumen.....	22
2.4.1.1.	Tablas estándar.....	23
2.4.1.2.	Tablas locales de volúmenes (tarifas)....	23
3.	PRESENTACIÓN DE RESULTADOS.....	25
3.1.	Objetivo 1	25
3.1.1.	Diagrama de dispersión.....	25
3.1.2.	Coeficiente de correlación	29
3.2.	Objetivo 2	31
3.2.1.	Porcentaje promedio de troza comercial y leña.....	31
3.2.2.	Intervalos de confianza para las proporciones	34

3.3.	Objetivo 3: Establecer el modelo de mejor ajuste para predecir el volumen total del árbol de teca, y diseñar una tabla de volumen.....	36
3.3.1	Modelos evaluados para el volumen total.....	36
3.3.2.	Verificación de supuestos del modelo.....	39
3.3.2.1.	Prueba de significancia del modelo seleccionado.....	39
3.3.2.2.	Supuesto de normalidad del modelo seleccionado.....	40
3.3.2.3.	Homocedasticidad.....	41
3.3.2.4.	Independencia.....	41
3.3.2.5.	Multicolinealidad.....	42
3.3.3.	Diseño de la tabla de volumen total	42
3.3.4.	Comparación del modelo cuadrático con FAO	43
3.4.	Objetivo general.....	45
4.	DISCUSIÓN DE RESULTADOS	47
4.1.	Análisis interno	47
4.2.	Análisis externo	48
	CONCLUSIONES	51
	RECOMENDACIONES	53
	REFERENCIAS	55
	APENDÍCES	61

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Diagrama de dispersión tridimensional	26
2.	Diagrama de dispersión DAP respecto a volumen.....	27
3.	Diagrama de dispersión altura en función de volumen.....	28
4.	Diagrama de dispersión DAP respecto altura.....	29
5.	Matriz de correlaciones	30
6.	Proporción de productos por árbol de Teca	33
7.	Ajuste del modelo seleccionado	39
8.	Gráfico de comparación de volúmenes	44

TABLAS

I.	Operativización de variables	XXI
II.	Modelos matemáticos evaluados para la volumetría de teca.....	8
III.	Correlación de Pearson.....	30
IV.	Valores en porcentaje para troza y leña en Teca	31
V.	Valores en porcentaje para leña, troza menor y troza mayor.....	32
VI.	Intervalos de confianza para los productos de Teca	36
VII.	Indicadores de los modelos evaluados	37
VIII.	Prueba de significancia del modelo seleccionado	40
IX.	Prueba de Shapiro-Wilks (modificado)	40
X.	Prueba de Harrison-Mcabe	41
XI.	Prueba de Durbin Watson	41
XII.	Índice de condición	42

XIII.	Tabla de volumen total en m ³	42
XIV.	Prueba de medias volumen real y modelo FAO	45
XV.	Prueba de medias volumen real y modelo seleccionado.....	45

LISTA DE SÍMBOLOS

Símbolo	Significado
R^2	Coeficiente de determinación
R^2_{adj}	Coeficiente de determinación ajustado
β_i	Coeficientes del modelo matemático
H_1	Hipótesis alternativa
H_0	Hipótesis nula
k	Número de variables independientes o regresoras
$\%$	Porcentaje
Σ	Sumatoria
n	Tamaño de la muestra
\hat{y}	Valor estimado de la variable dependiente
y_i	Valor observado de la variable dependiente

GLOSARIO

ANDEVA	Análisis de varianza.
DAP	Diámetro a la altura al pecho a una altura del suelo sobre el fuste de 1.3 metros.
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura.
FUSTE	Tronco o tallo de los árboles desde la base hasta el ápice o punta, sin incluir las ramas.
INAB	Acrónimo de Instituto Nacional de Bosques.
Latifoliadas	Son bosques conformados por especies de árboles de hoja ancha.
Modelo parsimonioso	Son modelos simples con gran poder predictivo explicativo, que explican con un número mínimo de parámetros.
PINFOR	Acrónimo de Proyecto de Incentivos Forestales.
PPMF	Acrónimo de parcelas permanentes de medición forestal.

Telerelascópio

Es un instrumento que permite medir y estimar gran número de parámetros forestales, como la medida de distancias horizontales, determinación de área basal, medición de pendientes, entre otros.

RESUMEN

El propósito de la presente investigación fue brindar a los silvicultores un modelo matemático para la estimación de la volumetría del árbol de Teca, y su respectiva tabla de volumen para su uso confiable en el plan de manejo y aprovechamiento forestal de esa especie.

A nivel nacional no existe un modelo matemático específico que permita estimar de manera precisa el volumen en pie de árboles individuales de Teca, usualmente se ha utilizado la expresión $V = 0.108337266 + 0.000046499 * (D^2 * H)$, la cual es generalizada para todas las especies latifoliadas.

La investigación empleó un enfoque cuantitativo, con un diseño observacional, de tipo transversal y con alcance correlacional.

El resultado al que se llegó fue la selección del modelo: $Y = -0.3545 - (-)0.0444 * H + 0.000013 * d^2 * H$, que presenta el mejor ajuste. A partir de esta expresión se elaboró la tabla de volumen que facilitará la estimación volumétrica sin la realización de cálculos, solamente será necesario medir de forma directa el DAP y de forma indirecta la altura total del árbol.

Se infiere que la relación entre las variables DAP, altura total y volumen son de tipo no lineal, en cuanto a los productos del árbol de Teca, la proporción de leña corresponde al 10.08 % (2.05 % a 18.11 %), de troza menor al 54.24 % (40.95 % a 67.53 %) y de troza mayor al 35.68 % (22.90 % a 48.46 %), valores que varían según su clase diamétrica. El modelo seleccionado permitió la

estimación del volumen total del árbol, el cual no presenta diferencia significativa en comparación a los volúmenes reales obtenidos del trabajo de campo.

De acuerdo con los resultados obtenidos, se sugiere usar el modelo propuesto porque presenta el mejor ajuste, así como los valores de las proporciones estimadas de los productos para planificar el manejo y aprovechamiento de las plantaciones de Teca.

PLANTEAMIENTO DEL PROBLEMA

- Contexto general

La teca es nativa de cuatro países en el continente asiático: India, Myanmar, Tailandia y Laos. La madera que proviene de los bosques naturales corresponde a una extensión de 2,933,700 hectáreas, de la cual 16,517.70 hectáreas se encuentran en Myanmar (Camino y Morales, 2013).

Derivado de la prohibición que se hizo a partir del 1 de abril de 2014 en Myanmar de obtener madera desde bosques naturales, se generó una crisis en las maderas duras tropicales, lo cual vino a potencializar el establecimiento de plantaciones para abastecer el mercado, Kollert y Walotek (citados en Mollinedo, Herrera y Muñoz, 2016).

Según el Instituto Nacional de Bosques, INAB (2015), la Teca (*Tectona grandis* L.f) es la especie exótica más plantada a nivel nacional a través del Programa de Incentivos Forestales (PINFOR), debido a la alta demanda de la madera a nivel internacional y que los precios se han incrementado haciéndola rentable.

En Guatemala, según registros del programa PINFOR del INAB (2019), del año 1998 al año 2016 se han registrado 22,139.61 hectáreas plantadas con teca, de las cuales el 22 % han sido establecidas en el departamento de Alta Verapaz, representando una inversión total del Estado de Guatemala por Q 260,694,208.82 en concepto de pago en incentivos forestales.

Por su importancia se hace necesario conocer el crecimiento, desarrollo y productividad de las plantaciones para el plan de manejo forestal que conlleve a obtener los mejores resultados económicos y financieros.

- Descripción del problema

A nivel nacional no existe una ecuación específica que permita estimar de manera precisa el volumen en pie de árboles individuales de Teca, por lo que el Instituto Nacional de Bosques (INAB) para este fin ha recomendado inicialmente utilizar la ecuación de la Organización de las Naciones Unidas para la Agricultura y la Alimentación del año 1989, la cual es útil para estimaciones de especies latifoliadas en general, la que consiste en la expresión $V = (0.108337266 + 0.000046499 * (D^2 * H))$ que permite estimar el volumen (V) a partir del diámetro de altura al pecho en centímetros (D) y la altura en metros (H), lo cual tiene el inconveniente que es en general para especies latifoliadas.

Posteriormente el INAB, a partir de la red de parcelas permanentes de muestreo forestal, estableció algunos modelos matemáticos (INAB, 2015), que buscan hacer estimaciones del crecimiento y productividad de Teca, pero en la práctica son muy generales, y la estimación no es por árbol individual, si no en volumen por hectárea. Esta ecuación estima el volumen total (m^3/ha) mediante la expresión $EXP(1.768595 - 12.819398/T + 0.181328*S + 0.001175*N)$, en donde se ingresa la edad en años de la plantación (T), el índice de sitio (S), y el número de árboles por hectárea (N).

Por lo tanto, al hacer uso de las ecuaciones recomendadas por el INAB se produce una sobre estimación o subestimación de los volúmenes, aunado a esto para realizar la estimación del volumen de troza comercial y leña, se usa como criterio general, que, del volumen total, el 80 % del volumen corresponde a troza

comercial y la leña al 20 % del volumen total, lo cual en la práctica es variable según las clases diamatrales.

- Formulación del problema

- Pregunta central

¿Qué modelo matemático es el que mejor se ajusta para predecir el volumen en metros cúbicos en árboles de Teca, a partir de las variables diámetro a la altura al pecho (centímetros) y la altura total (metros)?

- Preguntas auxiliares

- ¿Cuál es la tendencia de las variables diámetro a la altura al pecho (cm) y la altura total (m) que influyen en la estimación del volumen del árbol de Teca?
 - ¿Cuál es el porcentaje promedio de leña, troza mayor y troza menor que se obtiene de un árbol de Teca?
 - ¿Cuál es el modelo de mejor ajuste para diseñar una tabla de volumen, que a nivel local permite planificar el aprovechamiento de la plantación de Teca?

- Delimitación del problema

El objeto de estudio del problema de investigación es la especie forestal Teca (*Tectona grandis L.f.*), específicamente plantaciones con fines industriales

establecidas mediante semillas sexuales provenientes de la fuente de Peñas Blancas del Centro Agronómico Tropical de Investigación y Enseñanza (CATIE).

Los árboles por medir y cubicar se ubican en plantaciones establecidas en el departamento de Alta Verapaz, Guatemala.

La gestión y recopilación de datos (DAP, altura, volumen troza comercial, y volumen leña) se realizó en los meses de octubre, noviembre y diciembre del año 2021.

OBJETIVOS

- General

Evaluar y seleccionar mediante análisis de regresión múltiple el modelo matemático de mejor ajuste para estimar el volumen de áboles de Teca (*Tectona grandis L.f*) en plantaciones del departamento de Alta Verapaz, Guatemala.

- Específicos

- Describir la tendencia de las variables que influyen en la estimación del volumen del árbol de teca, mediante diagramas de dispersión y el coeficiente de correlación de Pearson, para inferir si dicha tendencia es lineal o no lineal.
- Estimar el porcentaje promedio de leña, troza mayor y troza menor resultante de la corta de un árbol de Teca, mediante intervalos de confianza, para la planificación del manejo de la plantación.
- Establecer el modelo de mejor ajuste para predecir el volumen total del árbol de teca, y diseñar una tabla de volumen.

RESUMEN DEL MARCO METODOLÓGICO

A continuación, se describe la metodología de investigación realizada en el presente estudio, detallando el enfoque, diseño, tipo de estudio, alcances, variables, fases y los resultados obtenidos.

- **Enfoque**

El enfoque del estudio desarrollado es cuantitativo, porque se recolectaron y analizaron datos de tipo cuantitativo obtenidos de la medición de árboles de Teca, como el diámetro a la altura al pecho (cm), altura (m), proporción de productos y volumen (metros cúbicos), usándose el método hipotético-deductivo, trabajándose con diagramas de dispersión, modelos de regresión, prueba de hipótesis para los coeficientes de regresión, e intervalos de confianza, con el fin de modelizar el fenómeno.

- **Diseño**

El presente estudio es de carácter no experimental u observacional, esto porque no se usaron ensayos de laboratorio o diseñaron experimentos para originar los resultados, si no que se realizaron mediciones en los árboles de Teca plantados sin ninguna manipulación, con el fin de evaluar y seleccionar modelos para predecir el volumen, y diseñar tablas de volumen.

- Tipo de estudio

El presente estudio realizado es de tipo transversal porque la recolección de los datos para su posterior análisis se realizó en una sola ocasión usando una boleta, siendo de carácter destructivo porque se tumbó el árbol para realizar las mediciones dasométricas, base para la evaluación y selección del modelo de mejor ajuste, utilizando para ello la técnica de regresión múltiple.

- Alcance

La investigación es de tipo correlacional porque se estudió la relación entre variables independientes y dependiente, en primer lugar, mediante diagramas de dispersión para observar el patrón existente, y en segundo lugar mediante la técnica de regresión para evaluar y seleccionar el modelo matemático de mejor ajuste que permitió la predicción del volumen individual del árbol de Teca.

- Unidad de análisis

La unidad de análisis está compuesta por las plantaciones de Teca del departamento de Alta Verapaz (población), las que representan el 22 % a nivel nacional y presentan los diámetros de interés para el estudio.

La subpoblación la conforman las plantaciones del municipio de Chahal, Alta Verapaz, en donde por medio de observación directa se midieron variables dasométricas de los árboles, los que se seleccionaron según clase diamétrica.

El individuo estudiado es el árbol de Teca reproducido a partir de semilla sexual, al cual se le midió el DAP a 1.3 m sobre el suelo, fue cortado y en el suelo, se midió la altura total, se cubicó la madera y estimó la leña.

Para el método de muestreo, se realizó un muestreo selectivo por clase diamétrica dentro de las plantaciones donde se obtuvieron los individuos.

Las variables que fueron consideradas en el presente estudio se describen a continuación:

Tabla I. Operativización de variables

Variable	Definición teórica	Definición operativa
DAP	Diámetro a la altura del pecho-DAP- medida del diámetro del árbol en cm, a 1.3 metros a partir del suelo; tipo de variable cuantitativa continua con escala de medición de razón.	La medición se hizo con cinta métrica, previo a tumbar el árbol, la dimensional es centímetros, se consideró como variable independiente que sirvió para predecir el volumen de la madera.
H	Altura del árbol -H-, medida del árbol en metros desde su base hasta su ápice; tipo de variable cuantitativa continua con escala de medición de razón.	La medición del árbol se realizó ya tumbado en el suelo, usando para ello cinta métrica, la dimensional es metros lineales, se consideró como variable independiente para predecir el volumen de madera del árbol de teca.
V	Volumen de madera del árbol de teca -V-, medida de la cantidad de madera sólida en metros cúbicos; tipo de variable cuantitativa continua con escala de medición de razón.	Se tumbó el árbol, ya en el suelo se dividió el fuste en secciones de dos metros semejantes a trozas, para luego cubicarlas usando la fórmula de Smalian, se utilizó para ello cinta métrica y flexómetro.

Fuente: elaboración propia.

Para determinar el volumen de la troza individual, se usó la fórmula de Smalian, que es el método de cuantificación directo recomendado por el (INAB, 2019), que considera las dimensiones y largo.

$$V(m^3) = \frac{\pi}{4} * \left(\frac{D_1+D_2}{2}\right)^2 * L \quad (\text{Ec. 1})$$

Donde:

- V : volumen de la troza (m^3).
- π : constante pi, equivalente a 3.1416.
- D_1 : diámetro promedio del extremo menor de la troza (m).
- D_2 : diámetro promedio del extremo mayor de la troza (m).
- L : largo de la troza (m).
- Fases del estudio

Para el desarrollo de la presente investigación fue necesario trabajar en cuatro etapas que se describen a continuación:

- Fase 1: revisión documental

Esta fase del estudio se orientó a la revisión de literatura, dividiéndola en fuentes para antecedentes y para el marco teórico, que permitieron conocer diferentes perspectivas de abordaje de solución del problema a investigar, estructurar el marco teórico y el proceso metodológico a seguir. Las fuentes principales son los artículos científicos, revistas, reportes y libros.

- Fase 2: recolección de datos

En un período de tres meses después de la aprobación del protocolo de investigación se recolectaron los datos de campo, con base en el cronograma de actividades propuesto, y la metodología e instrumento diseñado. Para esta fase se consideró la gestión de licencia ante el INAB para el aprovechamiento de los árboles que fueron seleccionados en la muestra, el corte de los árboles representativos de cada clase diametral, mediciones dasométricas, y conformación de base de datos.

- Fase 3: análisis e interpretación de datos

Luego de la fase de campo se conformó una base de datos en donde se organizaron las variables DAP, altura, proporción de troza comercial, leña, y volumen de los árboles por clase diamétrica. Seguidamente se revisaron los datos para validarlos, seguidamente se elaboraron los diagramas de dispersión para observar las posibles tendencias lineales o no lineales, se practicó el análisis de regresión múltiple que permitió evaluar y seleccionar el modelo de mejor ajuste, se calculó la proporción de troza pequeña, troza grande y leña, y finalmente se construyeron los intervalos de confianza.

Para los análisis estadísticos se utilizó el software Curve Expert Professional, Infostat y R.

- Fase 4: informe final

Al haberse obtenido los resultados de los análisis considerados, se redactó el informe final, presentándolos en tablas y gráficos, para luego

interpretarlos y discutirlos, y finalmente concluir y recomendar, esto con el fin de proponer alternativas de solución a la problemática analizada.

- Resultados obtenidos

Con el análisis de las variables que fueron consideradas en el presente estudio, se definió el patrón o tendencia de las variables independientes en conjunto con la variable dependiente, usando para ello el diagrama de dispersión, así como el modelo de regresión múltiple con mejor ajuste para predecir el volumen por árbol, la estimación de la proporción puntual y por intervalos de los productos del árbol de Teca, para que al final se alcanzara construir una tabla de volumen de doble entrada, para predecir el volumen por árbol sin necesidad de realizar cálculos.

Paralelo al análisis de las variables, a los residuos se les verificó los supuestos de normalidad, independencia, multicolinealidad y homogeneidad de varianzas para que las pruebas paramétricas usadas generaran conclusiones confiables.

- Población y muestra

La población estudiada consiste en reforestaciones de la especie forestal teca, establecidas con fines industriales a partir de semillas sexuales, plantaciones que están ubicadas en el departamento de Alta Verapaz. En la plantación se seleccionó de manera selectiva los árboles según las clases diamétricas establecidas. La muestra consistió en por lo menos 3 árboles por cada clase diamétrica, desde los 20 hasta los 50 centímetros de diámetro a la altura al pecho, totalizando una cantidad de 54 árboles considerados para el presente estudio.

INTRODUCCIÓN

Este estudio consistió en una innovación, en vista que a nivel país no existe un modelo matemático que permita predecir el volumen individual del árbol de Teca con exactitud, y tampoco una tabla de volumen que facilite conocer de manera fácil e inmediata la volumetría a partir de las variables diámetro a la altura al pecho y altura en metros, siendo indispensable para el manejo y aprovechamiento eficiente de las plantaciones, ya que esto facilita el plan de manejo a plantear ante el INAB para la gestión y emisión de la licencia de aprovechamiento, además contar con el volumen real para la negociación y comercialización de la madera.

En este orden de ideas la presente investigación se enfocó en la línea de pronósticos, ya que evaluó quince modelos matemáticos que se usan en otros países para predecir volumetría a partir de las variables diámetro a la altura al pecho (DAP) y altura en metros, para seleccionar el modelo de mejor ajuste que permita estimar en árboles individuales de teca el volumen total, y diseñar una tabla de volumen que permita en base al DAP y la altura contar con la volumetría para el árbol individual.

El problema radicó en que no existe una ecuación específica que permita estimar de manera precisa el volumen en pie de árboles individuales de teca, por lo que el Instituto Nacional de Bosques (INAB) para este fin ha recomendado inicialmente utilizar la ecuación de la Organización de las Naciones Unidas para la Agricultura y la Alimentación del año 1989, la cual es útil solamente para estimaciones de especies latifoliadas en general, la que consiste en la expresión

$V = 0.108337266 + 0.000046499 * (D^2 * H)$), que permite estimar el volumen (V) a partir del diámetro de altura al pecho en centímetros (D) y la altura en metros (H).

La importancia de realizar este estudio radicó en la necesidad de contar con un modelo matemático que permita pronósticos precisos de la volumetría del árbol individual de Teca a partir de variables fáciles de medir como el diámetro a la altura al pecho y la altura, o de contar con una tabla de volumen que permita sin hacer nuevos cálculos obtener el volumen de manera inmediata.

El esquema de solución se enfocó en seleccionar árboles de Teca representativos por clase diamétrica para obtener las variables diámetro a la altura al pecho, altura en metros, volumen, proporción de troza menor, troza mayor, y leña, para realizar análisis de regresión múltiple con el objetivo de evaluar y seleccionar el modelo matemático de mejor ajuste para predecir el volumen, y con este diseñar una tabla de volumen.

Se contó con todos los recursos para la realización del presente estudio, así como con los permisos para colectar los datos en la plantación de Teca.

El presente informe está dividido en cuatro capítulos:

- El primero consiste en el marco referencial, en donde se describen estudios realizados en otros países sobre la evaluación y selección de modelos matemáticos para la predicción de volumetría en Teca, experiencias que sirvieron para realizar la presente investigación.
- El segundo capítulo, corresponde al marco teórico, donde se realizó una constante revisión documental del problema a solucionar y las técnicas estadísticas a utilizar para el aporte en la solución.

- En el tercer capítulo, se presentan los resultados obtenidos del estudio de investigación para cada uno de los objetivos planteados, se describe la propuesta de solución, se verifican los supuestos, y se hace la comparación con el modelo que se usa actualmente para demostrar que el modelo propuesto es el que predice con mejor ajuste.
- En el cuarto capítulo, se presenta la discusión de los resultados considerando tanto la validez interna como externa, especificando las limitantes o problemas que se observaron al desarrollar el estudio y comparando con otras investigaciones realizadas en el ámbito internacional. Aunado a la discusión de resultados se presentan las conclusiones y recomendaciones coherentes a los objetivos alcanzados con el presente estudio.

1. MARCO REFERENCIAL

En relación a estudios sobre modelos matemáticos para estimar el volumen total y por productos, a nivel nacional se ha limitado nada más al modelo recomendado por el INAB, mediante el cual se puede predecir el volumen total por hectárea, y no así al volumen individual del árbol y sus productos (troza menor, troza mayor y leña), y a nivel de Mesoamérica se cuentan con algunos reportes en Venezuela y México, por lo que en este capítulo de antecedentes, se hace énfasis en investigaciones que aportaron para el presente estudio:

Moret, Jérez, y Mora (1998), realizaron en Venezuela el ajuste de quince modelos de regresión, con datos que provenían de 174 árboles seleccionados a partir de un muestreo estratificado por clases diamétricas, para obtener las ecuaciones de volumen. La evaluación y selección de los mejores modelos lo obtuvieron usando valores del cuadrado medio del error y del índice de Furnival, seguidamente validaron los modelos usando una muestra de 39 árboles, realizando para ello un análisis de varianza, probando si existían o no significancia entre el volumen real y el volumen estimado por los modelos ajustados, aunado a esto se usó el método de los predichos, suma de cuadrados del error y el coeficiente de determinación ajustado. El modelo de mejor ajuste fue el conocido como de Shumacher-Hall 1933, con mayor coeficiente de determinación de predicción y menor índice de Furnival.

Del anterior estudio fueron de utilidad los modelos evaluados y los criterios, se seleccionaron 5 modelos (Shumacher-Hall, Variable combinada logarítmica, semilogarítmico, Meyer modificado y Spurr) que consideran el diámetro a la altura al pecho y la altura para la predicción del volumen del árbol

individual, además los criterios (cuadrado medio del error y el coeficiente de determinación ajustado) y validarlos (normalidad, multicolinealidad, homocedasticidad e independencia).

Telles, Cárdenas, Rodríguez, Calderón y Pérez (2018), plantearon un modelo estadístico para pronosticar el volumen del fuste de una plantación de *Tectona grandis* L. f en Michoacán, considerando una muestra de árboles representativos a los cuales se les midió el diámetro del tocón, diámetro a la altura al pecho y a distintas secciones del árbol. Se hizo la evaluación de seis modelos (Burkhart, Australiana, Stoat, Meyer, Prodan, y Naslud) usando el software estadístico SAS, seleccionándolos en base a los criterios de la suma de cuadrados del error (SCE), la raíz del error medio cuadrático (REMC) y el coeficiente de determinación ajustados (R^2_{adj}), además de haber evaluado significancia en los parámetros. De los modelos evaluados, el de Meyer presentó la mejor bondad de ajuste, y con la totalidad de los parámetros significativos.

De este estudio se seleccionaron los seis modelos evaluados para predicción volumétrica, se consideró el criterio de muestreo para seleccionar árboles por clase diamétrica, la metodología para cubicación de los fustes de los árboles, los criterios para selección del modelo de mejor ajuste y su validación.

Asimismo, Barrantes, Murillo, Ávila, Fonseca, y Barquero (2021), indican que es necesario predecir con precisión el volumen de madera en reforestaciones clonales de *Tectona grandis* L.f, para facilitar el manejo y aprovechamiento forestal, valiéndose para ello de tablas de volumen resultantes de modelos matemáticos de mejor ajuste, para distintas calidades de sitio. El volumen total del árbol se calculó mediante la fórmula de Smalian, sumando cada sección cubicada. Por cada índice de sitio, se ajustaron los modelos usando el método de mínimos cuadrados ordinarios, valiéndose del software estadístico

Statgraphics Centurión XVI. Los modelos se seleccionaron según los criterios siguientes: buen ajuste del coeficiente de determinación, y bajos errores de estimación, usando el diámetro a la altura al pecho como variable regresora.

De este estudio fue de utilidad la metodología de Smalian para estimar el volumen de cada segmento del árbol, el método de ajuste de los modelos, criterio de evaluación y validación.

Por su parte, Rueda, *et al.*, (2020) manifiestan que el objetivo de la investigación fue predecir con modelos matemáticos el volumen del fuste y elaborar tablas de volumen para *Tectona grandis* L.f, valiéndose para ello de una muestra de 32 árboles por plantación, a los que se les midió variables dasométricas, usando telerelascópio. Se probaron cuatro modelos matemáticos usando el software estadístico SAS 9.2®. Para seleccionar el mejor modelo se consideró la suma de cuadrados del error, la raíz del error medio cuadrático y el coeficiente de determinación ajustado, además, si eran significativos los parámetros. El modelo de la variable combinada fue el de mejor ajuste, en comparación de los modelos Schumacher, Korsun y Thomber, y con este se elaboraron tablas de volumen para cada una de las localidades.

De este estudio se consideraron los cuatro modelos evaluados (Schumacher, variable combinada, Korsun, y Thomber), para determinar el que presenta mejor ajuste para la predicción volumétrica en la especie forestal teca, además los criterios de evaluación (cuadrado medio del error, suma de cuadrados del error significancia de los parámetros y la raíz del error medio cuadrático) y la validación (homocedasticidad y autocorrelación) mediante la verificación de los supuestos.

Tamarit, et al., (2019) comentan que, en México, la especie forestal Teca (*Tectona grandis* L. f) es importante por su potencial maderable, por ese motivo se diseñó un sistema de crecimiento y rendimiento maderable (Scrm) para plantaciones. Se constituyó de ecuaciones dinámicas ajustadas que usan variables dasométricas de una red de parcelas permanentes. La selección se fundamentó en una evaluación gráfica para comparar similitudes entre el patrón resultante, la tendencia observada, y en un análisis de bondad de ajuste, buscándose un equilibrio entre lo estadístico y lo biológico. Se eligieron los modelos 13 y 17 de los sistemas S4 y S6, respectivamente, que son tipo Schumacher para el volumen y el área basal, y entre ambos, se eligió finalmente el S4, que es el mejor para predecir y proyectar el crecimiento en volumen.

Del estudio antes descrito se consideraron los criterios de bondad de ajuste usados para evaluar el modelo para la predicción del volumen, siendo el coeficiente de determinación ajustado (R^2_{adj}), la raíz del cuadrado medio del error (RCME), y el criterio de información de Akaike (AIC); para validación se verificó la homocedasticidad mediante el método gráfico de residuales en comparación de los predichos y autocorrelación mediante la prueba de Durbin-Watson.

2. MARCO TEÓRICO

En el marco teórico se refleja el análisis documental realizado sobre la temática estadística y técnica desarrollada en el presente trabajo de investigación, útil para la metodología y discusión de los resultados:

2.1. Análisis de regresión múltiple

Es una herramienta estadística que se usa para el estudio de la interdependencia entre variables, que se centra en estudiar la relación entre variables, con el fin de hacer predicciones de valores de una o más variables dependientes a partir de un grupo de variables regresoras. Este análisis supone que la variable independiente no es aleatoria, o sea que se puede manejar por el investigador. La interdependencia entre variables es expresada por medio de relaciones o funciones matemáticas. Por medio de estas funciones se puede conocer cuánto cambia la variable dependiente (y) por unidad de cambio en la variable independiente (x), y a la vez pronosticarse el valor de y que está asociado a un valor específico de x .

La importancia de las funciones matemáticas radica en que a partir de una o más variables fáciles de medir se puede estimar otra variable que es difícil de medir, de alto costo o imposible de medir en condiciones normales de muestreo, esto permite incrementar la eficiencia y reducir costos de las estimaciones.

El problema básico en el análisis de regresión radica en estimar con precisión los coeficientes de la ecuación de regresión, a partir de datos contenidos en una muestra, y generalmente este se acompaña del análisis de

correlación que se ocupa de determinar la intensidad de la asociación entre variables, las que se suponen todas son aleatorias (Cancino, 2006).

2.1.1. Método de mínimos cuadrados

El método es un procedimiento en el que se utilizan los datos obtenidos de una muestra representativa para obtener la ecuación de la regresión estimada, o sea los valores de los parámetros de regresión ($b_0, b_1, b_2, b_3 \dots b_p$) que minimicen la suma de cuadrados de las desviaciones entre los valores observados de la variable dependiente y_i y los valores estimados de la variable dependiente \hat{y}_i . El criterio que emplea el método se basa en la expresión:

$$\min \varepsilon (y_i - \hat{y}_i)^2 \quad (\text{Ec. 2})$$

Donde:

- y_i : valor observado de la variable dependiente de la observación i-ésima.
- \hat{y}_i : valor estimado de la variable dependiente de la observación i-ésima.

Los valores estimados de la variable dependiente se determinan con la ecuación de regresión estimada:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p \quad (\text{Ec. 3})$$

En la regresión múltiple, las fórmulas para los coeficientes de regresión ($b_0, b_1, b_2, \dots b_p$) utilizan álgebra matricial, lo que complica el uso de fórmulas para su cálculo, siendo mejor el uso de software estadístico para obtener la ecuación de regresión estimada (Anderson, Sweeney, y Williams, 2012).

Para el caso del método de los mínimos cuadrados ponderados, un peso positivo se asocia a cada dato, y este se define por una relación que es proporcional con el inverso de la varianza de los residuos. La relación se simplifica para la biomasa o el volumen, y luego se traduce en un exponente adicional comparado con el método de mínimos cuadrados ordinarios.

El método de máxima verosimilitud consiste en la determinación de los argumentos que maximizan en las observaciones la verosimilitud. La variable respuesta se distribuye normalmente, por la cual la esperanza y la desviación estándar depende de los parámetros (θ , κ , c). El método consiste en encontrar los tres parámetros que maximizan la función. Para el caso de los modelos no lineales, una optimización se hace necesaria para comprobar que se obtenga el mínimo, o el máximo, respectivamente (Picard, Saint-André y Henry, 2012).

2.1.2. Modelos

Para el caso de los modelos no lineales anidados, la selección permite probar si los parámetros adicionales del modelo son significativos o no, en donde la hipótesis nula se rechaza si el valor de probabilidad de la prueba de Fischer es menor que el nivel de significancia. En los modelos que presenten la misma variable respuesta, el mejor será el que tenga el menor valor del criterio de información de Akaike (AIC) (Picard *et al.*, 2012).

Los modelos de volumen en árboles presentan modelos lineales y no lineales en los coeficientes, en cuanto a los primeros, se ajustan por medio del método de mínimos cuadrados ponderados u ordinarios, lo cual depende del comportamiento de la varianza, y en el caso del segundo tipo, se requiere de métodos de ajuste no lineales.

Depende de la estructura del modelo que se utiliza para el ajuste de la relación entre el volumen y otras variables, las funciones se clasifican en: a) funciones de volumen local, que incluyen únicamente como variable regresora al DAP; b) funciones de volumen general, que incluyen las variables regresoras diámetro a la altura al pecho y la altura; y c) funciones de volumen con clase de forma, que incluyen el diámetro a la altura del pecho, el factor de forma y la altura como variables regresoras (Cancino, 2006).

2.1.2.1. Modelos matemáticos evaluados

Diferentes modelos matemáticos han sido propuestos en otros países para la estimación de la volumetría en la especie *Tectona grandis* L.f, para esta investigación se han seleccionado algunos, de los que se han evaluado en Venezuela y México, que se presentan a continuación:

Tabla II. Modelos matemáticos evaluados para la volumetría de teca

Núm.	Modelo	Expresión matemática
1	Burkhart (1977)	$V = \beta_0 + \beta_1.d^{\beta_2} \cdot H^{\beta_3} + \varepsilon$
2	Australiana	$V = \beta_0 + \beta_1.d^2 + \beta_2.H + \beta_3.d^2.H$
3	Stoate (1945)	$V = \beta_0 + \beta_1.d^2 + \beta_2.d^3 + \beta_3.H + \beta_4/H + \varepsilon$
4	Meyer (1941)	$V = \beta_0 + \beta_1.d + \beta_2.d^2 + \beta_3.d.H + \varepsilon$
5	Prodan, Peters, Cox y Real (1997)	$V = \beta_0 + \beta_1.d^2 \cdot H + \beta_2.d^2 + \varepsilon$
6	Naslud (1936)	$V = \beta_1 + \beta_2.d^2 + \beta_3.d^2.H + \beta_4.H^2 + \beta_5.d.H^2$
7	Schumacher	$V = \beta_1.d^{\beta_2}.H^{\beta_3} + \varepsilon$
8	Variable combinada	$V = \beta_1(d^2.H)^{\beta_2} + \varepsilon$
9	Korsun	$V = \beta_1(d+1)^{\beta_2}.H^{\beta_3} + \varepsilon$
10	Thornber	$V = \beta_1(H/d)^{\beta_2}.d^2 \cdot H + \varepsilon$
11	Spurr	$V = \beta_0 + \beta_1(d^2.H)$
12		$V = \beta_0 + \beta_1.d.H + \beta_2.d^2$
13	Cuadrático	$V = \beta_0 \cdot \beta_1.H + \beta_2.d^2.H$
14	Semilogarítmico	$\ln V = \beta_0 + \beta_1 d^2 + \beta_2 H$

Continuación de la tabla II.

núm.	Modelo	Expresión matemática
15	Meyer modificado	$V = \beta_0 + \beta_1.d + \beta_2.(dH) + \beta_3.d^2 + \beta_4(d^2.H)$

Donde: V: es el volumen; d: el diámetro a la altura de 1.3 m; H: la altura total (m); β_i : parámetros a ser estimados; ln: logaritmo neperiano (2.71828) y ε : error aleatorio del modelo.

Fuente: elaboración propia.

2.1.3. Diagrama de dispersión

El patrón de tendencia en la relación entre el volumen y otras variables del árbol se puede observar fácilmente usando el análisis gráfico Cancino (2006), como en un diagrama de dispersión, el cual consiste en una técnica útil para entender la relación entre variables (Lind, Marchal, y Wathen, 2012).

Generalmente la asociación entre el volumen y variables como el DAP, altura y factor de forma, es de tipo no lineal, y las variables que resultan de las variables simples facilitan linealizarlas, reduciendo la varianza de los residuos del volumen (Cancino, 2006).

2.1.4. Supuestos del análisis de regresión

El análisis de los errores es la principal herramienta para evaluar si el modelo de regresión a emplear es el apropiado. El residuo de la observación i , se obtiene de la diferencia entre el valor observado de la variable dependiente (y) y el predicho de la variable dependiente (\hat{y}). En otras palabras, el residuo i ésimo es el error que resulta al usar el modelo de regresión estimado, para la predicción (Anderson *et al.*, 2012).

Después de ajustarse el modelo de regresión, es útil examinar los residuos para averiguar cómo se ajusta el modelo a los datos y los supuestos de regresión (Newbold, Carlson y Thorne, 2008).

2.1.4.1. Normalidad

El supuesto de normalidad es fundamental para todos los contrastes de hipótesis e intervalos de confianza asociados a la regresión normal múltiple, pudiéndose utilizar para verificar la normalidad métodos gráficos como los QQ-plots, en donde el ajuste a la normal ocurre cuando los puntos se colocan aleatoriamente alrededor de una recta. También se puede utilizar la prueba de Shapiro-Wilks, y Lilliefors como prueba de bondad de ajuste de la normalidad (Corral y López, 2019).

Según López *et al.*, (2010), al no cumplirse el supuesto, todas las pruebas estadísticas desarrolladas no serán válidas, esto porque la prueba de F y prueba de t lo condicionan.

2.1.4.2. Independencia

La independencia de observaciones implica que los residuales no tienen un patrón, que no están correlacionados entre sí, y que no existen corridas largas de residuos positivos o negativos. Cuando los errores o residuos están asociados, se dice que se presenta autocorrelación. La autocorrelación ocurre frecuentemente cuando los datos son colectados en un período de tiempo. Para verificar este supuesto existe una prueba denominada Durbin-Watson (Lind *et al.*, 2012).

2.1.4.3. Homocedasticidad

Es común detectar falta de homocedasticidad mediante la gráfica de residuos, una varianza grande del error produce residuos grandes (Walpole, Myers, R., Myers, S. y Ye, 2012).

Una variación constante con relación al modelo de regresión se denomina homocedasticidad, lo que significa que la variación de los valores de predicción es constante, no importando si los valores de predicción sean grandes o pequeños (Lind *et al.*, 2012).

2.1.4.4. Multicolinealidad

En el caso de la regresión múltiple, conforme aumenta el grado de correlación entre las variables independientes, los parámetros de regresión, regularmente se vuelven menos confiables. Si se presenta un alto nivel de correlación entre las variables independientes, implica que se tiene problema de multicolinealidad.

Con la multicolinealidad, al agregar una segunda variable que esté asociada con la primera, se distorsionan los valores de los coeficientes de regresión, por lo que, como sugerencia, el mejor modelo de regresión múltiple es el que explica la asociación entre conjunto de datos, con el menor número de variables independientes (Levin y Rubín, 2010).

Lind *et al.* (2012), indican que las variables regresoras que están correlacionadas complican las inferencias sobre los parámetros de regresión individuales, y sus efectos sobre la variable dependiente.

2.1.5. Evaluación y selección de modelos

Para evaluar y seleccionar los modelos se utilizaron los siguientes criterios:

2.1.5.1. Criterio de Akaike (AIC) y de Schwarz (BIC)

Estos dos criterios tienen una estructura muy similar, y a cambio de los coeficientes de determinación, estos indican mejores ajustes cuanto más bajos sean sus valores.

El estadístico AIC tiene la siguiente expresión:

$$AIC = -\frac{2l}{n} + \frac{2k}{n} \quad (\text{Ec. 4})$$

Donde l es el logaritmo de la función de verosimilitud evaluada para los valores estimados de los coeficientes.

El estadístico SC propuesto por Schwarz, tiene la siguiente expresión:

$$SC = -\frac{2l}{n} + \frac{k \ln(n)}{n} \quad (\text{Ec. 5})$$

Es importante destacar que: los estadísticos AIC y SC penalizan la introducción de nuevos regresores, no son medidas relativas como lo son los coeficientes de determinación, por lo tanto, su magnitud no ofrece ninguna información, se pueden utilizar para contrastar modelos en los que las variables externas tienen diferentes formas funcionales (Uriel, 2013).

2.1.5.2. Coeficiente de determinación múltiple

En el tema de regresión múltiple, el coeficiente es el porcentaje de variación de la variable explicada que se explica por las variables regresoras.

Este coeficiente se puede determinar a partir del análisis de varianza, como una razón de la suma de cuadrados de la regresión (SSR) entre la suma de cuadrados total (SST).

$$R^2 = \frac{SSR}{SS\ Total} \quad (\text{Ec. 6})$$

Las características principales del coeficiente de determinación múltiple son: a) se representa como R^2 , b) sus valores oscilan entre 0 a 1, valores cercanos a 0 implica poca asociación entre las variables independientes y la variable dependiente, un valor cercano a 1 significa una relación fuerte y c) no debe poseer valores negativos (Lind *et al.*, 2012).

2.1.5.3. Coeficiente ajustado de determinación

Según Lind *et al.* (2012) el coeficiente de determinación múltiple se incrementa con el número de variables regresoras en la ecuación de regresión múltiple. Cada variable independiente nueva hace que las predicciones sean más precisas, lo que reduce la sumatoria cuadrada del error (SSE) y aumenta la sumatoria cuadrada de la regresión (SSR).

El coeficiente se incrementa debido solamente al número de variables regresoras no es porque la variable que se agrega, aporte en predecir la variable dependiente. Para equilibrar el efecto de la cantidad de variables regresoras en

el coeficiente de determinación, se considera un coeficiente de determinación ajustado múltiple, que se calcula con la fórmula siguiente:

$$R^2 \text{ajustado} = \frac{\frac{SSE}{n-(k+1)}}{\frac{SS \text{ total}}{n-1}} \quad (\text{Ec.7})$$

Donde:

- SSE : suma de cuadrados del error.
- n : tamaño de muestra.
- k : número de variables independientes.
- SS total : Suma de cuadrados del total.

2.1.5.4. Bondad de ajuste de ANDEVA y coeficientes

Usando la distribución de F en el análisis de varianza-ANDEVA- como una prueba global del modelo, se puede probar si los parámetros de regresión en la población tienen un valor igual a cero, para ello se hace necesario primero formular las hipótesis nula y alterna. La hipótesis nula es: $H_0: \beta_1 = \beta_2 = \beta_3 = 0$, y la hipótesis alterna es: $H_1: \text{no todas las } \beta_i \text{ son igual a 0}$. Entonces si la hipótesis nula es verdadera, implica que la totalidad de los coeficientes tienen un valor de cero y, no son de utilidad para la estimación de la variable dependiente.

Al obtenerse significancia en la prueba global del análisis de varianza, implica que por lo menos uno de los parámetros de regresión es diferente a cero, lo que los hace significativos para las predicciones. El siguiente paso es contrastar las variables regresoras de manera individual para determinar que parámetros de regresión pueden tener un valor de cero y cuáles no. En esta prueba de significancia se usa como estadístico el valor de t de student con $n - k - 1$ grados de libertad.

(k+1) grados de libertad con el correspondiente valor del nivel de significancia (Lind *et al.*, 2012).

2.1.5.5. Índice de condición

La multicolinealidad entre las variables se puede evaluar por medio del índice de condición máximo (IC), que consiste en la raíz cuadrada del cociente entre los autovalores mayor y menor de la matriz de correlaciones de las variables regresoras, luego que hayan sido centradas y estandarizadas. Este procedimiento se considera que es el más recomendado de los disponibles en la actualidad.

El número de condición, $IC_{\max}(X)$, es igual a la raíz cuadrada de la razón entre la raíz característica más grande (λ_{\max}) y la raíz característica más pequeña (λ_{\min}) de la matriz $X'X$. Como esta matriz es de dimensión k * k resultan k raíces características, pudiéndose calcular para cada una de ellas un índice de condición (IC_i). (Jimenes, 2019, p. 200)

$$IC_{\max}(x) = \sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} \quad (\text{Ec. 8})$$

$$IC(\lambda_i) = \sqrt{\frac{\lambda_{\max}}{\lambda_i}} \quad (\text{Ec. 9})$$

Donde:

- λ : es la raíz característica, autovalor, valor propio o *eigenvalues* de una matriz.

El número de condición mide la sensibilidad de las estimaciones mínimos cuadráticos ante pequeños cambios en los datos. Valores menores a 10 significa que no hay colinealidad o es débil, valores de 10 a 30 significa la existencia de colinealidad moderada, y valores mayores a 30 significa que existe colinealidad severa. (Jimenes, 2019, p. 201)

2.1.5.6. Error estándar de estimación múltiple

El error estándar de la estimación múltiple mide que tan cercanos están los valores reales y los valores estimados con el modelo, e indica si un modelo matemático para predicción es más eficiente.

En el análisis de regresión se generan dos estadísticos que permiten evaluar la capacidad de un modelo de regresión para predecir una variable dependiente, el error estándar y el coeficiente de determinación.

Este estadístico se calcula al relacionar la suma de cuadrados de la regresión entre la resta del número de observaciones de la muestra con el número de parámetros del modelo, y obtener la raíz cuadrada, tal como se observa en la fórmula siguiente:

$$Sy. 123 \dots k = \sqrt{\frac{SSR}{n-(k+1)}} \quad (\text{Ec.10})$$

Donde:

- n : es el número de observaciones de la muestra.
- k : es el número de variables regresoras.
- SSR : es la suma de los cuadrados de la regresión de la tabla ANDEVA.

2.2. Muestreo

En la forestería existe una relación entre magnitudes que son fáciles de medir como el diámetro, altura, y densidad de la madera, con variables que son difíciles de medir, como por ejemplo la biomasa o el volumen, a lo cual técnicamente se le conoce como alometría, la cual consiste en estudiar las relaciones entre las características de los individuos de una población. Las mediciones onerosas y destructivas se pueden limitar a una muestra de árboles representativos, con el objetivo de determinar estadísticamente los coeficientes de regresión del conjunto de árboles de la población.

En el proceso de modelación estadística de volumen o biomasa existen dos fuentes de error: el primero, se asocia con que solo algunos árboles se miden de toda la población, a esto se le define como error de muestreo. El segundo, está relacionado con que el modelo resultante siempre permitirá solamente una aproximación de la biomasa o el volumen, siendo esto conocido como el error de predicción (Picard *et al.*, 2012).

2.2.1. Métodos de muestreo

Generalmente cuando se estudian las características de una población existen muchas razones prácticas para preferir algunas partes o muestras de esta para observar y medir, entre algunas razones son las siguientes: levantar la información de la población requeriría mucho tiempo, el costo por estudiar el universo de los elementos resultaría prohibitivo, es imposible verificar físicamente la totalidad de los individuos de la población, en algunas ocasiones las pruebas son de naturaleza destructiva y se ha determinado que los resultados muéstrales son adecuados.

2.2.1.1. Muestreo aleatorio simple

Este muestreo es el más común, y la muestra es elegida de manera que cada elemento o individuo de la población tiene las mismas posibilidades de que se le tome en cuenta (Lind *et al.*, 2012).

En la práctica, este muestreo se realiza enumerando las unidades de 1 a N, y posteriormente, se extraen n números aleatorios entre 1 y N. En cada extracción, el proceso debe de facilitar la misma oportunidad de selección para todos los elementos que no hayan salido.

Entre las ventajas de este muestreo es que es sencillo y de fácil comprensión, el cálculo de los estimadores es rápido, y existen muchos programas para realizar el análisis.

Entre las desventajas del muestreo es que requiere que se cuente con antelación el listado de toda la población (marco de muestreo), y si se trabaja con muestras menores a 30 individuos, se puede dar la posibilidad que no representen a la población de la mejor manera (López y González, 2017).

2.2.1.2. Muestreo aleatorio estratificado

Una muestra aleatoria estratificada se obtiene mediante la separación de los elementos que conforman la población en grupos que no presenten traslapos, que se les denomina estratos, y dentro de cada estrato se selecciona posteriormente una muestra aleatoria simple.

Una muestra debe ser elegida porque la estimación final puede tener un límite para el error de estimación mucho más pequeño, el costo de todo el

proceso de campo puede ser minimizado al seleccionar áreas geográficas compactas, y tal ahorro permite utilizar un tamaño de muestra mayor, lo que permite estimar individualmente los parámetros dentro de cada uno de los estratos.

Para seleccionar la muestra mediante este método se debe especificar claramente los estratos, seleccionar el tamaño de muestra idóneo para cada estrato, y seleccionar una muestra simple aleatoria para cada estrato.

Las muestras obtenidas presentan la ventaja de que cuando se diseñan técnicamente, representan adecuadamente las características de la población de la cual fueron extraídas, al compararlas con otro tipo de muestras (López y González, 2017).

2.2.1.3. Muestreo sistemático

Una muestra resultante al seleccionar aleatoriamente un individuo de los primeros k individuos en el marco de muestreo y después cada k -ésimo individuo es llamada como muestra sistemática de 1 en k . En este muestreo los elementos son seleccionados de la población dentro de un intervalo uniforme que se mide con respecto al tiempo, al orden o al espacio.

Este muestreo presenta la ventaja de ser más fácil de realizar en campo porque a cambio de las muestras simples aleatorias y estratificadas, se expone menos a los errores de selección que cometen los investigadores de campo.

La idea básica es que una muestra de n elementos se selecciona de una larga lista, para ello se elige en un intervalo apropiado, y selecciona los elementos a intervalos iguales a través de la lista (López y González, 2017).

2.2.1.4. Muestreo aleatorio por conglomerados

En este muestreo la unidad muestral es un grupo de elementos de la población que forman una unidad, denominados conglomerados. Este muestreo consiste en elegir de manera aleatoria un número de conglomerados y analizar posteriormente los elementos que los conformen.

Este muestreo es adecuado cuando los grupos en que se divide la población son muy similares entre sí, por lo tanto, no hay gran diferencia entre estudiar individuos de un grupo o de otro (López y González, 2017).

2.2.2. Tamaño de la muestra

Entre más variabilidad presente la población forestal, más grande debe ser el número de árboles a seleccionar mediante la técnica de muestreo. Una plantación industrial requiere un tamaño de muestreo más pequeño que un bosque natural. Para el primer caso los árboles son más parecidos entre ellos, porque tienen la misma edad, y son de la misma especie o del mismo material clonal. El tamaño de la muestra debe ser acorde a la heterogeneidad de los árboles en la plantación, a mayor heterogeneidad mayor debe ser el tamaño de muestra, y por el contrario a menor variabilidad menor el tamaño de la muestra (Picard *et al.*, 2012).

2.2.3. Selección de la muestra

La técnica de selección de los árboles en la población también tiene que ver con el error de muestreo, ya que, para la regresión, la precisión se incrementa al aumentar el tamaño de la muestra, siendo recomendable seleccionar árboles en todo el rango de la variable de entrada. En el caso de que la variable de entrada

elegida sea el diámetro, se prefiere seleccionar un número constante de árboles por clase diamétrica, para tener representación de los árboles con mayores diámetros (Picard *et al.*, 2012).

2.3. Descripción e importancia de la especie *Tectona grandis* L.f

La Teca es una especie latifoliada que pertenece a la familia botánica Verbenácea, originaria del sureste asiático, que incluye los países: Myanmar, Tailandia, India, Malasia, Java, Indochina y Laos. Es un árbol deciduo, que puede alcanzar hasta los 50 metros o más, y 2 metros de diámetro en su localidad de origen (Fonafifo, 2003).

Madeiros *et al.*, (citado en Barrantes, et al., 2021) indica que la Teca (*Tectona grandis* L.f) es una especie forestal de las más importantes a nivel mundial, con cerca de 6.8 millones de hectáreas cultivadas.

En Guatemala es la especie exótica más cultivada según registros del Programa de Incentivos Forestales (PINFOR), esto porque la demanda de la madera ha aumentado con el tiempo, y los precios de venta, la hace rentable. En el período de 1997 al año 2013 se tenía registrado un total de 18,222.36 hectáreas plantadas con teca, representando en incentivos forestales, pagados por el Estado de Q 211,238,309.78.

El 49 % de las plantaciones registradas en el PINFOR se localizan en el departamento de El Petén, en tanto el restante se distribuye en 4 departamentos, que son Izabal, Escuintla, Suchitepéquez y Alta Verapaz, siendo este último, el segundo lugar en importancia con el 22 % de la superficie plantada a nivel nacional (INAB, 2015).

2.4. Tablas de volumen

Una tabla de volumen se define según Ferreira (citado en Méndez y Lima, 2017) como una tabulación del volumen de los árboles de diversos tamaños, definidos por su DAP, altura y varios coeficientes de forma; considerándose el volumen como la variable dependiente, el DAP y la altura, como las variables regresoras.

Las tablas de volumen por lo general se generan a partir del desarrollo de fórmulas de volumen, las que consideran una o más variables, siendo las principales, DAP, altura y en algunos casos la forma del árbol (Romhan y Ramírez, 2010).

Son importantes ya que constituyen la base de los inventarios forestales, la silvicultura y de los planes de manejo sustentables.

2.4.1. Tipos de tablas de volumen

Existen diversos tipos de tablas de volumen, entre las que se pueden citar las regionales, locales, estándar, gráficas, de una entrada, de dos entradas, tarifas comerciales, totales, entre otros, en las que se basan los criterios siguientes: el número de variables incluidas, procedimiento de elaboración, área geográfica, unidades en que se diseña, cantidad del volumen de árboles en que se fundamentan, el tipo de material taxonómico, y de que sean aplicables para masas forestales o árboles individuales.

De todos los tipos mencionados, la que resulta más útil, es la basada en el área geográfica de aplicación y la especie, clasificándose en las categorías: tablas estándar y tablas en tarifas (Romhan y Ramírez, 2010).

2.4.1.1. Tablas estándar

Estas tablas por lo común están diseñadas para requerimientos de extensiones relativamente grandes, incluyendo como variables regresoras la altura, el DAP, y a veces algún coeficiente de forma (Romhan y Ramírez, 2010).

2.4.1.2. Tablas locales de volúmenes (tarifas)

Las tablas locales, consideran solamente rodales de carácter estrictamente local, pudiéndose construir en base a valores de campo o bien, derivadas de la tabla estándar directamente, considerando una relación diámetro-volumen (Romhan y Ramírez, 2010).

3. PRESENTACIÓN DE RESULTADOS

A continuación, se presentan los resultados de acuerdo con los objetivos planteados, considerando la evaluación de modelos estadísticos evaluados para determinar el de mejor ajuste para la predicción de volumen en *Tectona grandis* L.f:

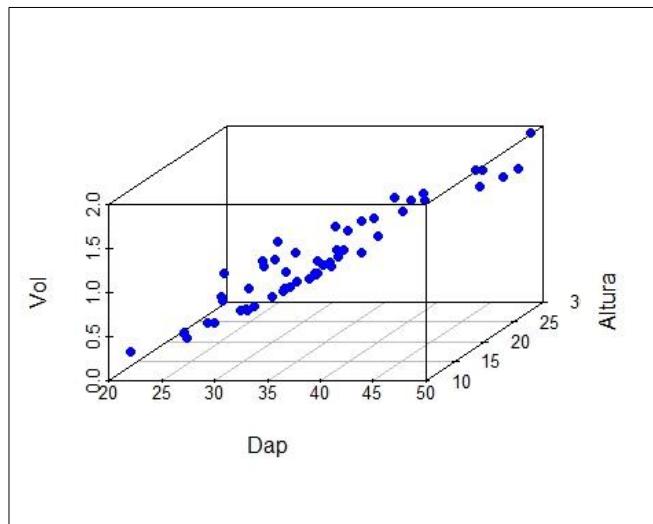
- 3.1. Objetivo 1: describir la tendencia de las variables que influyen en la estimación del volumen del árbol de Teca, mediante diagramas de dispersión y el coeficiente de correlación de Pearson, para inferir si dicha tendencia es lineal o no lineal.**

Como análisis exploratorio, se elaboraron diagramas de dispersión para observar el comportamiento de las variables entre la independiente y la dependiente, y un análisis de correlación para medir la fuerza y dirección de la asociación.

3.1.1. Diagrama de dispersión

Al graficar los valores del diámetro al pecho en centímetros (DAP) con altura en metros y volumen en metros cúbicos en un mismo plano se observa una relación directa, sugiriendo que al incrementarse el DAP de un árbol se incrementa la altura y en consecuencia el volumen, Figura 1.

Figura 1. **Diagrama de dispersión tridimensional**

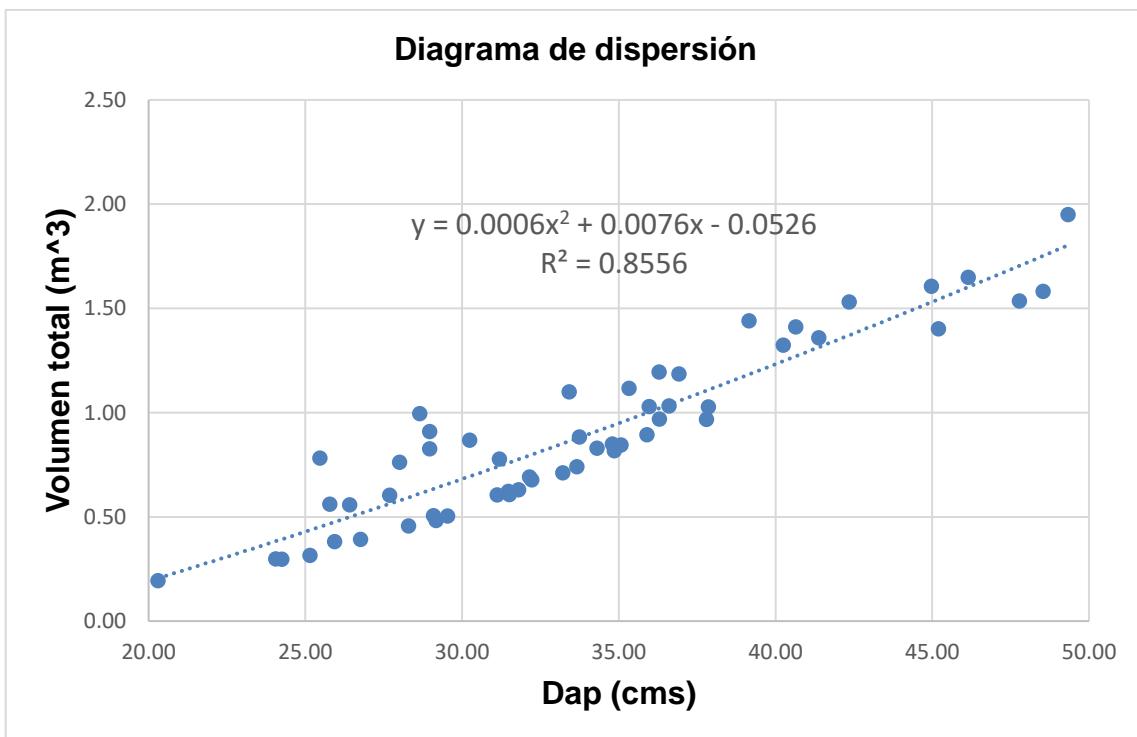


Fuente: elaboración propia, realizado con R.

El diagrama de dispersión muestra que individualmente existe una relación directa, al incrementarse el DAP del árbol y la altura se incrementa el volumen.

Al desarrollar de manera preliminar el análisis de regresión simple, se observó que la tendencia entre las variables es probablemente cuadrática (polinomio grado 2), obteniéndose un coeficiente de determinación de 0.8556, Figura 2.

Figura 2. **Diagrama de dispersión DAP respecto a volumen**

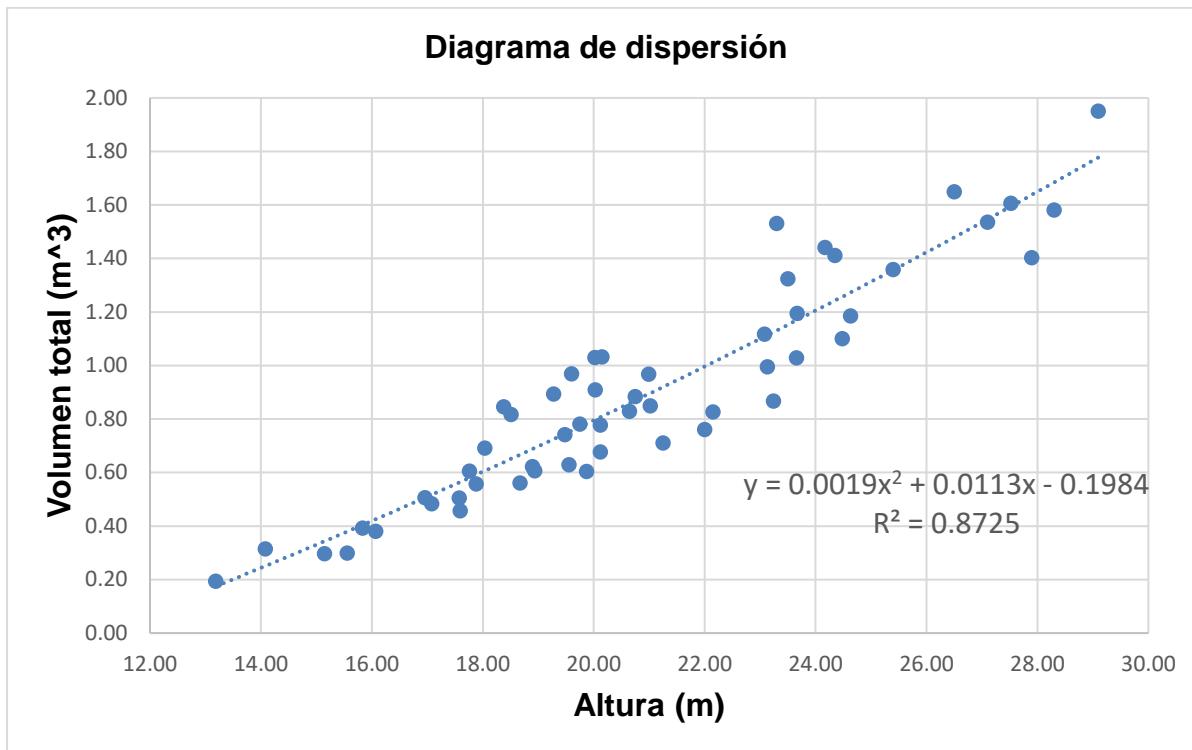


Fuente: elaboración propia, realizado con Microsoft Excel 365.

La gráfica con altura en el eje X y el volumen en el eje Y, indica que individualmente existe una relación directa entre variables, al incrementarse la altura del árbol se incrementa el volumen.

Al desarrollar de manera preliminar el análisis de regresión simple, se observó que la tendencia entre las variables probablemente es cuadrática (polinomio grado 2), obteniéndose un coeficiente de determinación de 0.8725, Figura 3.

Figura 3. **Diagrama de dispersión altura en función de volumen**

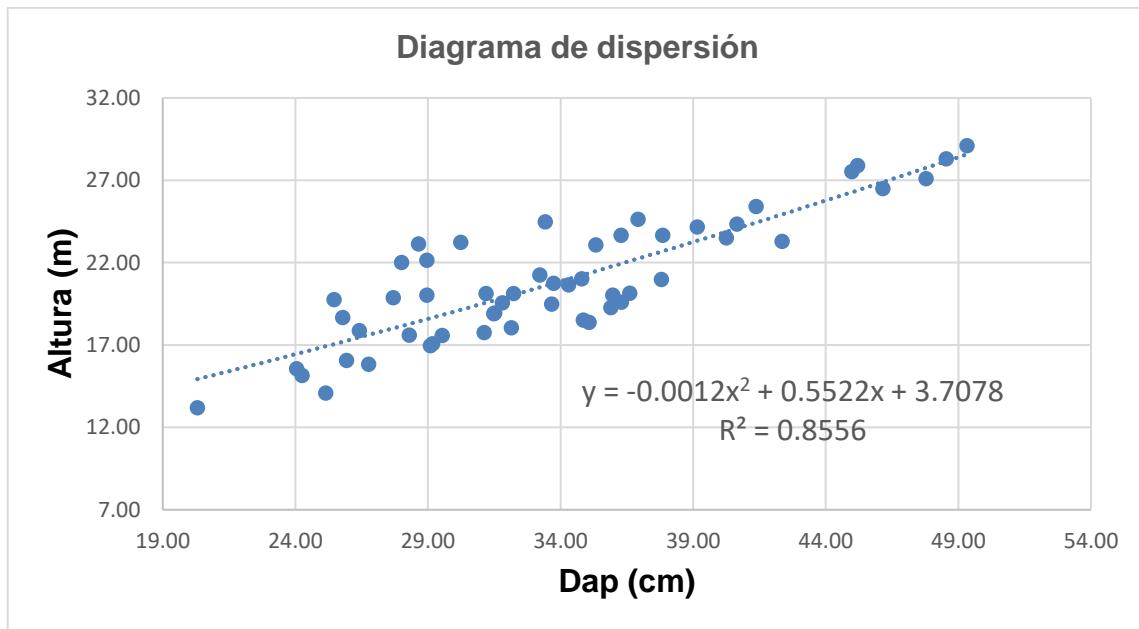


Fuente: elaboración propia, realizado con Microsoft Excel 365.

Para el caso de las variables independientes se elaboró un diagrama de dispersión, graficando la variable DAP en el eje X y la altura en el eje Y, observando que individualmente existe una relación directa, al incrementarse el DAP del árbol se incrementa la altura.

Al desarrollar de manera preliminar el análisis de regresión como simple, se observó que la tendencia entre las variables es probablemente cuadrática (polinomio grado 2), obteniéndose un coeficiente de determinación de 0.8556, Figura 4.

Figura 4. **Diagrama de dispersión DAP respecto altura**

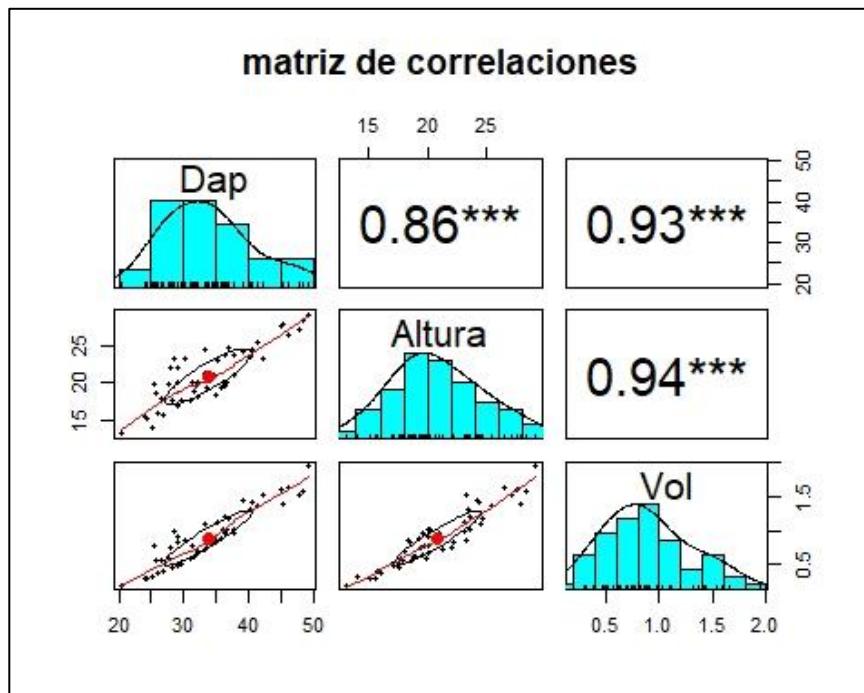


Fuente: elaboración propia, realizado con Microsoft Excel 365.

3.1.2. Coeficiente de correlación

Luego del análisis gráfico , se elaboró una matriz de correlaciones para conocer la intensidad y dirección de la relación existente, observándose que la variable DAP con la altura presentan un coeficiente de correlación de Pearson de 0.86, significando fuerte correlación positiva, así el DAP con el volumen 0.93 de correlación significando fuerte correlación positiva, y por último la variable altura con volumen presenta una correlación de 0.94, que implica una relación fuerte y directa, Figura 5.

Figura 5. **Matriz de correlaciones**



Fuente: elaboración propia, realizado con R.

La prueba de significancia de la correlación para las tres variables indica que es altamente significativa, ya que los valores de probabilidad (p-valor) son menores al nivel de significancia de 5 %, lo que aporta evidencia para concluir que muy probablemente existe asociación entre las variables, Tabla III.

Tabla III. **Correlación de Pearson**

Variable (1)	Variable (2)	n	Pearson	p-valor
DAP	Altura	54	0.86	<0.0001
DAP	Volumen	54	0.93	<0.0001
Altura	Volumen	54	0.94	<0.0001

Fuente: elaboración propia.

3.2. Objetivo 2: estimar el porcentaje promedio de leña, troza mayor y troza menor resultante de la corta de un árbol de Teca, mediante intervalos de confianza, para la planificación del manejo de la plantación

Se realizó la estimación de los productos del árbol de Teca mediante las proporciones de leña, troza mayor y menor, para seguidamente elaborar los intervalos de confianza en las clases diamétrales estudiadas.

3.2.1. Porcentaje promedio de troza comercial y leña

Analizando el árbol de Teca de manera general como se hace con las especies latifoliadas, este se clasifica con fines de aprovechamiento forestal en troza comercial (altura del árbol desde la base hasta donde el fuste se divide o presenta algún daño físico) y leña (altura del árbol a partir de donde el fuste se divide o presenta algún daño físico hasta el ápice).

Al calcular el volumen de troza comercial y leña por árbol, se relaciona cada valor con el volumen total para obtener el porcentaje específico, resultando que en promedio un árbol presenta un 89.41 % de troza comercial y un 10.59 % de leña, pero estas proporciones varían según clase diamétrica, tal como se observa en la tabla siguiente:

Tabla IV. Valores en porcentaje para troza y leña en Teca

Clase de DAP (cm)	Porcentaje volumen troza	Porcentaje volumen leña
[20 - 25)	85.09 %	14.91 %
[25 - 30)	86.94 %	13.06 %
[30 -35)	90.08 %	9.92 %

Continuación de la tabla IV.

Clase de DAP (cm)	Porcentaje volumen troza	Porcentaje volumen leña
[35 - 40)	90.96 %	9.04 %
[40 - 45)	91.96 %	8.04 %
[45 - 50)	91.41 %	8.59 %
Promedio	89.41 %	10.59 %

Fuente: elaboración propia.

Como se menciona anteriormente, a nivel general en especies latifoliadas un árbol con fines de aprovechamiento forestal se divide en troza comercial y leña, pero en el caso particular de la especie Teca, en la práctica por razones de mercado y valor económico, la troza comercial se clasifica en troza menor (medida de 15.91 a 22.28 cm de diámetro), troza mayor (medida mayor de 22.28 cm de diámetro), y la leña (diámetro menor de 15.91 cm).

Al calcular el volumen de troza menor, troza mayor y leña por árbol, se relaciona cada valor con el volumen total para obtener el porcentaje específico, resultando que en promedio un árbol presenta un 10.08 % de leña, 54.24 % de troza menor y un 35.68 % de troza mayor, pero estas proporciones varían según clase diamétrica, tal como se observa en la tabla siguiente:

Tabla V. **Valores en porcentaje para leña, troza menor y troza mayor**

Clase de DAP (cm)	Leña (< 15.91 cm Diámetro)	Troza menor (15.91 - 22.28 cm) Diámetro	Troza mayor (>22.28 cm) Diámetro
[20 - 25)	14.17	73.60	12.23
[25 - 30)	11.78	56.13	32.09
[30 -35)	9.04	53.48	37.48
[35 - 40)	7.42	47.97	44.61
[40 - 45)	10.07	47.98	41.95

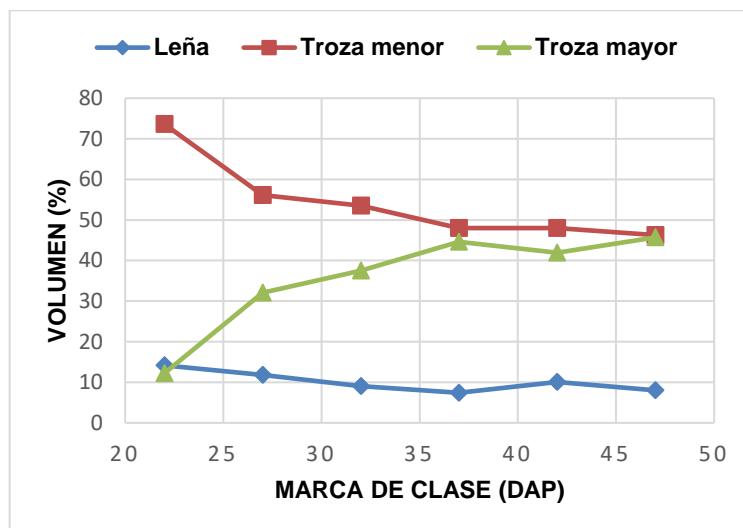
Continuación de la tabla V.

Clase de DAP (cm) [45 - 50]	Leña (< 15.91 cm diámetro)	Troza menor (15.91 - 22.28 cm) diámetro	Troza mayor (>22.28 cm) diámetro
% TOTAL	10.08	54.24	35.68

Fuente: elaboración propia.

En la figura 6 se observa claramente que el producto leña no varía mucho con respecto del tamaño del DAP, sin embargo, los productos troza menor y mayor cambian su proporción a medida que aumenta el DAP, igualándose estas proporciones al aproximarse a la clase de DAP de 45 a 50 centímetros.

Figura 6. Proporción de productos por árbol de Teca



Fuente: elaboración propia.

3.2.2. Intervalos de confianza para las proporciones

En el caso del intervalo de confianza para la proporción de la troza comercial que resulta en un 89.41 %, con el tamaño de la muestra n= 54, nivel de confianza del 95 % con un valor de Z de 1.96, presenta los valores siguientes:

$$0.8941 \pm 1.96 * \left(\frac{\sqrt{0.8941(1 - 0.8941)}}{54} \right)$$

Donde:

- Límite inferior = 81.20 %.
- Límite superior= 97.62 %.

Para el intervalo de confianza de la leña cuyo valor resulta en 10.59 % se obtienen los valores siguientes:

$$0.1059 \pm 1.96 * \left(\frac{\sqrt{0.1059(1 - 0.1059)}}{54} \right)$$

Donde:

- Límite inferior = 2.59 %.
- Límite superior= 18.79 %.

Para estructurar el intervalo de confianza para leña (0.1008) se relaciona con el total de troza menor y troza mayor (0.8992 %), considerando la totalidad de árboles (n=54 árboles) y un nivel de confianza del 95 % (valor de Z=1.96), resultan los valores siguientes:

$$0.1008 \pm 1.96 * \left(\frac{\sqrt{0.1008(1 - 0.1008)}}{54} \right)$$

Donde:

- Límite inferior = 2.05 %.
- Límite superior= 18.11 %.

En el caso del intervalo de confianza para troza menor (0.5424) se relaciona con el total de leña y troza mayor (0.4576), resultando los valores siguientes:

$$0.5424 \pm 1.96 * \left(\frac{\sqrt{0.5424(1 - 0.5424)}}{54} \right)$$

Donde:

- Límite inferior = 40.95 %.
- Límite superior= 67.53 %.

Para el intervalo de confianza de troza mayor (0.3568) se relaciona con el total de leña y troza menor (0.716), resultando los valores siguientes:

$$0.3568 \pm 1.96 * \left(\frac{\sqrt{0.3568(1 - 0.3568)}}{54} \right)$$

Donde:

- Límite inferior = 22.90 %.
- Límite superior= 48.46 %.

Según la muestra de árboles de teca la proporción por árbol de cada producto resultó, para leña con valores desde 2.05 % a 18.11 %, troza menor desde 40.95 % a 67.53 %, y troza mayor desde 22.90 % hasta 48.46 %, esto con un nivel de confianza del 95%, Tabla VI.

Tabla VI. Intervalos de confianza para los productos de Teca

Producto	Límite inferior	Límite superior
Leña	2.05 %	18.11 %
Troza menor	40.95 %	67.53 %
Troza mayor	22.90 %	48.46 %

Fuente: elaboración propia.

3.3. Objetivo 3: establecer el modelo de mejor ajuste para predecir el volumen total del árbol de Teca, y diseñar una tabla de volumen

La evaluación de los quince modelos se realizó para el volumen total del árbol de Teca, para posteriormente diseñar una tabla de volumen detallada por diámetro a la altura al pecho y altura, que facilite el cálculo de volumetría.

3.3.1. Modelos evaluados para el volumen total

En la Tabla VII se presentan los estadísticos de los modelos evaluados, que sirvieron para la elección del modelo de mejor ajuste para la predicción del volumen, los que se ordenan desde los mejor calificados a los menos calificados.

Tabla VII. Indicadores de los modelos evaluados

Modelo núm.	Modelos	Error estándar	R ² adj	AIC	b _i	Valor	P-valor
15	$\beta_0 + \beta_1.d + \beta_2.(dH) + \beta_3.d^2 + \beta_4(d^2.H)$	0.1032	0.9278	-309.93	b ₀ 4.2762 b ₁ -0.4592 b ₂ 0.0102 b ₃ 0.0096 b ₄ -0.0002	0.0018 0.0002 <0.0001 0.0001 <0.0001	
6	$\beta_1 + \beta_2.d^2 + \beta_3.d^2.H + \beta_4.H^2 + \beta_5.D.H^2$	0.1164	0.9081	-293.25	b ₁ -0.8365 b ₂ 0.0000 b ₃ 0.0000 b ₄ 0.0067 b ₅ -0.0002	<0.0001 0.7490 0.1791 0.0022 0.0559	
2	$\beta_0 + \beta_1.d^2 + \beta_2.H + \beta_3.d^2.H$	0.1427	0.9018	-266.41	b ₀ -0.3746 b ₁ 0.0001 b ₂ 0.0422 b ₃ 0.0000	<0.0001 0.0004 <0.0001 0.0788	
13	$\beta_0 - \beta_1.H + \beta_2.D^2.H$	0.1429	0.9000	-267.36	b ₀ -0.3545 b ₁ -0.0444 b ₂ 0.0000	0.0007 <0.0001 <0.0001	
1	$\beta_0 + \beta_1.d\beta_2.H\beta_3$	0.1446	0.8991	-264.57	b ₀ -0.1193 b ₁ 0.0026 b ₂ 0.6964 b ₃ 1.1565	0.0848 0.4667 0.0015 0.0027	
4	$\beta_0 + \beta_1.d + \beta_2.d^2 + \beta_3.d.H$	0.1457	0.8976	-263.54	b ₀ -0.0487 b ₁ -0.0005 b ₂ -0.0001 b ₃ 0.0016	0.0139 0.0395 0.0249 <0.0001	
12	$\beta_0 + \beta_1.D.H + \beta_2.D^2$	0.1446	0.8976	-265.73	b ₀ -0.0561 b ₁ 0.0016 b ₂ -0.0001	0.0015 <0.0001 0.3672	
7	$\beta_1.d\beta_2.H\beta_3$	0.1453	0.8966	-265.07	b ₁ 0.0009 b ₂ 0.8259 b ₃ 1.2827	0.0041 <0.0001 <0.0001	
9	$\beta_1(D+1)\beta_2.H\beta_3$	0.1454	0.8965	-264.98	b ₁ 0.0008 b ₂ 0.8499 b ₃ 1.2836	0.0048 <0.0001 <0.0001	
3	$\beta_0 + \beta_1.d^2 + \beta_2.d^3 + \beta_3.H + \beta_4/H$	0.1254	0.8933	-282.99	b ₀ -0.6921 b ₁ 0.0010 b ₂ -0.0000 b ₃ 0.0458 b ₄ -2.4108	0.5669 0.1247 0.5028 0.1025 0.9505	
8	$\beta_1(d^2.H)\beta_2$	0.1422	0.8566	-269.1	b ₁ 0.0007 b ₂ 0.7068	0.0063 <0.0001	
5	$\beta_0 + \beta_1.d^2.H + \beta_2.d^2$	0.1687	0.8606	-244.41	b ₀ 0.1314 b ₁ 0.0000 b ₂ 0.0004	0.0400 0.0075 0.6078	
11	$\beta_0 + \beta_1(d^2H)$	0.1728	0.8516	-242.22	b ₀ 0.2375 b ₁ 0.0000	<0.0001 <0.0001	
14	$\exp(b_0+b_1*(x_1)^2+b_2*(x_2))$	0.1762	0.8479	-238.40	b ₀ -2.0190 b ₁ 0.0001 b ₂ 0.0782	>0.9999 >0.9999 >0.9999	

Continuación de la tabla VII.

Modelo núm.	Modelos	Error estándar	R ² adj	AIC	bi	Valor	P-valor
10	$\beta_1(H/d)\beta_2 \cdot d^2 \cdot H$	0.1976	0.8058	-223.6	b1 b2	0.0000 1.4881	<0.0001 <0.0001

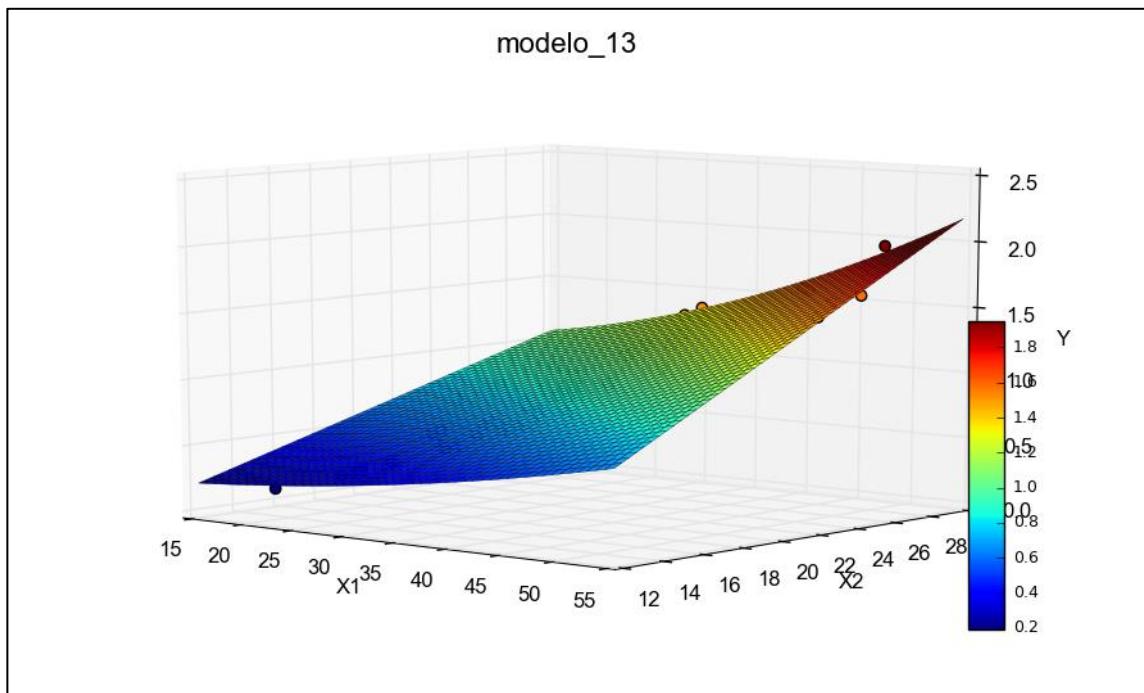
Fuente: elaboración propia.

Los modelos que presentaron el mejor ajuste fueron el denominado Meyer modificado ($\beta_0 + \beta_1.d + \beta_2.(dH) + \beta_3.d^2 + \beta_4(d^2.H)$) y el modelo cuadrático ($\beta_0 - \beta_1.H + \beta_2. D^2.H$), ya que sus valores de coeficientes de determinación resultante son de 92.78 % y 90 %, respectivamente, además todos los coeficientes del modelo son significativos al 5 %, pero se eligió el segundo por ser más parsimonioso, ya que presenta menor número de coeficientes, y cumple con los supuestos.

El modelo de Meyer modificado no cumplió con los supuestos, por ese motivo se descartó, a pesar de que presentó mejor coeficiente de determinación y AIC que el modelo denominado cuadrático.

Aunado al análisis de los indicadores, en el gráfico siguiente se observa el ajuste que presenta el modelo seleccionado.

Figura 7. Ajuste del modelo seleccionado



Fuente: elaboración propia, realizado con Curve Expert Professional 2.7.3.

3.3.2. Verificación de supuestos del modelo

Al modelo seleccionado se le verifico la significancia de sus coeficientes, el supuesto de normalidad, homocedasticidad, independencia, y multicolinealidad, presentándose a continuación las pruebas correspondientes:

3.3.2.1. Prueba de significancia del modelo seleccionado

Todos los coeficientes del modelo seleccionado son significativos al 5 % como se puede observar en la tabla siguiente:

Tabla VIII. Prueba de significancia del modelo seleccionado

$$Modelo y = \beta_0 - \beta_1 \cdot H + \beta_2 \cdot D^2 \cdot H$$

Variable	N	CMError	Sigma	AIC	BIC	Iteración
y	54	0.01	0.12	-75.40	-67.44	6

Parámetros	Cota inf.	Cota sup.	Val.Ini.	Estimación	E.E.	T	p-valor
B0	-1*10^30	1*10^30	1*10^-03	-0.55	0.15	-3.59	0.0007
B1	-1*10^30	1*10^30	1*10^-03	-0.05	0.01	-5.35	<0.0001
B2	-1*10^30	1*10^30	1*10^-03	1.3*10^-05	2.4*10^-06	5.44	<0.0001

Matriz de correlación de las estimaciones

	B0	B1	B2
B0	1.00	0.98	0.81
B1	0.98	1.00	0.90
B2	0.81	0.90	1.00

Fuente: elaboración propia.

3.3.2.2. Supuesto de normalidad del modelo seleccionado

A los residuos del modelo seleccionado se le realizó la prueba de Shapiro-Wilks, obteniéndose una probabilidad (p-valor) de 0.59, lo que indica que presentan normalidad al 5 % de significancia, como se observa en la tabla siguiente:

Tabla IX. Prueba de Shapiro-Wilks (modificado)

Shapiro-Wilk normality test

data: mod2\$residuals
W = 0.982, p-value = 0.59

Fuente: elaboración propia.

3.3.2.3. Homocedasticidad

Se verificó el supuesto de homogeneidad de varianzas, mediante la prueba de Harrison-Mcabe, obteniéndose un valor de probabilidad (p-valor) de 0.14, concluyéndose que el modelo cumple con el supuesto de homocedasticidad al 5 % de significancia.

Tabla X. **Prueba de Harrison-Mcabe**

Harrison-McCabe test
data: mod2
HMC = 0.401, p-value = 0.14

Fuente: elaboración propia.

3.3.2.4. Independencia

Se realizó la prueba de Durbin Watson para verificar la independencia de los residuos, determinándose que existe independencia, presentando una probabilidad (p-valor) mayor al 5 % de significancia.

Tabla XI. **Prueba de Durbin Watson**

lag	Autocorrelation	D-W Statistic	p-value
1	0.15546	1.6839	0.168
Alternative hypothesis: rho! = 0			

Fuente: elaboración propia.

3.3.2.5. Multicolinealidad

Para verificar la existencia de multicolinealidad entre los coeficientes del modelo se calculó el índice de condición sin intercepto, con intercepto y se estableció el porcentaje de incremento, determinando que no existe multicolinealidad, porque los valores obtenidos son menores a un valor de 30.

Tabla XII. Índice de condición

\$`Condition Number without intercept	\$`Condition Number with intercept`	\$`Increase (in percentage)`
19.958	25.103	20.497

Fuente: elaboración propia.

3.3.3. Diseño de la tabla de volumen total

Posteriormente a seleccionar el modelo de mejor ajuste ($\beta_0 - \beta_1 \cdot H + \beta_2 \cdot D^2 \cdot H$) considerando que las clases diamatrales de la muestra oscilan entre 20 a 50 centímetros de diámetro al pecho y que los árboles presentaron una altura máxima de 30 metros, se elaboró la tabla de volumen siguiente:

Tabla XIII. Tabla de volumen total en m³

(CM)	DAP											ALTURA TOTAL (METROS)										
	10	12	14	16	18	20	22	24	26	28	30	10	12	14	16	18	20	22	24	26	28	30
20	0.141	0.241	0.340	0.439	0.538	0.637	0.737	0.836	0.935	1.034	1.134											
22	0.152	0.254	0.355	0.457	0.558	0.659	0.761	0.862	0.964	1.065	1.166											
24	0.164	0.268	0.372	0.476	0.579	0.683	0.787	0.891	0.995	1.098	1.202											
26	0.177	0.284	0.390	0.496	0.603	0.709	0.816	0.922	1.028	1.135	1.241											
28	0.191	0.301	0.410	0.519	0.628	0.737	0.847	0.956	1.065	1.174	1.283											

Continuación de tabla XIII.

(CM)	10	12	14	16	18	20	22	24	26	28	30
30	0.206	0.319	0.431	0.543	0.655	0.767	0.880	0.992	1.104	1.216	1.329
32	0.223	0.338	0.453	0.569	0.684	0.800	0.915	1.031	1.146	1.261	1.377
34	0.240	0.359	0.477	0.596	0.715	0.834	0.953	1.072	1.191	1.309	1.428
36	0.258	0.380	0.503	0.625	0.748	0.870	0.993	1.115	1.238	1.360	1.483
38	0.277	0.404	0.530	0.656	0.783	0.909	1.035	1.162	1.288	1.414	1.541
40	0.297	0.428	0.558	0.689	0.819	0.949	1.080	1.210	1.341	1.471	1.602
42	0.319	0.453	0.588	0.723	0.857	0.992	1.127	1.261	1.396	1.531	1.665
44	0.341	0.480	0.619	0.759	0.898	1.037	1.176	1.315	1.454	1.593	1.733
46	0.365	0.508	0.652	0.796	0.940	1.084	1.227	1.371	1.515	1.659	1.803
48	0.389	0.538	0.686	0.835	0.984	1.133	1.281	1.430	1.579	1.727	1.876
50	0.414	0.568	0.722	0.876	1.030	1.183	1.337	1.491	1.645	1.799	1.953

Fuente: elaboración propia.

3.3.4. Comparación del modelo cuadrático con FAO

El modelo de FAO para la predicción de volumetría total en especies latifoliadas es el siguiente: $= 0.108337266 + 0.000046499 * (D^2 * H)$.

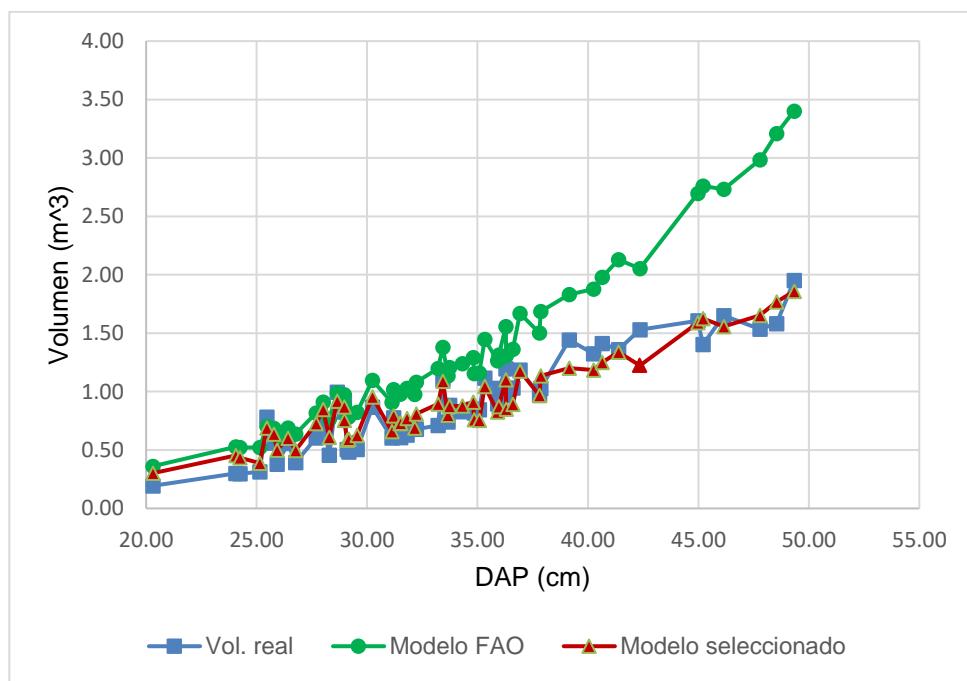
Donde:

- V: Volumen total en metros cúbicos.
- D: DAP en cm.
- H: Altura total en metros.

Se realizó la estimación de volumetría utilizando el modelo FAO y el modelo seleccionado, para su comparación con el volumen real por árbol, cálculos que se presentan en el apéndice 2.

Al graficar los valores estimados de volumen con el modelo FAO, el modelo seleccionado, y los valores reales, se observa diferencia significativa entre el modelo FAO y el volumen real, no así entre el volumen real con el modelo seleccionado, en donde se ve un comportamiento entre estos, similar, Figura 8.

Figura 8. Gráfico de comparación de volúmenes



Fuente: elaboración propia, realizado con Excel 365.

Al realizar una prueba de muestras apareadas usando la distribución de T Student se observa que hay diferencia significativa entre la volumetría estimada con el modelo FAO y el volumen real, ya que la probabilidad (p-valor) de <0.0001 es menor que el nivel de significancia del 5 %, resultando en una sobre estimación del modelo en comparación del volumen real, Tabla XIV.

Tabla XIV. Prueba de medias volumen real y modelo FAO

Prueba T (muestras apareadas)								
Obs(1)	Obs(2)	N	Media(dif)	Media(1)	Media(2)	DE(dif)	T	Bilateral
Vol. real	Modelo FAO	54	-0.43	0.89	1.33	0.37	-8.67	<0.0001

Fuente: elaboración propia.

Al comparar el volumen real del árbol de Teca con lo estimado utilizando el modelo seleccionado se determinó que la diferencia no es significativa al 5 % de significancia, porque la probabilidad (p-valor) de 0.2776 es mayor a este nivel, lo que sugiere que el modelo predice la volumetría acorde al comportamiento real de la especie estudiada, Tabla XV.

Tabla XV. Prueba de medias volumen real y modelo seleccionado

Prueba T (muestras apareadas)								
Obs(1)	Obs(2)	N	media(dif)	Media(1)	Media(2)	DE(dif)	T	Bilateral
Vol. real	Mod. seleccionado	54	-0.02	0.89	0.91	0.12	-1.10	0.2776

Fuente: elaboración propia.

3.4. Objetivo general: evaluar y seleccionar mediante análisis de regresión múltiple el modelo matemático de mejor ajuste para estimar el volumen de áboles de teca (*Tectona grandis L.f*) en plantaciones del departamento de Alta Verapaz, Guatemala

Se realizó la evaluación de los quince modelos propuestos en base a referencias, tomándose como criterio de selección: el coeficiente de determinación ajustado , el criterio de información Akaike (AIC), el error estándar,

la significancia de los parámetros de regresión, y el cumplimiento de los supuestos de normalidad, homogeneidad de varianzas, independencia y multicolinealidad, estableciendo que el modelo de mejor ajuste, para la volumetría total, es el denominado cuadrático cuya expresión algebraica es $\beta_0 - \beta_1.H + \beta_2.d^2.H$, ya que además de cumplir con los criterios antes mencionados, es el más parsimonioso.

4. DISCUSIÓN DE RESULTADOS

El estudio se realizó con el fin de evaluar quince modelos estadísticos, que han sido estudiados en otros países para la predicción de volumen de Teca, considerando las variables regresoras DAP (cm) y altura total (m), y que en nuestro medio se ha estado usando modelos generales que generan sobre estimación de la madera, lo que incide en errores a la hora de planificar el aprovechamiento de las plantaciones.

4.1. Análisis interno

El modelo seleccionado no presenta diferencia significativa al 5 % con los valores reales de volumetría, lo cual sugiere que existe un buen ajuste en la estimación del volumen.

El modelo propuesto por la FAO, recomendado por el INAB, presenta diferencia significativa al 5 % con los valores reales de volumetría, observándose sobre estimación con relación al volumen observado.

En la recolección de datos se tuvo dificultades principalmente por las condiciones topográficas donde se ubicaba la plantación, ya que eran terrenos escarpados lo que provocó que esta actividad fuera compleja y no se pudiera realizar con celeridad.

Al desarrollar el análisis exploratorio de datos, elaborando histogramas y pruebas de normalidad para comprender la distribución de frecuencias se notó que al incluir árboles desde los diez centímetros hasta los cincuenta centímetros

de diámetro a la altura al pecho, la distribución se distorsionaba, provocando que los datos de manera preliminar no presentaran normalidad, siendo necesario hacer una depuración de los mismos, decidiendo finalmente trabajar solamente con clases diametrales desde los veinte centímetros hasta los cincuenta centímetros.

En este mismo tema, se había contemplado modelar para todos los productos del árbol de Teca, como troza menor, troza mayor, leña, y volumen total, pero en los análisis preliminares se detectó que el producto leña no presentaba normalidad, lo cual ocasionaría que la tabla de volumen a diseñarse quedaría incompleta, por lo que se decidió realizar la evaluación de los modelos solamente con el volumen total del árbol sin realizar la desagregación en productos, considerando que esta evaluación sobre el volumen es un importante aporte para el medio forestal porque daría certeza en cuanto al volumen de los árboles de las plantaciones en la planificación del manejo y aprovechamiento.

4.2. Análisis externo

En el análisis exploratorio se elaboraron diagramas de dispersión tanto bidimensionales como tridimensional, relacionando las variables diámetro a la altura al pecho con el volumen total, altura con el volumen total, diámetro a la altura al pecho, resultando que existe asociación directa fuerte entre estas variables, relación que es de tipo polinómica grado dos (cuadrática), o sea no lineal, lo que resulta acorde a lo afirmado por Cancino (2006), quien indicó que generalmente esas relaciones no son lineales.

En el caso del diagrama de dispersión tridimensional se observa que, al incrementarse el diámetro a la altura al pecho, y la altura del árbol, se incrementa el volumen total, siendo relaciones directas entre todas las variables.

El árbol de teca por razones comerciales se secciona en los productos troza mayor, troza menor y leña, para el presente estudio se observó que en todas las clases diametrales el producto existente en mayor proporción es la troza menor (diámetros de 15.91 a 22.28 cm), y en menor proporción la leña (diámetros menores a 15.91 cm), lo que es consistente con lo afirmado por Mora y Hernández (2007), en relación con las proporciones de mencionados productos.

Entre los modelos mejor evaluados destaca el modelo de Meyer modificado y el cuadrático, resultando un coeficiente de determinación ajustado de 92.78 % y 90 %, respectivamente, pero a pesar de estar mejor calificado el primero, se eligió finalmente el segundo, porque cumplió con todos los supuestos. En el caso del estudio de Telles *et al.* (2018), al evaluar seis modelos matemáticos para predecir el volumen fustal se seleccionó el modelo de Meyer ya que fue el de mejor ajuste, y cumplió con los supuestos, en el presente estudio no se seleccionó este modelo porque no cumplió con los supuestos.

Al realizar estimaciones con el modelo cuadrático se obtiene mejor ajuste que con el modelo recomendado por el INAB (2015), esto porque se realizó una prueba de hipótesis para medias dependientes, para probar significancia entre los valores del volumen real y el modelo seleccionado, entonces se infiere que probablemente no existe diferencia significativa entre estas estimaciones, no así con el modelo de FAO que sí presenta diferencia significativa a un nivel del 5 %, esto porque al usarse este modelo se sobre estima el volumen real.

La tabla de volumen propuesta es útil para realizar estimaciones volumétricas de árboles comprendidos entre clases diametrales de 20 a 50 centímetros, y de alturas totales de 10 a 30 metros en vista que el conjunto de datos colectado en campo presenta este dominio de estudio, situación que indica,

Rueda *et al.* (2020) que debe considerarse al hacer el uso de las tablas, así como su validación.

CONCLUSIONES

1. La relación de las variables diámetro a la altura al pecho y altura con el volumen total del árbol de Teca presenta una asociación de tipo cuadrática, la cual es directa y fuerte, lo que significa que, al incrementarse el diámetro a la altura al pecho y la altura, se incrementa el volumen total del árbol.
2. El porcentaje promedio de troza y leña del árbol de teca presenta un volumen de troza del 89.41 % (IC 81.20 % - 97.62 %) y 10.59 % de leña (IC 2.59 % - 18.79 %), valores que varían según la clase diamétrica, notándose que al incrementarse el diámetro disminuye el volumen de leña. Para los árboles con fines comerciales, las proporciones de troza mayor en promedio corresponden a un 35.68 % (IC 22.90 % - 48.46 %), la troza menor un 54.24 % (IC 40.95 % - 67.53 %), y la leña un 10.08 % (IC 2.05 % - 18.11 %).
3. El modelo de mejor ajuste para la predicción del volumen total de Teca corresponde a la expresión $V= -0.3545 - (-) 0.0444 * H + 0.000013 * D^2 * H$, la cual es la más parsimoniosa, sus coeficientes son significativos, y cumple con todos los supuestos. Este modelo permitió elaborar la tabla de volumen de doble entrada a nivel local, que facilitará la planificación del aprovechamiento de la plantación de Teca.
4. La evaluación de los modelos volumétricos mediante regresión múltiple permitió seleccionar el de mejor ajuste, el cual predice el volumen total del

árbol de Teca con exactitud, esto porque al comparar el volumen estimado y los valores reales de volumen, no existe diferencia significativa.

RECOMENDACIONES

1. Se hace necesario continuar analizando plantaciones en diferentes sitios, para estudiar las tendencias de las variables diámetro a la altura al pecho, altura y volumen, utilizando diagramas de dispersión y el análisis de correlación para comprender mejor el comportamiento de las relaciones y la volumetría del árbol de Teca.
2. Se debe continuar haciendo estimaciones del porcentaje promedio de los productos del árbol de Teca como troza menor, troza mayor y leña en plantaciones forestales de diferentes sitios para determinarlos con precisión, ya que existen variaciones entre clases diamatrales y posiblemente localidades diferentes, para que se permita planificar adecuadamente el manejo y aprovechamiento forestal en las diferentes regiones del país.
3. Se sugiere realizar investigaciones similares al presente estudio para establecer el modelo matemático de mejor ajuste para predecir el volumen del árbol de Teca, en otros sitios con diferentes condiciones biofísicas para contar con datos acordes a la realidad del crecimiento y desarrollo de las plantaciones.
4. Se sugiere el uso del modelo cuadrático, para predecir el volumen total del árbol de Teca, porque es el de mejor ajuste y cumple con los supuestos, lo que garantiza predicciones precisas.

REFERENCIAS

1. Anderson, D., Sweeney, D., y Williams, T. (2012). *Estadística para negocios y economía*. México, D.F: Cengage, Learning. Recuperado de: https://www.academia.edu/37099878/Estad%C3%ADstica_para_negocios_y_econom%C3%ADA.
2. Barrantes, K., Murillo, R., Ávila, C., Fonseca, W., y Barquero, A. (Enero, 2021). Tabla de volumen comercial para teca clonal en diferentes índices de sitio en la Zona Norte de Costa Rica. *Revista de Ciencias Ambientales*, 55(1), 221–229. Recuperado de: <https://www.scielo.sa.cr/pdf/rca/v55n1/2215-3896-rca-55-01-211.pdf>.
3. Camino, R., y Morales, J. (2013). *Las plantaciones de teca en América Latina: Mitos y realidades*. Turrialba, Costa Rica,: In Serie técnica / CATIE. Recuperado de: https://www.researchgate.net/publication/259908884_Las_plantaciones_de_teca_en_America_Latina_Mitos_y_realidades.
4. Cancino, J. (2006). *Dendrometría básica*. Concepción, Chile: Universidad de Concepción. Recuperado de: http://repositorio.udec.cl/bitstream/11594/407/2/Dendrometria_Basica.pdf.

5. Corral, N., y López, M., (2019). *Regresión Lineal Múltiple*. Oviedo, España: Universidad de Oviedo. Recuperado de: https://bellman.ciencias.uniovi.es/norberto/ModelizacionEstadistica/Apuntes/regresion_2019.pdf.
6. Fondo Nacional de Financiamiento Forestal. (2003). *Manual para productores de Teca en Costa Rica*. Heredia, Costa Rica: Autor. Recuperado de: <https://www.fonafifo.go.cr/media/1332/manual-para-productores-de-teca.pdf>.
7. Instituto Nacional de Bosques. (2015). *Crecimiento y productividad de plantaciones forestales de Teca (Tectonis grandis L.f) en Guatemala*. Guatemala: Autor. Recuperado de: http://www.ppm.inab.gob.gt/docs/crecimiento_y_productividad_de_teca.pdf.
8. Instituto Nacional de Bosques. (2019). *Guía práctica para la cubicación de productos forestales*. Guatemala: Autor. Recuperado de: [https://inab.gob.gt/images/centro_descargas/manuales/Guia practica para la cubicación de productos forestales Nov 2019.pdf](https://inab.gob.gt/images/centro_descargas/manuales/Guia_practica para la cubicación de productos forestales Nov 2019.pdf).
9. Jimenes, U. (2019). *Introducción a la econometría*. Valencia, España: Universidad de Valencia. Recuperado de: <https://www.uv.es/uriel/material/multicolinealidad3.pdf>.
10. Levin, R., y Rubin, D. (2010). *Estadística para administración y economía* (7ma. edición). México, D.F.: Pearson, Educación. Recuperado de: <content/uploads/2017/12/Estad%C3%ADstica-para-administraci%C3%B3n-y-economia-Richard-L.-Levin.pdf>.

11. Lind, D., Marchal, W., y Wathen, S., (2012). *Estadística aplicada a los negocios y la economía*. México, D.F.: Mc Graw Hill. Recuperado de:https://eduvirtual.cuc.edu.co/moodle/pluginfile.php/523770/mod_resource/content/1/Estadistica%20para%20Administraion%20y%20Negocios.pdf.
12. López, E., y González, B. (2017). *Principio de muestreo estadístico*. Guatemala: Universidad de San Carlos de Guatemala. Recuperado de:<https://archive.org/details/PrincipiosDeMuestreoEstadsticoEd.Agosto2017/page/n1/mode/2up>.
13. López, F., Pérez, T., Prieto, M., Ramírez, F., Reyes, A., y Treviño, F. (2010). *Introducción a los modelos de Análisis multivariante*. México, D.F.: Aprenda ediciones. Recuperado de:<https://www.manuelprieto.com/publicaciones/libroIntroduccionalosModelos.pdf>.
14. Méndez, B., y Lima, L., (2017). *Influencia de la distribución de productos de plantaciones forestales establecidas por PINFOR en la continuidad de la sostenibilidad de la actividad forestal en plantaciones de pino maximinoi en el Norte de Guatemala*. Guatemala, Guatemala: Dirección General de Investigación. Recuperado de:<https://digi.usac.edu.gt/bvirtual/informes/resumenes2016/inf1636.html>.
15. Mollinedo, M., Herrera, M., y Muñoz, F. (Julio, 2016). Caracterización del crecimiento de plantaciones jóvenes de Teca y estimación de

- curvas de índice de sitio. *Revista Madera y Bosques*, 22(2), 89–103. Recuperado de: <https://www.redalyc.org/pdf/617/61749306007.pdf>.
16. Mora, F., y Hernández, W. (Marzo, 2007). Estimación del volumen comercial por producto para rodales de teca en el pacífico de Costa Rica. *Revista Agronomía Costarricense* 31(1), 101–112. Recuperado de: <https://www.redalyc.org/pdf/436/43631112.pdf>.
 17. Moret, A., Jerez, M., y Mora, A. (1998). Determinación de ecuaciones de volumen para plantaciones de teca (*Tectona grandis* L.F) en la unidad experimental de la reserva forestal Caparo, Estado barinas - Venezuela. *Revista Forestal* 42(1), 41–50. Recuperado de: https://www.researchgate.net/publication/255993942_Determinacion_de_ecuaciones_de_volumen_para_plantaciones_de_teca_Tectona_grandis_L_en_la_Unidad_Experimental_de_la_Reserva_Forestal_Caparo_Estado_Barinas-Venezuela.
 18. Newbold, P., Carlson, W., y Thorne, B. (2008). *Estadística para Administración y Economía*. Madrid, España: Pearson educación. Recuperado:https://www.academia.edu/15542328/Estadistica_para_Administracion_y_Economia_6ta_Ed_Paul_Newbold_William_Berry.
 19. Picard, N., Saint-André, L. y Henry, M. (2012). *Manual de construcción de ecuaciones alométricas para estimar el volumen y la biomasa de los árboles, del trabajo de campo a la predicción*. Paris, Francia: Las Naciones Unidas para la Alimentación y la Agricultura y el Centre de Coopération Internationale en Recherche Ag. CIRAD y

- FAO. Recuperado de:
<http://www.fao.org/docrep/018/i3058s/i3058s.pdf>.
20. Romhan, C., y Ramírez, H. (2010). *Dendrometría*. Chapingo, México: Universidad Autónoma de Chapingo. Recuperado de:
<http://dicio.chapingo.mx/pdf/publicaciones/dendrometria.pdf>.
21. Rueda, A., Benavides, J., García, D., Saénz, T., Castillo, D., y García, E. (Octubre, 2020). Tablas de volumen para *Tectona grandis L.f* en plantaciones del centro-occidente de México. *Revista Latinoamericana de Recursos Naturales*, 16(3), 96–102. Recuperado de:
https://www.itson.mx/publicaciones/rln/Documents/RLRN_2020_3_2.pdf.
22. Tamarit, J., De los Santos, H., Aldrete, A., Valdez, J., Ramírez, H., y Guerra, V. (Marzo, 2019). Sistema de crecimiento y rendimiento maderable para plantaciones de teca (*Tectona grandis L. f*) en Campeche, México. *Revista Madera y Bosques*, 25(3), 1–16. Recuperado de:
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-04712019000300214.
23. Telles, R., Cárdenas, M., Rodríguez, E., Calderón, O., y Pérez, J. (Mayo, 2018). Ajuste y selección de modelos matemáticos para predecir el volumen fustal de *Tectona grandis L. f* en Nuevo Urecho, Michoacán, México. *Revista Madera Bosques*, 24(3), 1–12. Recuperado:

https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-04712018000300215.

24. Uriel, E., (2013). *Regresión lineal múltiple: estimación y propiedades*. Valencia, España: Universidad de Valencia. Recuperado de: https://www.academia.edu/21768094/3_Regresi%C3%B3n_lineal_m%C3%BAltiple_estimaci%C3%B3n_y_propiedades.
25. Walpole, R., Myers, R., Myers, S., y Ye, K. (2012). *Probabilidad y estadística para ingeniería y ciencias*. México, D.F.: Pearson educación. Recuperado de: https://vereniciafunez94hotmail.files.wordpress.com/2014/08/8va-probabilidad-y-estadistica-para-ingenier-walpole_8.pdf.

APÉNDICES

Apéndice 1. Matriz de coherencia

Preguntas de investigación	Objetivos	Metodología	Resultados	Conclusiones	Recomendaciones
01. ¿Cuál es la tendencia de las variables diámetro a la altura al pecho (cm) y la altura total (m) que influyen en la estimación del volumen del árbol de Teca?	01. Describir la tendencia de las variables que influyen en la estimación del volumen del árbol de teca, mediante diagramas de dispersión y el coeficiente de correlación de Pearson, para inferir si dicha tendencia es lineal o no lineal.	Se elaboró un diagrama de dispersión en directa fuerte 3D para entre las variables que influyen en la tendencia estimación del volumen del árbol de teca, mediante diagramas de dispersión y el coeficiente de correlación de Pearson, para inferir si dicha tendencia es lineal o no lineal.	Existe asociación entre las variables analizadas, la relación que es de tipo polinómica donde se graficó eno sea novolumen totaldiagramas de parejas todaslineal, lo quedel árbol de tecadispersión y el	Se realizó la descripción para conocer la tendencia de las plantaciones en variables, diferentes sitios, estableciéndose las tendencias de que la relación de las variables diámetro a la altura al pecho, altura con el utilizando presentan unaanálisis de acuerdo a loasociación decorrelación para mediante elafirmado portipo cuadrática,comprender mejor análisis deCancino la cual esel comportamiento correlación se(2006), quiendirecta y fuerte,de las relaciones y evaluó la fuerza ygeneralmente que, la volumetría del dirección existente entre variables.	Se hace necesario continuar analizando las variables, estableciéndose las tendencias de diámetro a la altura al pecho, altura y volumen, altura con el utilizando presentan unaanálisis de acuerdo a loasociación decorrelación para mediante elafirmado portipo cuadrática,comprender mejor análisis deCancino la cual esel comportamiento correlación se(2006), quiendirecta y fuerte,de las relaciones y evaluó la fuerza ygeneralmente que, la volumetría del dirección existente entre variables.

Continuación del apéndice 1.

02. ¿Cuál es el porcentaje promedio de leña, troza mayor y troza menor que se obtiene de un árbol de teca?	02. Estimar el porcentaje promedio de troza comercial y leña que se obtiene del aprovechamiento de un árbol de teca, mediante la construcción de intervalos de confianza, para la planificación del manejo de la plantación.	En la plantación forestal de Teca, mediante muestreo selectivo se eligieron los árboles representativos para cada clase diamétrica, la cual cuenta con intervalos de 5 cm a partir de un diámetro de 20 cm hasta 50 cm de diámetro, se cortaron como mínimo 3 a 5 árboles por cada clase, y previo a la corte se le midió el Dap. Ya tumbado el árbol, el fuste se seccionó cada dos metros, se cubicó usando la fórmula de Smalian para obtener el volumen de troza comercial y leña, esta sumatoria de volúmenes representan el total de volumen, al relacionar cada volumen entre el total resultaron las proporciones por producto. Se construyen intervalos de confianza para estimar verdadero parámetro.	El árbol de teca por razones comerciales se secciona en los productos troza mayor, troza menor y leña, para el presente estudio se observó que en todas las clases diametrales el producto existente en mayor proporción es la troza menor (diámetros de 15.91 a 22.28 cm), y en menor proporción la leña o leña (diámetros menores a 15.91 cm), lo que es consistente con lo afirmado por (Mora y Hernández, 2007), en relación con las proporciones de mencionados productos.	Se estimó el porcentaje promedio de troza y leña del árbol de teca, estableciendo que corresponde a un volumen de troza del 89.41 % y 10.59 % de leña, valores que varían según la clase diamétrica, notándose que al incrementarse el diámetro disminuye el volumen de leña. Al diferenciar el árbol con fines comerciales, las proporciones de troza mayor en promedio corresponden a un 35.68 %, la troza menor un 54.24 %, y la leña un 10.08 %.	Se debe continuar realizando estimaciones del porcentaje promedio de los productos del árbol de Teca como troza menor, troza mayor y leña en plantaciones forestales de diferentes sitios para determinarlos con precisión, ya que existen variaciones entre clases diametales y posiblemente localidades diferentes, para que se permita planificar adecuadamente el manejo y aprovechamiento forestal en las diferentes regiones del país.
--	--	---	---	--	--

Continuación del apéndice 1.

<p>03. ¿Cuál es el modelo de mejor ajuste para estimar el volumen y diseñar la tabla de volumen que a nivel local permite planificar el aprovechamiento de la plantación de teca?</p>	<p>03. Seleccionar el modelo de mejor ajuste para estimar el volumen total del árbol de Teca, y permitir diseñar una tabla de volumen para una determinada plantación forestal.</p>	<p>La evaluación de los modelos se realizó usando CurveExpertProfessional, seguido se realizó la prueba de significancia para los coeficientes de todos los modelos, posteriormente se verificó los supuestos de los modelos de mejor ajuste, eligiendo finalmente el modelo, con esta expresión se diseñó la tabla de volumen de dos entradas, usando el DAP y la Altura para estimar el volumen.</p>	<p>Entre los modelos mejor evaluados destaca el modelo de Meyer modificado y el cuadrático, resultando un coeficiente de determinación ajustado de 92.78 % y 90 %, respectivamente, pero a pesar de estar mejor calificado el primero, se eligió finalmente el segundo, porque cumplió con todos los supuestos.</p>	<p>Se estableció el modelo de mejor ajuste para la predicción del volumen total de Teca, y el diseño de la tabla de volumen, siendo la expresión $V = -0.3545 - (-) 0.0444 * H + 0.000013 * D^2$</p>	<p>* H, la cual es la más parsimoniosa, sus coeficientes son significativos, y cumple con los supuestos. La validación de este modelo permitió que se elaborará la tabla de volumen que será de utilidad para la estimación de volumen de manera rápida y precisa, lo que facilitará el plan de manejo y aprovechamiento forestal.</p>
---	---	--	---	---	--

Continuación del apéndice 1.

<p>¿Qué modelo matemático es el que mejor se ajusta para predecir el volumen en metros cúbicos en árboles de Teca, a partir de las variables diámetro a la altura al pecho (cm) y la altura total (m).?</p>	<p>Evaluar y seleccionar mediante análisis de regresión múltiple el modelo matemático de mejor ajuste para estimar el volumen de árboles de teca (<i>Tectona grandis L.f.</i>) en plantaciones del departamento de Alta Verapaz, Guatemala.</p>	<p>Con software especializado y regresión múltiple se evaluaron los modelos propuestos para la estimación de volumen total de Teca, considerando la selección del modelo de mejor ajuste, proponiendo como variables independientes el DAP y la altura.</p>	<p>Al realizar estimaciones con el modelo seleccionado se obtiene mejor ajuste que con el modelo de FAO recomendado por el INAB (2015), en vista que mediante una prueba de hipótesis se determinó que no existe diferencia significativa entre el volumen real y el volumen estimado seleccionado, no así con el modelo de FAO que sí presenta diferencia significativa al 5 %. La tabla de volumen propuesta es útil para realizar estimaciones volumétricas de árboles comprendidos entre clases diamatrales de 20 a 50 centímetros, y de alturas totales de 10 a 30 metros , situación que indica, Rueda Sánchez et al (2020) que debe considerarse al hacer el uso de las tablas, así como su validación.</p>	<p>La evaluación de los modelos volumétricos mediante regresión múltiple permitió seleccionar el de mejor ajuste, el cual predice el volumen total del árbol de Teca con exactitud, esto porque al comparar el volumen estimado con el modelo con los valores reales de volumen, se determinó que no existe diferencia significativa.</p>	<p>Se sugiere el uso del modelo seleccionado para el presente estudio, el cual denomina cuadrático, para predecir el volumen total del árbol de Teca, porque es el de mejor ajuste y cumple con los supuestos, lo que garantiza predicciones cercanas a la realidad.</p>
---	---	---	--	---	--

Fuente: elaboración propia.

Apéndice 2. Cálculo de volumen total utilizando diferentes modelos

DAP (cm)	Altura (m)	Vol. Total real (m ³)	Modelo FAO	Modelo seleccionado
32.15	18.04	0.69	0.98	0.69
25.46	19.75	0.78	0.70	0.69
28.97	22.15	0.83	0.97	0.87
28.01	22.00	0.76	0.91	0.85
28.97	20.03	0.91	0.89	0.75
27.69	19.87	0.60	0.82	0.73
28.65	23.13	0.99	0.99	0.92
31.19	20.12	0.78	1.02	0.79
30.24	23.24	0.87	1.10	0.95
33.42	24.48	1.10	1.38	1.09
25.78	18.67	0.56	0.69	0.64
36.61	20.15	1.03	1.36	0.89
35.33	23.08	1.12	1.45	1.04
25.15	14.08	0.31	0.52	0.39
26.42	17.88	0.56	0.69	0.60
33.74	20.75	0.88	1.21	0.87
36.92	24.63	1.19	1.67	1.18
36.29	23.67	1.20	1.56	1.10
35.97	20.02	1.03	1.31	0.87
39.15	24.17	1.44	1.83	1.20
34.31	20.64	0.83	1.24	0.88
31.12	17.76	0.60	0.91	0.66
34.80	21.02	0.85	1.29	0.91
31.81	19.55	0.63	1.03	0.77
31.51	18.94	0.61	0.98	0.73
32.23	20.12	0.68	1.08	0.81
33.22	21.25	0.71	1.20	0.89
33.66	19.48	0.74	1.13	0.80
29.10	16.96	0.51	0.78	0.59
37.80	20.99	0.97	1.50	0.97
36.30	19.60	0.97	1.31	0.85
28.29	17.59	0.46	0.76	0.61
31.48	18.90	0.62	0.98	0.73

Continuación del apéndice 2.

DAP (cm)	Altura (m)	Vol. Total real (m^3)	Modelo FAO	Modelo seleccionado
37.86	23.65	1.03	1.68	1.14
35.08	18.38	0.85	1.16	0.76
35.90	19.28	0.89	1.26	0.82
34.85	18.51	0.82	1.15	0.76
29.54	17.58	0.50	0.82	0.63
25.93	16.07	0.38	0.61	0.50
24.05	15.56	0.30	0.53	0.45
20.30	13.19	0.19	0.36	0.30
24.25	15.14	0.30	0.52	0.43
26.76	15.84	0.39	0.64	0.50
29.18	17.08	0.48	0.78	0.59
40.25	23.50	1.32	1.88	1.18
40.65	24.35	1.41	1.98	1.25
41.38	25.40	1.36	2.13	1.34
42.35	23.30	1.53	2.05	1.22
44.98	27.52	1.61	2.70	1.59
45.20	27.90	1.40	2.76	1.63
46.15	26.50	1.65	2.73	1.56
47.78	27.10	1.54	2.99	1.65
48.54	28.30	1.58	3.21	1.77
49.33	29.10	1.95	3.40	1.86

Fuente: elaboración propia.