UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS AGRARIAS

ESCUELA PROFESIONAL DE INGENIERÍA FORESTAL Y AMBIENTAL

Estimación del volumen de árboles de *Pinus patula* en plantaciones forestales de la provincia de Chota, mediante el uso de algoritmos de inteligencia artificial

TESIS PARA OPTAR EL TITULO PROFESIONAL DE

INGENIERO FORESTAL Y AMBIENTAL

AUTOR

Flor Erlita Benavides Huamán

ASESOR

M.Sc. Duberli Geomar Elera Gonzales

M.Sc. Duberlí Geomar Elera Gonzáles
Docente UNACH

CHOTA-PERÚ

AGOSTO, 2025

Universidad Nacional Autónoma de Chota

FACULTAD DE CIENCIAS AGRARIAS

Unidad de Investigación de la Escuela Profesional de Ingeniería Forestal y Ambiental

RESOLUCIÓN DE COORDINACIÓN N° 251-2025-FCA/UNACH

"Año de la recuperación y consolidación de la economía peruana"

CONSTANCIA DE ORIGINALIDAD

El que suscribe, Director de la Unidad de Investigación de la Escuela Profesional de Ingeniería Forestal y Ambiental de la Facultad de Ciencias Agrarias de la Universidad Nacional Autónoma de Chota, **hace constar** que la tesis de investigación Titulada **"Estimación del volumen de árboles de *Pinus patula* en plantaciones forestales de la provincia de Chota, mediante el uso de algoritmos de inteligencia artificial"**; ejecutado por la Bachiller **Flor Erlita Benavides Huamán** de la Escuela Profesional de Ingeniería Forestal y Ambiental, **asesorado por el M. Sc. Duberli Geomar Elera Gonzales (Asesor)**; presenta un **ÍNDICE DE SIMILITUD DEL 10%**, sin incluir bibliografía; por lo tanto, cumple con el criterio de evaluación de originalidad establecido en el REGLAMENTO DE GRADOS Y TÍTULOS DE LA UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA aprobado mediante RESOLUCIÓN DE COMISIÓN ORGANIZADORA N°120-2022-UNACH.

Se expide la presente, a petición de la parte interesada para los fines que estime conveniente.

Chota, 14 de octubre de 2025.

Atentamente

Dr. Jim Jairo Villena Velásquez
Unidad de Investigaciónde EPIFA - UNACH

Flor Benavides

Flor Erlita Benavides Huamán

Flor Erlita Benavides Huamán
Flor Erlita Benavides Huamán
Universidad Nacional Autonoma de Chota

Detalles del documento

Identificador de la entrega
trn:oid:::1:3372613677

Fecha de entrega
14 oct 2025, 12:46 a.m. GMT-5

Fecha de descarga
14 oct 2025, 1:02 a.m. GMT-5

Nombre del archivo
nforme_Final_de_Tesis-_Flor_Erlita_Benavides_Huam_n_v1_222_1.odt

Tamaño del archivo
39.6 MB

118 páginas

15.564 palabras

93.078 caracteres

10% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Coincidencias menores (menos de 17 palabras)

Fuentes principales

5%	Fuentes de Internet
1%	Publicaciones
8%	Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.° de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

Fuentes principales

5% Fuentes de Internet
1% Publicaciones
8% Trabajos entregados (trabajos del estudiante)

Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Trabajos del estudiante Universidad Nacional Autonoma de Chota	5%
2	Internet repositorio.unach.edu.pe	1%
3	Internet alicia.concytec.gob.pe	<1%
4	Internet myb.ojs.inecol.mx	<1%
5	Trabajos del estudiante Integración Moodle Presencial 4.3	<1%
6	Internet es.scribd.com	<1%
7	Internet repositorio.uta.edu.ec	<1%
8	Internet cienciasforestales.inifap.gob.mx	<1%
9	Trabajos del estudiante Universidad TecMilenio	<1%
10	Internet de.slideshare.net	<1%
11	Internet repositorio.unprg.edu.pe	<1%

12 Trabajos del estudiante
Universidad Ricardo Palma <1%

13 Trabajos del estudiante
Universidad Santo Tomas <1%

14 Trabajos del estudiante
Universidad Internacional de la Rioja <1%

15 Internet
www.lillo.org.ar <1%

16 Internet
www.repositorio.usac.edu.gt <1%

17 Internet
revistas.udem.edu.co <1%

18 Internet
www.scielo.org.mx <1%

Anexo 01:

ACTA DE SUSTENTACIÓN DE TESIS

REG. N° 034 -2025-FCA

Siendo las ....9.... horas, del día ....14.... de ....agosto.... de ..2025.., los miembros del Jurado de Tesis titulada: Estimación del volumen de árboles de Pinus patula en plantaciones forestales de la provincia de Chota, mediante el uso de algoritmos de inteligencia artificial ....., integrado por:

1. Mtr. Denisse Milagros Alva Mendoza ..... Presidente
2. M. Sc. Yuli Anabel Chávez Juanito ..... Secretario
3. Dr. Jim Jairo Villena Velásquez ..... Vocal

Sustentada de manera presencial (X), virtual ( ) por Bach. Flor Elita Benavides Huamán ....., con la finalidad de obtener el Título Profesional de (Ingeniero Agroindustrial/Ingeniero Forestal y Ambiental) Terminada la sustentación, con las preguntas formuladas por los integrantes del Jurado y las respuestas otorgadas por el graduando, luego de deliberar, acuerda… (Aprobar, desaprobar) … la tesis, calificándola con la nota de: (.....Dieciseis.....), se eleva la presente Acta al Coordinador de la Facultad de Ciencias Agrarias, a fin de que se le declare EXPEDITO para conferirle el correspondiente título profesional

Colpa Matara, ..14.. de .....agosto..... del 2025

.....
Presidente
Mtr. Denisse Milagros Alva Mendoza

.....
Secretario
M.Sc. Yuli Anabel Chávez Juanito

.....
Vocal
Dr. Jim Jairo Villena Velásquez

.....
Asesor

DEDICATORIA

Con profundo agradecimiento y afecto, dedico este trabajo a mis padres, hermanas y a toda mi familia, quienes han sido el sostén esencial en mi vida. Su amor incondicional y los sacrificios realizados para que yo pudiera alcanzar una educación superior son el motor de mi esfuerzo. Esta meta es tan suya como la mía, y merecen este logro y mucho más.

A mis compañeros y amigos, gracias por confiar en mí y acompañarme en este camino. Su amistad, solidaridad y constante apoyo emocional han sido fundamentales para mantener vuva mi motivación y perseverancia durante la etapa universitaria, convirtiéndola en una experiencia inolvidable.

A todos los docentes que formaron parte de mi proceso académico, mi reconocimiento por su valiosa contribución en mi formación profesional. Gracias por transmitirme conocimientos y valores que hoy me permiten avanzar con responsabilidad y compromiso en mi carrera.

A todos los que de una u otra manera estuvieron pendientes de mi trabajo para darme apoyo y ánimo para iniciar y culminar esta investigación.

Flor Erlita Benavides Huamán

AGRADECIMIENTOS

En primer lugar, expreso mi gratitud a Dios por brindarme la oportunidad de culminar este proyecto y por darme la fortaleza necesaria para superar cada obstáculo que se presentó al largo de esta investigación.

A mis padres, Gonzalo y Rosa, quienes han sido el motor de mis sueños y esperanzas. Su compañía constante en los momentos más exigentes de mi formación, tanto en los días como en las noches de estudio, ha sido invaluable. Han sido mis guías de vida, y esta meta alcanzada es también suya. Me siento profundamente orgullosa de contar con usted en este momento tan significativo.

A mis hermanas, abuela, tíos C y M, y a toda mi familia, por el respaldo incondicional que me ofrecieron durante mi trayectoria académica. Su fe en mí, incluso en los momentos más difíciles, ha sido el fundamento de este logro.

Al MSc. Duberli Geomar Elera Gonzales, mi sincero agradecimiento por su paciencia, constancia y sabios consejos. Su orientación fue clave en los momentos de incertidumbre, cuando las ideas parecían esquivas. Sus palabras de aliento llegaron justo cuando más las necesitaba, y fueron esenciales para concretar este trabajo.

A mis amigos KBC, AVR, ABS, ABC, LDH, EMG y a mis compañeros, gracias por su apoyo constante y por estar presentes en cada etapa de este proceso.

A la Universidad Nacional Autónoma de Chota y a la Escuela Profesional de Ingeniería Forestal y Ambiental, por acogido durante cinco años de mi proceso formativo, y por facilitarme el entorno académico necesario para llevar a cabo el presente proyecto de investigación.

Muchas gracias a todos

Flor Erlita Benavides Huamán

ÍNDICE

RESUMEN 17

ABSTRACT 19

CAPÍTULO I 20

INTRODUCCIÓN 20

1.1. Planteamiento del problema 20

1.2. Formulación del problema 23

1.2.1. Problema general 23

1.3. Justificación 23

1.4. Objetivos de la investigación 25

1.4.1. Objetivo general 25

1.4.2. Objetivos específicos 25

CAPÍTULO II 26

MARCO TEÓRICO 26

2.1 Antecedentes 26

2.2 Bases teórico–científicas 28

2.1.1. Modelamiento de volumen de árboles 28

2.3 Marco conceptual 34

2.2. Hipótesis 37

2.3. Operacionalización de variables 37

CAPÍTULO III 39

MARCO METODOLÓGICO 39

3.1. Tipo y nivel de investigación 39

3.2. Diseño de la investigación 39

3.3. Métodos de investigación 39

3.3.1. Área de estudio 39

3.3.2. Trabajo de campo 41

a. Localización y delimitación de parcelas 41

b. Inventario 42

c. Tabulación de datos y distribución diamétrica 42

d. Cubicación rigurosa de árboles 43

3.3.3. Ajuste y selección de modelos 44

3.4. Población, muestra y muestreo 46

3.4.1. Población 46

3.4.2. Muestra 47

3.4.3. Muestreo 47

3.5. Técnicas e instrumentos de recolección de datos 47

3.5.1. Técnicas de recolección de datos 47

3.5.2. Instrumentos para recolección de datos 47

3.6. Técnicas de procesamiento y análisis de datos 48

3.7. Aspectos éticos 51

CAPÍTULO IV 52

RESULTADOS Y DISCUSIÓN 52

4.1. Información general de las plantaciones inventariadas 52

4.2. Contrastación de hipótesis 64

4.3. Discusión de resultados 64

Información general 64

CAPÍTULO V 69

CONCLUSIONES Y RECOMENDACIONES 69

5.1 Conclusiones 69

5.2 Recomendaciones 70

CAPÍTULO VI 71

REFERENCIAS 71

CAPÍTULO VII. ANEXO 83

ÍNDICE DE TABLAS

Tabla 1 Operacionalización de variables 37

Tabla 2 Formato para registrar de los árboles inventariados 42

Tabla 3 Base de datos para el modelamiento del volumen individual de árboles 48

Tabla 4 Estadísticas descriptivas de las variables dasométricas del inventario forestal 52

Tabla 5 Resultados del test de normalidad de Kolmogórov-Smirnov en datos de inventario 54

Tabla 6 Resumen de las variables dasométricas de los árboles cubicados 55

Tabla 7 Resultados del test de normalidad de Kolmogorov-Smirnov en datos de árboles muestra 56

Tabla 8 Estadísticos de bondad de ajuste de modelos entrenados 60

ÍNDICE DE FIGURAS

Figura 1 Representación de la cubicación rigurosa del fuste de un árbol. 29

Figura 2 Estructura de una red neuronal artificial 33

Figura 3 Ubicación de área de estudio, en la provincia de Chota, zona central de la región Cajamarca, norte de Perú. 40

Figura 4 Delimitación de parcelas de inventario forestal, con área de 400 m^2 (20m x 20m), en las plantaciones de Pinus patula 41

Figura 5 Gráfico de frecuencias de los árboles inventariados, a) distribución de frecuencias del dap, b) distribución de frecuencias de h. 53

Figura 6 Normal Q-Q plot para los datos de dap y h provenientes del inventario forestal 54

Figura 7 Gráfico de frecuencias de 122 árboles cubicados rigurosamente a) diámetro de frecuencia del dap, b) distribución de frecuencias de h 55

Figura 8 Normal Q-Q plot para los datos de dap y h provenientes del inventario forestal 57

Figura 9 Correlación de Pearson de la variable dependiente (volumen) y variables independientes analizadas. 58

Figura 10 Vista tridimensional de la relación volumen, dap y altura total de los árboles cubicados 59

Figura 11 Comparación de las estadísticas de bondad de ajuste calculados con datos de entrenamiento (barra negra) y datos de validación (barra blanca) de los cuatro modelos ajustados. 61

Figura 12 Comparación de las estadísticas de bondad de ajuste calculados con datos de entrenamiento y datos de validación de los cuatro modelos ajustados. 63

RESUMEN

La estimación precisa del volumen individual de los árboles es esencial para monitorear la productividad de las plantaciones forestales y mejorar las decisiones sobre el manejo forestal. La presente investigación se llevó a cabo en la Provincia de Chota, en plantaciones ubicadas entre los 2367 hasta 3478 ms.n.m., en la parte central del departamento de Cajamarca. El objetivo principal fue determinar la eficiencia de tres algoritmos de inteligencia artificial para estimar el volumen de árboles de *Pinus patula* en plantaciones forestales; los objetivos específicos fueron: 1) determinar el volumen individual de árboles de *Pinus patula* de plantaciones forestales a través del método de cubicación rigurosa, 2) determinar las variables de inventario forestal de mayor importancia para el modelamiento del volumen individual de árboles de *Pinus patula* utilizando algoritmos de inteligencia artificial. Se utilizaron datos de 21 parcelas de medición, con un total de 1006 registrados de diámetro a la altura del pecho y altura total. El modelamiento del volumen se realizó con datos de 122 árboles cubicados rigurosamente. Las variables más importantes para estimar el volumen de árboles fueron el diámetro, altura, altura dominante y altitud. Los algoritmos Random Forest, Support Vector Machine y Artificial Neural Network, junto al modelo de Schumacher & Hall fueron utilizados. El algoritmo Artificial Neural Network mostró los mejores resultados, con $R^2 = 0.9523$ y RSME= 0.0881, sin embargo, no superó al modelo S&H, que obtuvo un $R^2 = 0.9588$ y RSME= 0.0799 pero a la vez, no existió diferencias estadísticas entre estos dos, sugiriendo que, tanto el algoritmo Artificial Neural Network y el modelo Schumacher & Hall pueden ser utilizados para estimar el volumen individual de *Pinus patula* procedentes de plantaciones forestales de la Provincia de Chota.

Palabras claves: Cubicación rigurosa, Redes Neuronales, inteligencia artificial, modelo volumétrico, manejo forestal, modelamiento forestal.

ABSTRACT

Accurate estimation of individual tree volume is essential for monitoring the productivity of forest plantations and improving forest management decisions. This research was conducted in the province of Chota, in plantations located between 2,367 and 3,478 meters above sea level, in the central part of the department of Cajamarca. The main objective was to determine the efficiency of three artificial intelligence algorithms for estimating the volume of Pinus patula trees in forest plantations. the specific objectives were: 1) to determine the individual volume of Pinus patula trees in forest plantations using the rigorous cubing method, 2) to determine the most important forest inventory variables for modeling the individual volume of Pinus patula trees using artificial intelligence algorithms. Data from 21 measurement plots were used, with a total of 1,006 trees recorded for diameter at breast height and total height. Volume modeling was performed using data from 122 rigorously measured trees. The most important variables for estimating tree volume were diameter, height, dominant height, and altitude. The Random Forest, Support Vector Machine, and Artificial Neural Network algorithms were used, along with the Schumacher & Hall model. The Artificial Neural Network algorithm showed the best results, with R2 = 0.9523 and RSME = 0.0881, however, it did not outperform the S&H model, which obtained an R2 = 0.9588 and RSME = 0.0799. At the same time, there were no statistical differences between the two, suggesting that both the Artificial Neural Network algorithm and the Schumacher & Hall model can be used to estimate the individual volume of Pinus patula from forest plantations in the province of Chota.

Keywords: Rigorous cubication, Neural Networks, artificial intelligence, volumetric model, forest management, forest modeling.

CAPÍTULO I

INTRODUCCIÓN

1.1. Planteamiento del problema

Según Sánchez (2019), uno de los vacíos más relevantes en la investigación forestal en el Perú es la limitada disponibilidad de información relacionada con el valor económico de los tocones y del suelo en el que se establecen las plantaciones forestales. Esta carencia representa un obstáculo para la valoración integral de los recursos forestales y para la toma de decisiones en el manejo sostenible de los ecosistemas intervenidos. En ese sentido Zamora et al. (2004) destacan que, para realizar una adquisición eficiente de madera en pie, es indispensable contar con estimaciones precisas del volumen y la calidad de los productos que se obtendrán tras el proceso de aserrado ya que esta precisión es clave para mantener la competitividad de los productos forestales. Por su parte, Guariguata et al. (2017) destacan la importancia de contar con un marco sistemático que facilite la producción y divulgación de conocimientos técnicos orientados a la industria y a los distintos sectores del mercado, sumada la insuficiencia de información clara y actualizada, constituye una barrera relevante que desincentiva tanto a inversionistas nacionales e internacionales como a los gobiernos locales y regionales a invertir recursos públicos en el desarrollo de plantaciones comerciales, lo cual repercute negativamente en el avance del sector forestal.

Cajamarca se posiciona entre las regiones con mayor superficie destinada a plantaciones forestales del Perú. Actualmente, se ha iniciado el aprovechamiento de las primeras áreas establecidas, aunque persiste una carencia de información precisa sobre el volumen real de las reservas forestales en pie. Esta falta de datos técnicos impide que los propietarios obtengan una

valorización justa por sus recursos, afectando directamente la rentabilidad y sostenibilidad de las actividades forestales (Orrillo, 2019).

Hynynem (2011) sostiene que los modelos de crecimiento y rendimiento son herramientas valiosas para la toma de decisiones, ya que proporcionan información sobre el desarrollo forestal potencial. Un ejemplo de ello es el modelamiento volumétrico, el cual emplea herramientas estadísticas para estimar el volumen de los árboles. Esta información resulta fundamental para calcular las existencias maderables presentes en un rodal y proyectar la distribución de los productos forestales antes de realizar las actividades de corta (Santiago et al., 2020); en este mismo contexto Cruz et al. (2008) proponen la implementación de ahusamiento y volumen, basado en una estructura geométrica común y parámetros compartidos. Esta propuesta facilita la asignación proporcional del volumen entre los distintos productos forestales, contribuyendo a la eliminación de las discrepancias que suelen presentarse cuando se ajustan por separado los modelos de volumen total y comercial.

La creciente aplicación de la Inteligencia Artificial (IA) como herramienta para abordar problemáticas en diversas disciplinas ha despertado un notable interés en los últimos años (Ortega & Cruz, 2020). Según Reaño (2020), esta tecnología constituye un avance innovador de gran relevancia, con impacto transversal en diversos sectores como las telecomunicaciones, el marketing, la economía, la industria de las tecnologías de la información, así como en areas clave como la medicina y las ciencias naturales. Los algoritmos de IA han evolucionado de manera constante, perfeccionando sus estrategias de resolución y generando beneficios sin precedentes en términos de eficiencia, precisión y alcance.

Una evaluación precisa de la productividad forestal es esencial para una gestión eficiente y sostenible de los recursos forestales. No obstante, el desarrollo de modelos altamente precisos

enfrenta desafíos significativos, debido a la necesidad de recopilar, procesar e integrar grandes volúmenes de datos y de alta dimensión. Este proceso requiere técnicas avanzadas de procesamiento y análisis con capacidad de alto rendimiento computacional.

En este contexto, los avances recientes en sistemas de aprendizaje automático basados en la nube representan una solución efectiva. Estas tecnologías permiten el procesamiento ágil y eficiente de datos forestales a gran escala, optimizando la precisión de las estimaciones y mejorando la toma de decisiones en la gestión forestal (Sotomayor et al., 2023)

Diversas investigaciones han evidenciado el potencial de los algoritmos de inteligencia artificial como herramientas eficaces para modelar el crecimiento de masas forestales, incluyendo aquellos enfoques orientados a la estimación de la producción forestal. Estos avances respaldan la aplicabilidad de la IA en el análisis y proyección de dinámicas forestales, optimizando la toma de decisiones en el manejo sostenible de los recursos (Martins et al., 2021; Boechat et al; 2004), altura total (Goycochea et al., 2022), diámetro y altura (Correia et al., 2018), índice de sitio (Pinno et al., 2012) , biomasa (Wang et al., 2016) y volumen del fuste (Özçelik et al., 2010; Zhu et al., 2021).

De acuerdo con Aarón et al. (2012), las plantaciones de *Pinus patula* establecidas en zonas altoandinas desempeñan un papel importante en la restauración de sueños erosionados y degradados. Además, estas especies contribuyen al mejoramiento de la calidad del suelo mediante la incorporación progresiva de materia orgánica, producto de la caída natural de hojas, ramas y acículas. En concordancia, un estudio realizado en el distrito de Chalamarca señala que las plantaciones de *P. patula* presentan una acumulación progresiva de biomasa aérea: a los cinco años alcanzan 15,84 t/ha, la de diez años 32,53 y t/ha y la de quince años 51,56 t/ha. La mayor proporción de esta biomasa se concentra en el fuste, seguida por el follaje y, en menor medida en as rama (Mirez, 2021).

Este proyecto propone la aplicación de algoritmos de inteligencia artificial para estimar el volumen individual de árboles de *Pinus patula*, utilizando como base datos obtenidos a partir de procedimientos de cubicación rigurosa precisos en plantaciones forestales situados en la provincia de Chota, Perú. Está metodología busca mejorar la precisión en la cuantificación del recurso forestal, contribuyendo a una gestión más eficiente y sostenible.

1.2. Formulación del problema

1.2.1. Problema general

¿Es posible estimar el volumen individual de árboles de *Pinus patula* de plantaciones forestales de la provincia de Chota, mediante el uso de los algoritmos de inteligencia artificial?

1.3. Justificación

En el sector forestal, existe una larga trayectoria en la parametrización y aplicación de modelos orientados a cuantificar y predecir el crecimiento y rendimiento de las masas forestales, considerados elementos clave para optimizar las prácticas silvícolas y promover un manejo forestal sostenible. Sin embargo, continúa siendo imperativo el desarrollo de herramientas biométricas más eficaces que contribuyan a optimizar la planificación y el aprovechamiento sostenible de los recursos forestales. Actualmente, existe una amplia variedad de modelos con aplicaciones específicas, que abarcan desde estimaciones de volumen y ahusamiento, hasta sistemas compatibles ahusamiento-volumen. También se incluyen modelos altura-diámetro, de clasificación de la productividad de terrenos forestales, de biomasa, carbono y contenido nutrimental de ecosistemas, así como modelos para el control de la densidad, distribución diamétrica y de árbol individual (García et al., 2020).

Esa alta disponibilidad de modelos ha sido rápidamente complementada con el uso de nuevas tecnologías, entre ellas, los sistemas de visión por computadora que ya se usan en el monitoreo y planificación de las operaciones forestales permitiendo incrementar la competitividad en el mercado (Hildt et al., 2020).

En el manejo de una operación de plantaciones forestales, contar con proyecciones confiables sobre el crecimiento y la producción de madera resulta esencial. Estas estimaciones dependen principalmente de dos factores: i) la medición de variables dendrometrías como el diámetro, la altura, el área basal, el volumen, el índice de sitio, la distribución diamétrica y la edad de los rodales; ii) la aplicación de modelos que permitan estimar dichas variables, facilitando la predicción del estado actual y futuro de los bosques bajo manejo. Aunque en muchos casos se emplean modelos simplificados que describen la evolución de una variable en el tiempo, en escenarios de manejo forestal intensivo se requiere el uso de sistemas avanzados de modelación de crecimiento y rendimiento, capaces de simular el comportamiento de los rodales bajo distintos esquemas de gestión y prácticas silvícolas (Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), 2010) ; García et al., 2020).

Para conocer las características de crecimiento de los rodales, las variables previamente mencionadas son obtenidas mediante la medición sistemática en parcelas permanentes, complementada con el uso de herramientas analíticas como la regresión y técnicas de inteligencia artificial, entre ellas las redes neuronales artificiales (ARN), y las máquinas de vector de soporte (MSV). Estas metodologías modernas ofrecen ventajas significativas en términos de velocidad de procesamiento, precisión en las estimaciones y reducción de costos operativos (Dávila, 2018).

Disponer con información clara y precisa sobre el crecimiento y el comportamiento de distintas especies arbóreas en diversas condiciones de crecimiento de manera clara y concisa,

significativamente favorece a una toma de decisiones más eficiente en cada una de las etapas del manejo forestal Hynynem (2011). En el contexto de las plantaciones de *Pinus patula* ubicadas en la provincia de Chota, esta información proporcionará información útil para el conocimiento del stock actual maderables y la implementación de futuros proyectos de plantaciones forestales. En este marco, el estudio propone la aplicación de diversos algoritmos de inteligencia artificial con el objetivo de estimar el volumen individual de los árboles, contribuyendo así el desarrollo de una herramienta técnica que facilite la toma de decisiones por parte de los responsables del manejo de bosques establecidos con esta especie.

1.4. Objetivos de la investigación

1.4.1. Objetivo general

– Determinar la eficiencia de algoritmos de inteligencia artificial para estimar el volumen de árboles de *Pinus patula* en plantaciones forestales en la provincia de Chota.

1.4.2. Objetivos específicos

– Determinar el volumen individual de árboles de *Pinus patula* de plantaciones forestales a través del método de cubicación rigurosa.

– Determinar las variables de inventario forestal de mayor importancia para el modelamiento del volumen individual de árboles de *Pinus patula* utilizando algoritmos de inteligencia artificial.

– Aplicar diferentes algoritmos de inteligencia artificial y seleccionar el mejor para estimar el volumen individual de árboles de *Pinus patula* de plantaciones forestales de la provincia de Chota.

CAPÍTULO II

MARCO TEÓRICO

2.1 Antecedentes

Şenyurt et al. (2020) investigaron el uso de redes neuronales artificiales (ANN) como herramienta para modelar la relación entre el diámetro del tocón y el diámetro a la altura del pecho (DAP), con el objetivo de evaluar la capacidad predictiva de este enfoque en la estimación del DAP. Su estudio destaca el potencial de las ANN para mejorar la precisión en contextos donde las mediciones directas del DAP no son viables.

Para llevar a cabo el estudio, se recopilaron 583 pares de daros correspondientes al diámetro a la altura del pecho (DAP) y el diámetro del tocón, distribuidos en 61 parcelas de muestreo de *Pinus nigra subsp. Pallasiana* (pino de Crimea), ubicadas en el bosque experimental de ÇAKÜ (Turquía). Los resultados del estudio evidencian que el potencial de las ANN para estimar con precisión esta relación, destacando su utilidad en contextos donde el DAP no puede ser medido directamente.

La creciente demanda de madera y sus derivados ha convertido a las plantaciones forestales industriales en una actividad de alcance internacional. En el caso del Perú, desde inicios del siglo XXI se han impulsado diversas iniciativas orientadas a fomentar el desarrollo de este tipo de plantaciones. De acuerdo con (Guariguata et al., 2017), se calcula que en el país existen entre 1 y 2 millones de hectáreas con potencial para plantaciones forestales, dentro de un conjunto más amplio de aproximadamente 10,5 millones de hectáreas que aún no han sido georreferenciadas. La mayor parte de estas áreas se localizan en las cuencas amazónicas andinas, lo que representa una oportunidad estratégica para el desarrollo forestal sostenible. Un ejemplo destacado de este modelo

exitoso en el país es la plantación forestal de Porcón, en Cajamarca, donde se proyecta transformar importantes recursos forestales en una zona con alto valor agregado (Llerena et al., 2016).

Martins et al. (2019) realizaron una comparación entre dos enfoques metodológicos aplicados a la estimación del volumen de eucalipto en Brasil. En dicho análisis, tanto el sistema de inferencia difusa basado en redes adaptativas (ANFIS) como el algoritmo de Random Forest demostraron una alta capacidad predictiva, permitiendo realizar estimaciones precisas sobre el comportamiento volumétrico del bosque. Ambos métodos demostraron ser herramientas complementarias con alto potencial para apoyar la gestión forestal, al ofrecer alternativas confiables en el análisis y pronóstico de variables en la gestión forestal.

El volumen de los árboles puede ser calculado mediante modelos de regresión lineal de menor complejidad, como propuesto por Revilla et al. (2021) en su estudio sobre ejemplares de bolaina **(***Guazuma crinita*) de 31 meses de edad en la Amazonia peruana. En dicha investigación, se ajustó el modelo de regresión lineal de Schumacher y Hall, utilizando como variables predictoras el diámetro a la altura del pecho (DAP), medido a 1,3 metros sobre el nivel del suelo, y a la altura total del árbol (H), con el fin de estimar el volumen (V).

Marques et al. (2014) demostraron que las redes neuronales artificiales pueden emplearse eficazmente para estimar el volumen de árboles con o sin corteza en plantaciones de eucalipto; facilitando la formulación de ecuaciones volumétricas con un costo operativo reducido. Por su parte, Tian et al. (2022) y Martins et al. (2021) señalan que el modelo de Random Forest presenta una alta confiabilidad estadística y resulta útil para predecir el crecimiento y el rendimiento de rodales forestales. Sin embargo, se ha observado que el desempeño de este modelo fue superior al estimar el volumen de árboles en bosques naturales en comparación con plantaciones forestales, debido a ciertas limitaciones inherentes al algoritmo Random Forest. Aun asi, la implementación

de estas técnicas puede mejorar significativamente la precisión en las estimaciones de volumen forestal.

Jevšenak & Skudnik (2021) aplicaron el algortimo Random Forest (RF) para estimar y analizar el incremento del área basal (BAI) en especies como *Pinus sylvestris*, *Abies alba*, *Fagus sylvatica* y *Quercus* sp. El modelo fue entrenado utilizando un conjunto de 18 variables independientes y un total de 15,580 registros provenientes del inventario forestal nacional (NFI) de Eslovenia. Los resultados evidenciaron una alta precisión en las predicciones, lo que confirma la eficacia de esta técnica como herramienta avanzada para el análisis forestal.

Entre los estudios más resientes sobre uso de técnicas de inteligencia artificial en modelamiento forestal, destaca el desarrollado por Goycochea et al., (2022) quienes utilizaron redes neurales profundas para estimar altura total de *Guazuma crinita* en plantaciones de la Amazonia peruana en función de variables ambientales.

2.2 Bases teórico–científicas

2.1.1. Modelamiento de volumen de árboles

El modelamiento volumétrico de los árboles se basa en la estimación del volumen individual o grupal mediante la aplicación de modelos o ecuaciones matemáticas, las cuales emplean como variables explicativas dimensiones fácilmente medibles, como el diámetro y la altura (Campos & Leite, 2007).

La importancia de obtener estimaciones precisas del volumen arbóreo en el contexto del manejo forestal ha impulsado a los profesionales del sector a emplear modelos matemáticos previamente desarrollados. Estos modelos permiten calcular de manera eficiente tanto el volumen

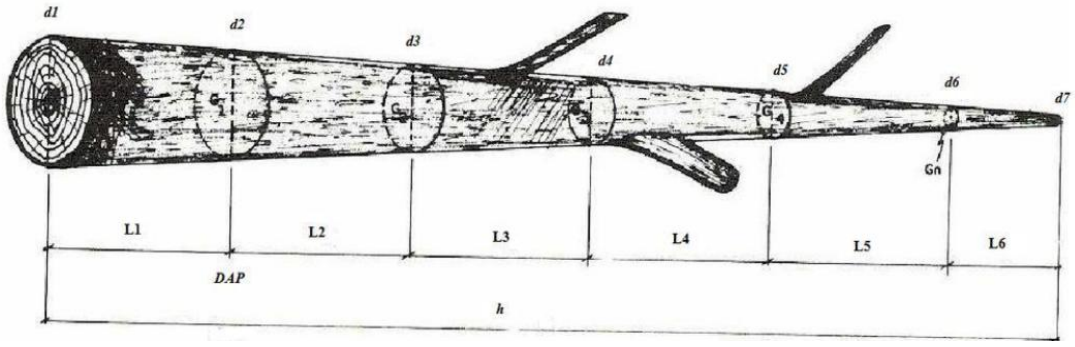
total como el volumen comercial de los árboles, utilizando variables de fácil medición como la altura total y el diámetro a la altura del pecho (DAP) (Galán et al., 2008).

2.1.2. Cubicación rigurosa de árboles

Los modelos matemáticos utilizando datos del volumen calculados por la medición completa de fustes de muestras representativas de árboles talados y cubicados de forma rigurosa. La cubicación rigurosa permite estimar el volumen total del fuste a través de la suma de los volúmenes de sus secciones individuales, determinados a partir de la aplicación de las expresiones matemáticas de Huber, Smalian o Newton (Campos & Leite 2017); como se muestra en la siguiente figura:

Figura 1

Representación de la cubicación rigurosa del fuste de un árbol.



2.1.3. Inteligencia artificial aplicada al modelamiento forestal

Guirado y Martínez (2021) señalan que existen estrategias basadas en inteligencia artificial (IA), las cuales se dividen en métodos supervisados y no supervisados. Los métodos supervisados son los más comunes y se caracterizan por requerir una base de datos etiquetada proporcionada por el usuario para entrenar el modelo. Dentro del aprendizaje automático, los enfoques supervisados incluyen algoritmos como los K-vecinos más cercanos, Random Forest (Rf), Support Vector Machine, perceptrones multicapa y redes neuronales (redes neuronales convolucionales) y redes neuronales.

Poco a poco, el uso de la IA va abriéndose camino en el sector forestal. Encontramos desarrollos desde escalas de territorio hasta niveles celulares, tanto ligados a la gestión forestal como en desarrollos industriales. A nivel nacional, algunas herramientas permiten identificar daños forestales por factores abióticos, como heladas tardías (Rodríguez et al., 2021).

El uso de la inteligencia artificial puede ayudar a realizar un inventario forestal en tiempo real y detectar cambios en los usos del suelo (cortas, incendios, plagas, etc.) (Rodríguez et al., 2021). De la misma manera menciona que las aplicaciones también sirven para el contexto urbano, donde la visión artificial permite identificar, geolocalizar y etiquetar arbolado urbano con árboles. La Inteligencia Artificial ya es una realidad en el sector forestal y nos ayuda en distintas escalas de trabajo.

2.1.4. Random Forest como método de regresión.

El modelo de bosque aleatorio se compone de múltiples árboles de regresión individuales, cada uno entrenando a partir de muestras generadas mediante técnicas de bootstrapping sobre el conjunto de datos original (Amat, J., 2020). Al concluir el proceso de entrenamiento, se obtienen predicciones independientes de cada árbol, luego se combinan para generar una estimación global más robusta y precisa.

La regresión de Bosque Aleatorio (RFR) representa una técnica flexible y eficaz dentro del aprendizaje automático, con aplicaciones en el ambiento forestal. Se ha utilizado e tareas como la estimación del diámetro del tronco de los árboles y el cálculo de la biomasa aérea (Diamantopoulou, 2022). Esta metodología permite mejorar la precisión de resultados, siendo una herramienta valiosa para la gestión y el análisis de recursos naturales.

2.1.5. Support Vector Machine para regresión, aplicaciones en el sector forestal.

Es un método utilizado en clasificación y regresión. Su objetivo es garantizar que los clústeres se separen mediante un vector de soporte que se ubique lo más alejado posible de las muestras, que de las propias muestras. Para ello, las observaciones se representan espacialmente a través de un hiperplano que maximiza la separación entre las clases.

La regresión mediante máquinas de vectores de soporte (SVM) se ha convertido en una potente herramienta para estimar parámetros forestales, especialmente en el contexto de los datos obtenidos mediante escaneo láser aéreo y teledetección (Monnet et al., 2011). Este método resulta especialmente ventajoso en entornos forestales heterogéneos, donde los modelos de regresión tradicionales pueden tener dificultades (Ying et al., 2011).

2.1.6. Artificial Neural Networks (ANN): Estructura y entrenamiento de redes neuronales para problemas no lineales

Caparrine (2015) define una Red Neuronal Artificial (RNA) como un modelo matemático inspirado en el comportamiento biológico de las neuronas y en cómo se organizan formando la estructura del cerebro.

Las RNA constituyen modelos computacionales inspirados en el funcionamiento del cerebro humano, conformados por sistemas de procesamiento paralelo conectados mediante estructuras dirigidas para el tratamiento eficiente de información. Cada nodo en la red representa una unidad de procesamiento -o neurona – capaz de recibir y transformar datos complejos (Coello et al., 2015).

En el contexto forestal, el uso de RNA ha demostrado ser una alternativa prometedora en frente a las técnicas de regresión tradicionales. Diversas investigaciones han validado su aplicabilidad para tareas como la estimación del volumen de árboles, la evaluación del crecimiento

y de la producción forestal. Gracias a su capacidad de modelar matemáticamente variables no lineales y fenómenos complejos que escapan a los métodos convencionales y las RNA ofrecen un gran potencial para mejorar la precisión y eficiencia en el manejo de recursos naturales (Nunes et al., 2010).

Las ANN también se han utilizado para estimar la biomasa aérea en bosques tropicales, combinando datos obtenidos por teledetección con mediciones de campo. Un modelo desarrollado en la Amazonia brasileña logró un error de estimación de aproximadamente el 20 % al integrar índices de vegetación como el NDVI, lo que demuestra la eficacia de las ANN en la cuantificación de la biomasa (Costa et al., 2023)

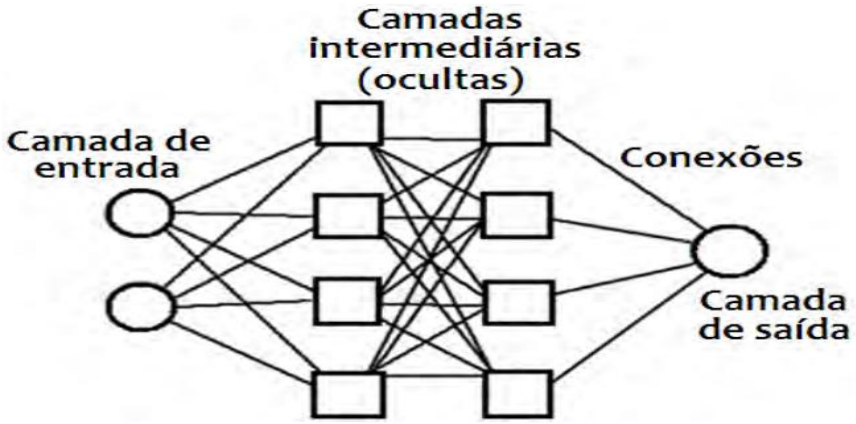
Las redes Neuronales Artificiales (RNA) ofrecen una mayor capacidad para describir con precisión la productividad promedio en extensas areas; que permiten la utilización de un gran número de variables. Esto se debe a su habilidad para incorporar un amplio conjunto de variables, incluyendo factores topográficos, climáticos y fisiográficos, dentro de una misma red, lo que las convierte en herramientas especialmente útiles para realizar pronósticos de crecimiento y producción. Asimismo, en función de la cobertura y representatividad del muestreo, las redes neuronales tienen la capacidad de simular escenarios asociados a la escasez hídrica o eventos de sequía, permitiendo analizar su impacto potencial sobre el crecimiento o desarrollo de las especies forestales valuadas (Flores et al., 2019).

Montaño (2002) resalta que, entre los distintos algoritmos de inteligencia artificial, las Redes Neuronales Artificiales se posicionan como las influyentes en la actualidad, gracias a su notable aplicabilidad en diversos contextos prácticos. En años recientes, esta técnica ha generado un creciente interés entre expertos en estadística y análisis de datos, quienes la han adoptado como

una herramienta metodológica eficaz para la identificación de patrones y la estimación de variables continuas.

Figura 2

Estructura de una red neuronal artificial



2.1.7. Sobreajuste

El sobreajuste es el fenómeno en el que un modelo de aprendizaje automático ajusta demasiado los datos de entrenamiento, lo que impide generalizarlos a datos nuevos e inéditos. Esto resulta en un rendimiento deficiente con los datos de prueba, ya que el modelo aprende ruido en lugar de la distribución subyacente (Gygi et al., 2023).

2.1.8. Evaluación de modelos predictivos

La evaluación del rendimiento de los modelos de aprendizaje automático requiere el uso de métricas estadísticas adecuadas, seleccionadas según el tipo de tarea a analizar. Entre las más utilizadas se encuentran: la exactitud, que representa el porcentaje global de predicciones correctas;

la precisión, que valora específicamente la calidad de las predicciones positivas; la puntuación F1, que combina en equilibrio la precisión y la capacidad de recuperación de modelo (Geng, 2024).

2.3 Marco conceptual

2.3.1. Volumen total del árbol

El volumen es una variable dasométricas que representa en unidades de metros cúbicos (m3), la cantidad de madera contenida en un árbol. Es fundamental para determinar la disponibilidad de recurso maderable, facilitando la planificación eficiente y promoviendo el manejo sostenible de los recursos forestales (Augusto, 2009).

2.3.2. Diámetro a la altura del pecho (DAP)

El diámetro a la altura del pecho (DAP), es una medida estándar utilizada en estudios dasométricas para representar el grosor del tronco del árbol en pie(Li et al., 2023). Ferreyra (2005) define como una medición directa tomada a 1.3 metros sobre el nivel del suelo, siguiendo una convención ampliamente aceptada en el ámbito forestal. El DAP se registra incluyendo la corteza del árbol.

2.3.3. Altura total

Solano (2013) establece que la altura de un árbol corresponde a la medida vertical que va desde la base en el nivel del suelo hasta su punto más alto del ápice. Asimismo, señala que la altura comercial se refiere a la distancia desde el suelo hasta el extremo superior de la última sección del árbol que puede ser aprovechada comercialmente.

2.3.4. Forma del árbol

La forma del tronco de un árbol se caracteriza a través del factor de forma, que influye en la estimación precisa del volumen. Este factor refleja las variaciones por diversos elementos como

las condiciones de la calidad del sitio, el tipo de suelo, el clima, la especie, la edad y otras características (Oluwajuwon et al., 2024).

2.3.5. Variables predictoras

Las variables predictoras, también denominadas variables independientes, constituyen fundamento estructural de los modelos de análisis, ya que proporcionan información relevante sobre las relaciones entre los distintos puntos de datos. Son elementos clave que los investigadores utilizan para interpretar y explicar el comportamiento de una variable dependiente (Espinoza, 2018).

2.3.6. Un modelo IA

Peter (2023) define que un modelo de inteligencia artificial es un marco computacional que simula la inteligencia humana para realizar tareas como el razonamiento, el aprendizaje y la toma de decisiones. Incorpora diversas técnicas de IA para resolver problemas reales y mejorar la automatización en diversas áreas de aplicación.

2.3.7. Errores de predicción

Un error de predicción se define como la discrepancia entre el valor estimado de una variable y su valor observado en la realidad. Esta diferencia puede originarse por múltiples factores, como la calidad deficiente de datos, fallos en el diseño o desempeño de los algoritmos; el error de predicción simple una función critica como el mecanismo de retroalimentación, contribuyendo al ajuste y mejora de los modelos aplicados (Tronson, 2020).

2.3.8. Validación cruzada

La validación cruzada es un método estadístico empleado para medir la eficacia de los modelos predictivos. Esta técnica se basa en la partición del conjunto de datos en varios subconjuntos utilizada para evaluar el rendimiento de los modelos predictivos. También permite

estimar de manera robusta el error de predicción y seleccionar los modelos más adecuados en función de su capacidad de generalización (Wager, 2019)**.**

2.3.9. El sobreajuste

El sobreajuste es un problema en el aprendizaje automático, donde un modelo logra una precisión perfecta en su conjunto de datos de entrenamiento, pero resulta un rendimiento deficiente con datos inéditos. Esto resulta en una mayor complejidad, mayores costos computacionales e inestabilidad en el rendimiento del modelo (Rabin et al., 2024).

2.3.10. El error cuadrático medio (RMSE)

El error cuadrático medio (RMSE) es una métrica estadística utilizada para medir la precisión de los modelos predictivos. Su función consiste en cuantificar la diferencia entre los valores estimados por el modelo y los valores reales observados. Se calcula como la raíz cuadrada del promedio de los cuadrados de los errores; lo que hace esta medida sea particularmente sensible a errores de gran magnitud debido al efeto de elevar cuadrados. La sensibilidad convierte al RMSE en una herramienta útil para detectar desviaciones significativas y mejorar el ajuste de modelos en análisis forestales (Hodson., 2022).

2.3.11. Error Absoluto Medio (MAE)

El Error Absoluto Medio (MAE) es un método para medir la precisión de los modelos de pronóstico. Representa el error promedio entre los valores predichos y los valores reales, otorgando la misma ponderación a todos los puntos de datos para una evaluación intuitiva de la precisión general de la predicción (Suryanto & Muqtadir., 2019)**.**

2.3.12. Coeficiente de determinación (R^2)

El coeficiente de determinación (R^2) mide la bondad de ajuste de los modelos lineales e indica la proporción de varianza en la variable dependiente explicada por las variables independientes y se utiliza ampliamente en la regresión lineal para la selección e interpretación de modelos (Bonghee., 2023).

2.2. Hipótesis

H0: Todos los algoritmos de inteligencia artificial utilizados son estadísticamente similares y significativos para estimar el volumen de árboles de *Pinus patula* de plantaciones forestales de la provincia de Chota.

H1: Por lo menos un algoritmo de inteligencia artificial utilizado es estadísticamente superior para estimar el volumen individual de árboles de *Pinus patula* de plantaciones forestales de la provincia de Chota.

2.3. Operacionalización de variables

Tabla 1
Operacionalización de variables

Variable	Concepto	Dimensión	Indicador	Instrumento de medición	Escala de medición
Volumen de árboles de *P. patula*	Volumen maderable del fuste de los árboles de pino determinado por cubicación rigurosa.	Volumen total del fuste de los árboles.	Volumen	wincha y forcípula	m^3
Características dasométricas.	Dimensiones de las características	Dimensiones dasométricas de los árboles.	Altura Total	Hipsómetro	m
			Altura dominante	Hipsómetro	m

	dasométricas de los árboles (DAP, DAPD, H, HD).	DAP	forcípula o cinta diamétrica	cm
		Densidad de plantación	Conteo de árboles por unidad de área	arb/ha
		Edad	Entrevista	años
		Altitud	GPS o Modelo digital de elevación	m

CAPÍTULO III

MARCO METODOLÓGICO

3.1. Tipo y nivel de investigación

La investigación fue de tipo descriptiva porque se han medido y colectados datos de volumen de árboles para posteriormente testar y determinar el mejor modelo de para estimación del volumen individual de árboles de *Pinus patula* (Hernández & Mendoza, 2018).

Por naturaleza del estudio, esta investigación se enmara dentro de un nivel básico, dado que se centró en el análisis comparativo y en la determinación de correlaciones entre variables específicas, con el propósito de identificar patrones y relaciones significativas dentro del contexto evaluado.

3.2. Diseño de la investigación

Se empleó el diseño no experimental, las variables no se manipularon, solo se tuvo en cuenta la distribución, edad y tamaño de los árboles en las plantaciones forestales, a partir de las cuales se determinó el volumen real y posteriormente se utilizó y seleccionó el mejor algoritmo de inteligencia artificial para estimación de este (Hernandez, 2014).

3.3. Métodos de investigación

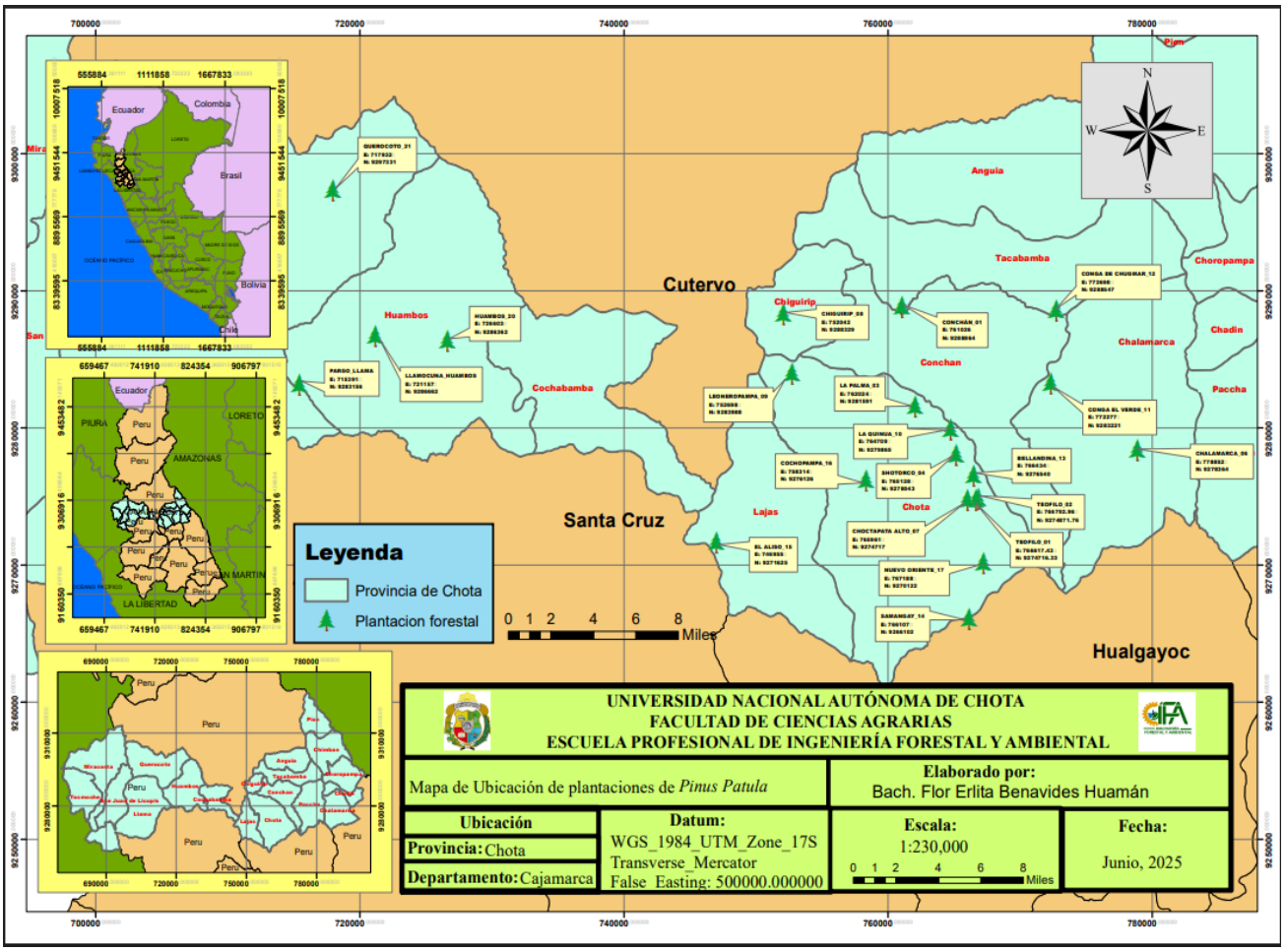
3.3.1. Área de estudio

La presente investigación se desarrolló en la provincia de Chota, situada en la región andina del norte del Perú. Su capital está ubicada en la meseta de Acunta, a una altitud de 2,388 metros sobre el nivel del mar, y se encuentra aproximadamente 150 km al norte de la ciudad de Cajamarca o a 219 km al este de Chiclayo, departamento de Lambayeque. Geográficamente, se localiza entre

las coordenadas 6° 33' 42" Latitud Sur y 78° 38' 56" Longitud Oeste; las plantaciones forestales de *Pinus patula* consideradas en este estudio se distribuyen principalmente en los distritos Chota, Conchan, Chalamarca, Chiguirip, Lajas, Huambos, Tacabamba, Llama y Querocoto dentro de un rango altitudinal que va desde los 2367 hasta 3478 metros sobre el nivel del mar.

Figura 3

Ubicación de área de estudio, en la provincia de Chota, zona central de la región Cajamarca, norte de Perú.



3.3.2. Trabajo de campo

a. Localización y delimitación de parcelas

Las plantaciones forestales fueron localizadas de forma preliminar utilizando la plataforma Google Earth, consultas a especialistas, pobladores locales, y visitas de campo.

En cada plantación evaluada, fueron localizadas al azar una o más parcelas, las cuales fueron delimitadas estableciendo un primer vértice (V1) y luego utilizando una wincha fueron medidos los cuatro lados de 20 m de distancia, formando un cuadrado con la ayuda de una brújula. Fueron establecidos los vértices 2 (V2), 3 (V3) y 4 (V4) en sentido antihorario, iniciando con el azimut 0º N, conformando una parcela cuadrangular de 400m^2 en la que se midieron entre 45 y 55 árboles. Los vértices de cada parcela fueron georreferenciados con GPS (figura 5).

Figura 4

Delimitación de parcelas de inventario forestal, con área de 400 m^2 (20m x 20m), en las plantaciones de Pinus patula

b. Inventario

Fueron medidos el 100% de los árboles de las parcelas, el diámetro a la altura del pecho (cinta diamétrica), altura total (H) utilizando el nivel Abney, además fueron registrados los cuatro árboles dominantes de cada parcela considerando el procedimiento establecido por Hart (1928). Los datos fueron registrados en el formato de inventario (tabla 2).

Tabla 2
Formato para registrar de los árboles inventariados

Lugar:				
N° de parcela:		**Fecha de plantación:**		
Coordenadas	**Este:** **Norte:**			
N° de árbol	**DAP (cm)**	**H (m)**	**HD (m)**	**Observación**
1				
2				
…				
n				

c. Tabulación de datos y distribución diamétrica

Los datos del inventario forestal fueron procesados mediante su digitalización y organizada en una hoja de cálculo de Microsoft Excel, lo que permitió su sistematización para facilitar el análisis estadístico y posterior a ello fue determinada la distribución diamétrica en la que se calculó en el número de clases (F1) y la amplitud de clases (F2).

$$k = 1 + 3.322 \times \log(N) \quad \text{(F1)}$$

$$C = \frac{dap_{max} - dap_{min}}{k} \quad \text{(F2)}$$

Donde: k= número de clases, N= número total de árboles inventariados, C= amplitud de clases, dap_{max} y dap_{min} el diámetro máximo y mínimo respectivamente.

d. Cubicación rigurosa de árboles

Preliminarmente fueron cubicados rigurosamente 122 árboles distribuidos de forma proporcional en todas las clases diamétricas determinadas y aleatoriamente en todas las parcelas o plantaciones inventariadas, en seguida fueron aplicadas la formula del volumen de un cilindro (F3) para el tocón, formula de *Smalian* (F4) para la secciones intermedias y formula del volumen de un cono (F5) para la punta del fuste; finalmente el volumen de cada árbol fue calculado por la sumatoria sucesiva del volumen de cada sección cubicada (F6) , luego a partir de estos datos fue ajustado el modelo volumétrico propuesto por Schumacher and Hall (1933) con la finalidad de calcular el tamaño de muestra (F7) representativo de los árboles que deberían ser cubicados.

$$Vsi = S_0 \times L \quad \text{(F3)}$$

$$Vsi = \frac{S_0 + S_1}{2} \times L \quad \text{(F4)}$$

$$Vsi = \frac{S_0 \times L}{3} \quad \text{(F5)}$$

$$v = Vs_1 + Vs_2 + \cdots \ldots + Vs_{n-1} + Vs_n \quad \text{(F6)}$$

$$n = \frac{t_{\propto}^2 (CV)^2}{E\%^2} \qquad \text{(F7)}$$

Donde: Vsi = volumen de la sección de cada árbol(m^3), $S_0 + S_1$ = área de las secciones transversales menor y mayor (m^2), L = longitud de la sección (m), v= volumen de cada árbol (m^3), Vs_1 hasta Vs_n = volumen de cada sección cubicada del fuste desde la sección 1 en la base hasta la sección "n" en la punta del árbol, n = número de árboles cubicados, $t =$ valor tabular de t de Student, CV =con eficiente de variación del volumen, $E\%$= error de muestreo admitido.

3.3.3. Ajuste y selección de modelos

Los datos de cubicación rigurosa fueron digitalizados, sistematizados y verificados en una planilla de Microsoft Excel, organizados por número de árbol, diámetro a la altura del pecho, altura total, altura dominante y volumen.

Fueron probados un modelo de regresión y tres algoritmos de inteligencia artificial, Random Forest (RF), Support Vector Machine (SVM) y Redes Neuronales Artificiales (RNA). Para el ajuste de modelos, la base de datos fue dividida en datos de entrenamiento o prueba (70%) y datos de validación (30%).

a. Regresión

Se empleó el modelo intrínsecamente linealizado de Schumacher (ecuación 8).

$$V = \beta_0 \times dap^{\beta_1} \times H^{\beta^2} \qquad \text{Ecuación 8}$$

Donde: v= volumen total del árbol, dap= diámetro normal, H= altura total, Ai= parámetros

b. Random Forest

El algoritmo Random Forest desarrollado por Breiman (2001) compuesto por un conjunto de árboles de regresión construidos aleatoriamente y agrupados en particiones con valores similares para la variable de respuesta, fue aplicado sobre los datos entrenamiento utilizando como variable de respuesta el volumen de los árboles (m^3) y como variables explicativas el dap (cm) y altura total (m). Como criterio de parada fueron definidos 500 árboles de regresión y la minimización del error de la variable estimada.

c. Suport Vector Machine

Campos (2016) define este algoritmo de aprendizaje automático supervisado se ha mostrado altamente eficiente en la segmentación de datos lineales y no lineales**.** El procedimiento de construcción de vectores basado en un proceso de aprendizaje automático estadístico ha sido detallado por (Haykin, 2001) **y** (Christmann & Van, 2008)**.** La aplicación de este modelo se basó en determinar la dependencia del volumen de los árboles en relación con las variables independientes dap y altura total. El criterio de parada se basó en la minimización del error de la variable estimada.

d. Redes Neuronales Artificiales

Las RNA son sistemas que procesan información intentando replicar el comportamiento del cerebro humano desarrolladas por autor el científico Frank Rosenblatt, inspirado en el trabajo de Warren McCulloch y Walter Pitts creó el Perceptrón, la unidad desde donde nacería y se potenciarían las redes neuronales artificiales. Una red neuronal tiene la capacidad de aprender a partir de los datos, lo que le permite ser entrenada para

identificar patrones, realizar clasificaciones y efectuar predicciones sobre eventos futuros (Castaneda et al., 2023).

El entrenamiento y validación de los modelos fue realizado en el software R versión 4.4.1, utilizando la función de nls del paquete base de R para el modelo de regresión, el paquete Random Forest (Liaw & Wiener, 2002) para el algoritmo RF, el paquete e1071 (Meyer et al., 2024) para Suport Vector Machine y el paquete Neuralnet (Fritsch et al. 2019) para RNA.

La selección del modelo más adecuado se fundamentó en el análisis integral de diversos indicadores estadísticos que permitieron valorar su capacidad predictiva. Entre estos se consideraron el coeficiente de determinación (R^2), la suma de los cuadrados de los errores, el error porcentual, el error absoluto (MAE), el coeficiente de correlación (r) y la raíz del error cuadrático medio (RMSE), los cuales proporcionaron criterios objetivos para comprar la precisión y eficacia de los modelos aplicados. Además, se complementó el análisis con representaciones gráficas de los residuos y comparaciones entre los valores estimados y los observados, lo que permitió validar la precisión y consistencia del modelo seleccionado (Campos & Leite, 2017).

3.4. Población, muestra y muestreo

3.4.1. Población

La población evaluada de esta investigación comprende un total de 21 plantaciones de *Pinus patula* distribuidas en la provincia de Chota, las cuales presentan distintas edades de desarrollo. Estas areas forestales abarcan una superficie total de 41.53 hectáreas.

3.4.2. Muestra

El tamaño de la muestra fue determinado aplicando la formula F7. Considerado un error de muestreo del 20% (E%), primero para realizar la instalación de parcelas de medición en la que fueron medidos un total de 1005 árboles y luego se determinó el tamaño de la muestra de los árboles que deberían ser cubicados, calculado en 122 árboles, que representa el 12.14 % de la población total de árboles inventariados.

3.4.3. Muestreo

La presente investigación se desarrolló utilizando un diseño de muestreo estratificado, dado que se consideraron especies distribuidas en distintos rangos altitudinales, edades y ubicaciones geográficas. Dentro de cada estrato definido, la selección de las unidades maestrales se realizó de forma aleatoria, garantizando así la representatividad y reduciendo posibles sesgos en la recolección de datos.

3.5. Técnicas e instrumentos de recolección de datos

3.5.1. Técnicas de recolección de datos

La recolección de datos se llevó a cabo mediante una fase inicial de observación en campo, seguida de mediciones directas a través de un inventario forestal realizado en parcelas permanentes. Posteriormente, los datos de volumen fueron obtenidos mediante procedimientos de cubicación precisa de árboles, garantizando así la fiabilidad de la información utilizada en el análisis.

3.5.2. Instrumentos para recolección de datos

Formularios de campo, paja rafia, cinta métrica, spray, jalones y equipos de medición como GPS, Wincha, forcípula y Nivel Abney.

3.6. Técnicas de procesamiento y análisis de datos

Los datos del inventario forestal fueron organizados y procesada mediante su digitalización en hoja de cálculo Microsoft Excel, lo que permitió estructurar los datos de manera ordenada para su análisis estadístico y modelación, luego fue calculado la distribución diamétrica de árboles y finalmente determinado el tamaño de la muestra de los árboles que fueron cubicados rigurosamente.

Todos los datos de cubicación fueron procesados de acuerdo a las fórmulas detalladas previamente y la base de datos fue organizada tal como se detalla en la tabla 3.

Tabla 3
Base de datos para el modelamiento del volumen individual de árboles

N°	Volumen	DAP	HT (m)	HTDOM	Edad
1					
2					
…					
122					

Donde: DAP= Diámetro a la altura del pecho; HT= altura; HTDOM= altura dominante; S= índice de sitio.

Previo a la configuración de algoritmos y modelamiento, fue determinada la normalidad de los datos mediante la prueba de Kolmogórov-Smirnov. Así mismo se determinó el coeficiente de correlación de Pearson entre todas las variables.

La evaluación de cada algoritmo en la estimación del volumen individual de los árboles fue evaluada el conjunto de datos de validación, mediante el cálculo de diversos indicadores estadísticos:

❖ Coeficiente de correlación (r) se emplea como indicador estadístico para determinar el grado de asociación entre los valores observados y los estimados. Esta métrica permite evaluar la fuerza y dirección de la relación lineal entre ambas variables, siendo fundamental para validar la precisión del modelo en términos de ajuste predictivo.

$$r_{xy} = \frac{S_{xy}}{S_x \; S_y} = \frac{\sum(xi - \bar{x})(yi - \bar{y})}{\sqrt{(\sum xi - \bar{x})^2)(\sum(yi - \bar{y})^2}} \qquad \textbf{Ecuación 09}$$

Donde: yi = datos observados del volumen y xi = datos del volumen estimado por el modelo.

❖ Coeficiente de determinación (R^2) fue utilizado como indicador para cuantificar el grado en que cada algoritmo logra explicar la variabilidad observada en los datos. Esta métrica permite evaluar la capacidad explicativa del modelo, expresada en términos porcentuales, y constituye un criterio fundamental para valorar su desempeño en la estimación del volumen arbóreo.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \qquad \textbf{Ecuación 10}$$

Donde: y_i = dato observado, $\hat{y}_i$ = dato estimado, $\bar{y}$ = promedio de los datos observados.

❖ Error medio absoluto (MAE) permitió evaluar de manera general la varianza del error.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$ **Ecuación 11**

Donde: MAE = error medio absoluto, y_i = dato estimado, x_i = dato observado y n = número total de observaciones.

❖ Raíz del error cuadrático medio (RMSE) se emplea como una métrica estándar para evaluar el rendimiento de los modelos predictivos, ya que permite cuantificar la desviación promedio entre los valores observados y los estimados.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2}{n}}$$ **Ecuación 12**

Donde: Y_i = valores estimados, $\hat{Y}_i$ = valores observados y n = número de observaciones.

❖ El error porcentual (E%) se obtiene a partir de la diferencia entre los valores observados y estimados del volumen, y permiten evaluar la precisión relativa del modelo. Esta métrica expresa el grado de desviación en términos porcentuales, facilitando la comparación entre distintos algoritmos o enfoques de estimación.

Ecuación 13

$$E(\%) = \frac{Yi - \hat{Y}i}{Yi} 100$$

Donde: Yi = valor observado en campo de la variable dependiente, $\hat{Y}i$ = valor estimado por el modelo.

Así mismo se realizó la verificación de los supuestos estadísticos asociados al análisis de residuos. La normalidad de los errores fue evaluada mediante la prueba de Kolmogórov-Smirnov, mientras que la homocedasticidad se analizó utilizando la prueba de Breusch – Pagan. Por otro lado, la independencia de los residuos fue determinada a través de la prueba de Durbin-Warson, con el fin de garantizar la validez de los modelos aplicados en la estimación volumétrica.

Complementariamente, se realizaron análisis gráficos que incluyeron la dispersión de los residuos, la distribución de sus frecuencias y la comparación entre los valores estimados y los observados, con el fin de validar la consistencia y precisión del modelo ajustado.

3.7. Aspectos éticos

La presente investigación no implicó la participación directa de seres humanos ni de animales en sus experimentos. La información recabada de los pobladores, obtenida mediante consultas, será utilizada exclusivamente para los fines establecidos en este estudio, respetando los principios éticos de confidencialidad y consentimiento informado. Finalmente, la presente investigación se desarrolló garantizando el uso adecuado de datos.

CAPÍTULO IV.

RESULTADOS Y DISCUSIÓN

4.1. Información general de las plantaciones inventariadas

El inventario forestal se realizó en 21 plantaciones distribuidas en la provincia de Chota y en cada plantación se instaló una parcela de medición, siendo inventariados un total de 1005 árboles cuyas estadísticas descriptivas se muestra a continuación:

Tabla 4
Estadísticas descriptivas de las variables dasométricas del inventario forestal

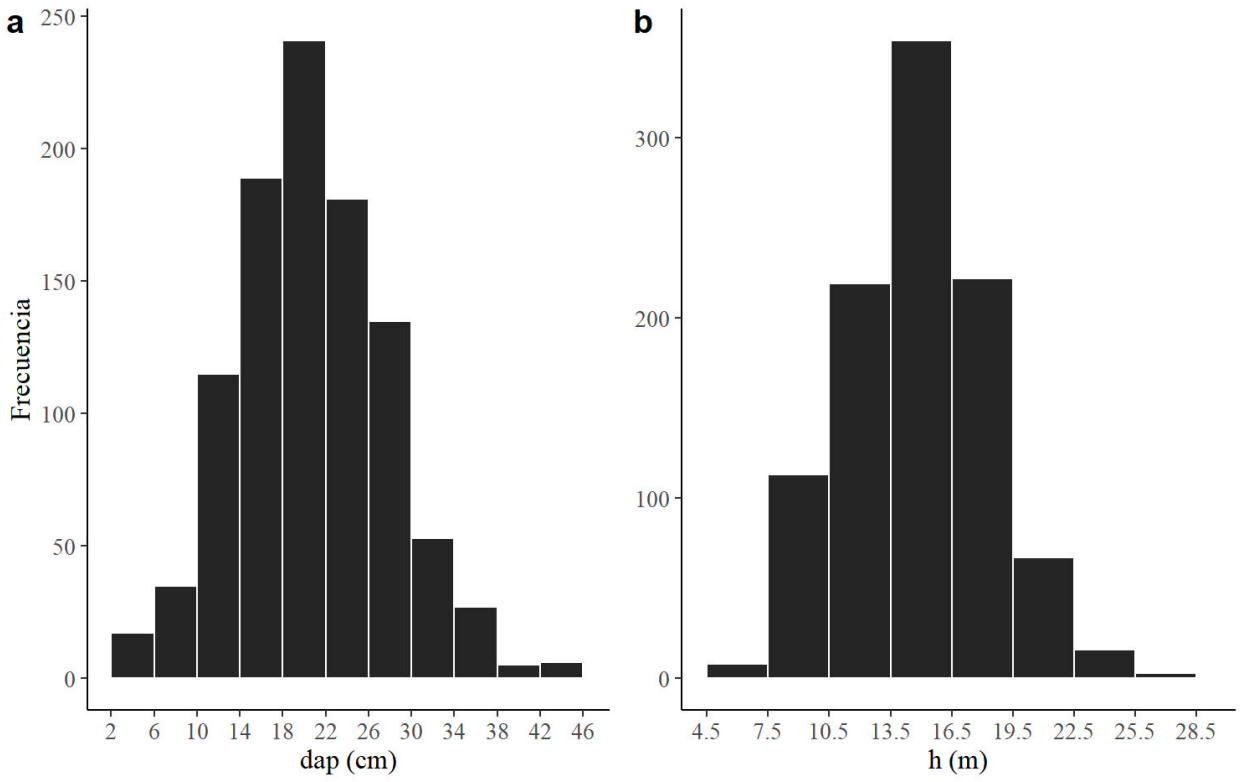
Variable	Mínimo	Promedio	Mediana	Máximo	Desviación estándar	Coefiente de Variacion
dap	3.9	20.8	20.4	46.3	7.02	33.75%
h	4.6	15.0	15.0	37.2	3.57	23.80%
hd	11.8	18.36	19.10	23.98	2.95	16.07%
Altitud	2367	2979	3123	3478	378.56	12.71%

Nota: *dap*: diámetro a la altura del pecho medido a 1,30 m sobre el nivel del suelo, h: altura total, hd: altura dominante.

A partir de la aplicación de la fórmula de Sturges (F1) se determinaron 11 clases diamétricas y una amplitud de clase (F2) de 3.85 cm. Para facilitar nuestro trabajo, fue definida una amplitud de clases de 4 cm, resultando en una distribución diamétrica con 11 clases tal como se muestra en la figura 5a.

Figura 5

Gráfico de frecuencias de los árboles inventariados, a) distribución de frecuencias del dap, b) distribución de frecuencias de h.



En relación a la distribución de los datos, los resultados del test de normalidad de Kolmogoroy-Smirnov *(H0: Los datos provienen de una distribución normal)* aplicado a los datos de dap y h provenientes del inventario forestal muestran que no existe evidencia estadística para rechazar la hipótesis nula (p-valor > 0.005) por lo tanto se asume que los datos provienen de una distribución normal (Tabla 5, figura 6).

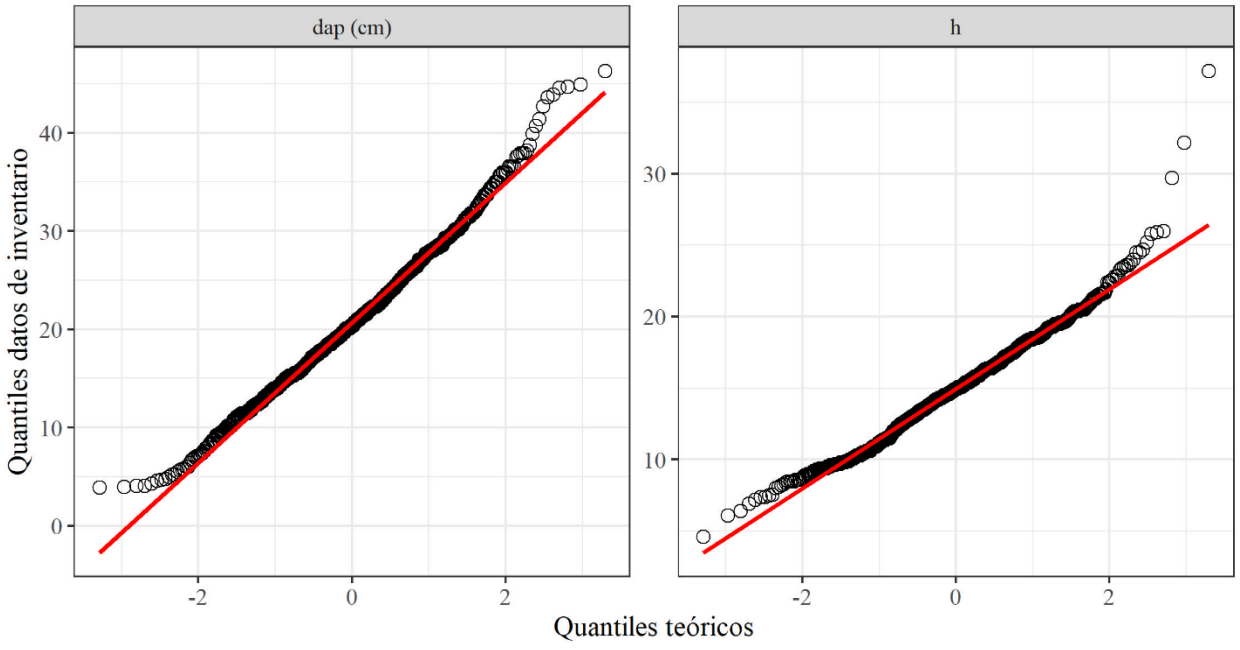
Tabla 5

Resultados del test de normalidad de Kolmogórov-Smirnov en datos de inventario

Variables	Kolmogórov-Smirnov	
	Estadística D	p-valor
DAP	0.038511	0.1015
h	0.028515	0.3873

Figura 6

Normal Q-Q plot para los datos de dap y h provenientes del inventario forestal



.

Tabla 6

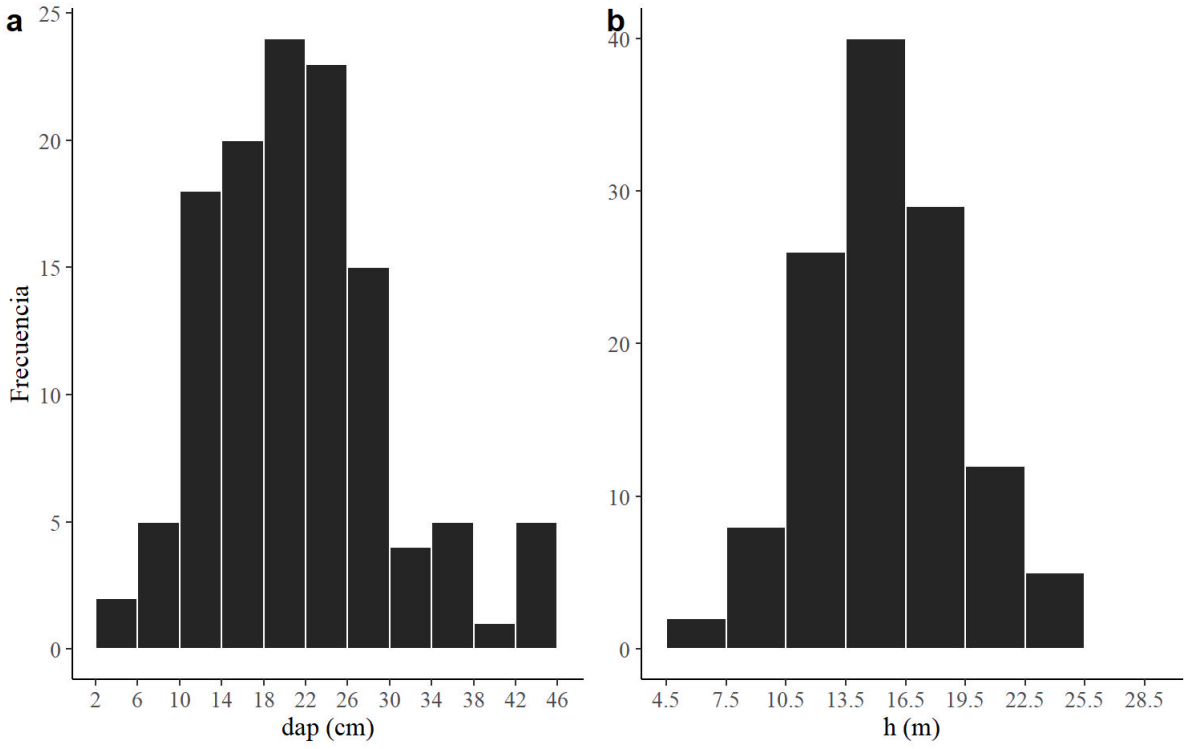
Resumen de las variables dasométricas de los árboles cubicados

Variables	*dap*	*h*	*hd*	*vol*	*altitud*
Min	4.8	5.88	15.97	0.0061	2404
Promedio	21.4	15.43	19.27	0.2757	2972
Mediana	20.4	15.38	19.20	0.1826	3142
Max	44.9	23.6	23.98	1.5827	3303
Desviación estándar	8.70	3.69	2.17	0.2787	358
Coeficiente de variación	40.67%	23.91%	11.26%	97.93%	12.04%

Los valores mínimos y máximos de las variables que se muestran en la tabla 6, esta misma se relaciona con los datos del inventario en la tabla 4.

Figura 7

Gráfico de frecuencias de 122 árboles cubicados rigurosamente a) diámetro de frecuencia del dap, b) distribución de frecuencias de h



Del mismo modo, se realizó la verificación de la normalidad de los datos provenientes de árboles cubicados rigurosamente, los resultados del test de normalidad de Kolmogorov-Smirnov (*H0: Los datos provienen de una distribución normal*) muestran que no existe evidencia estadística para rechazar la hipótesis nula (p-valor > 0.05) por lo tanto se asume que los datos provienen de una distribución normal (tabla 7; figura 8).

Tabla 7
Resultados del test de normalidad de Kolmogorov-Smirnov en datos de árboles muestra

Variables	**Kolmogorov-Smirnov test**	
	Estadística D	p-valor
DAP	0.070441	0.5802
h	0.04162	0.9841

Figura 8

Normal Q-Q plot para los datos de dap y h provenientes del inventario forestal

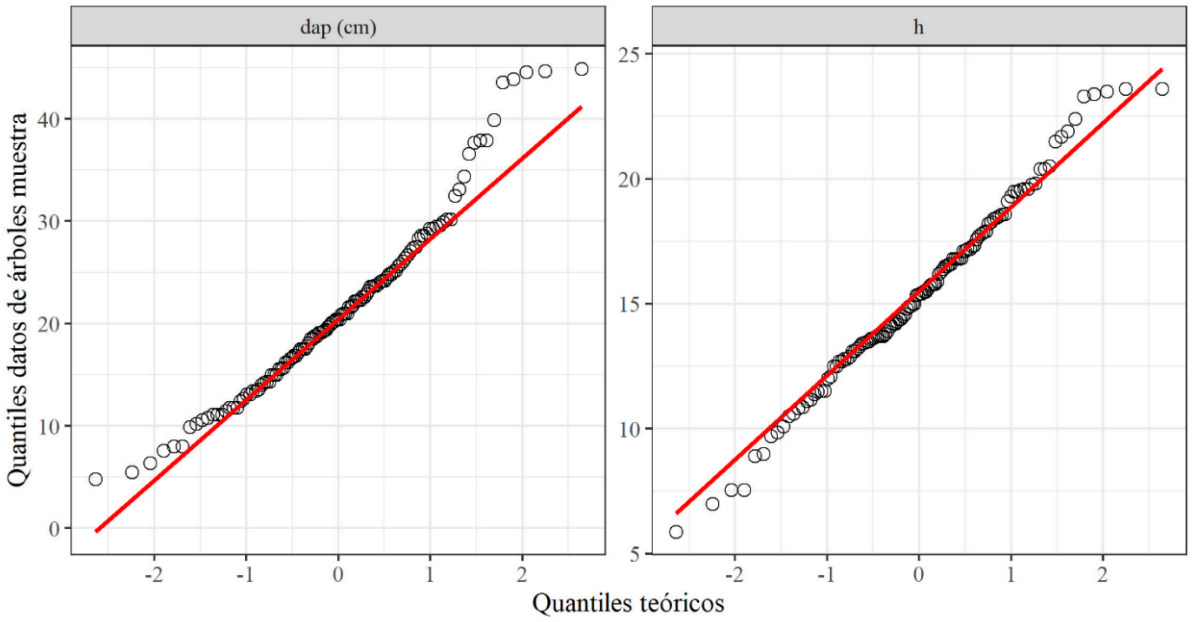
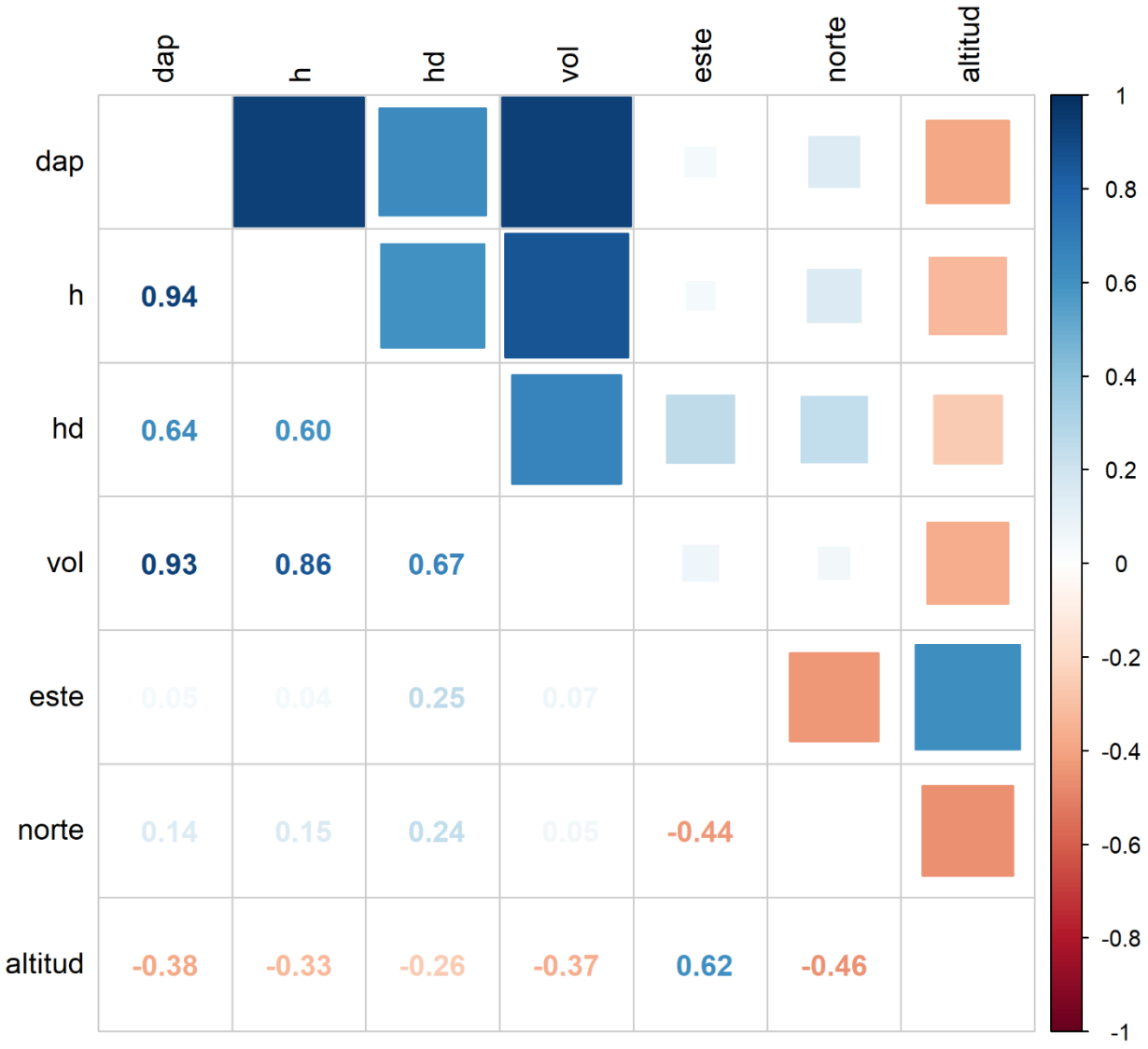


Figura 9

Correlación de Pearson de la variable dependiente (volumen) y variables independientes analizadas.



Nota: Valores en tono azul representan correlación positiva y por lo contrario en tonalidad roja representan correlación negativa.

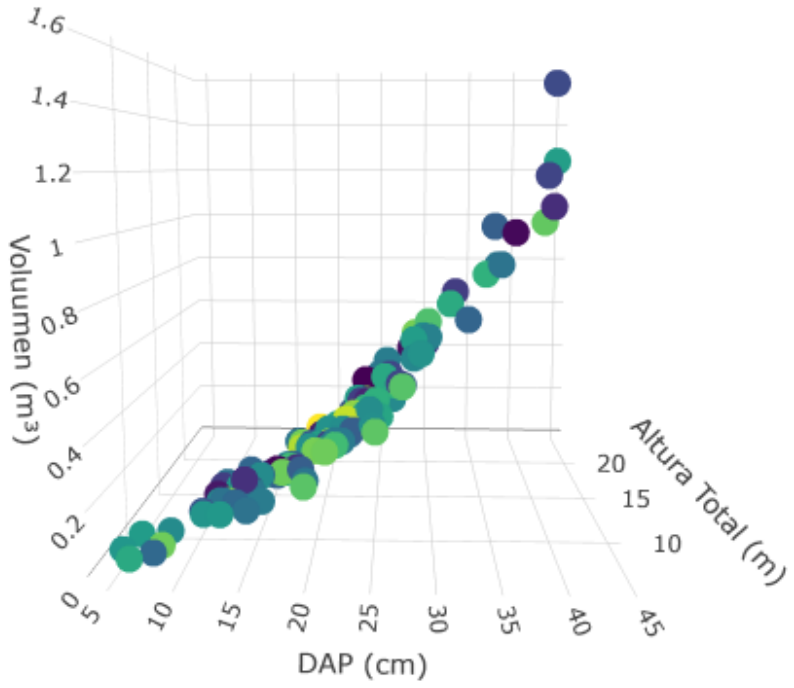
El volumen presenta correlación positiva muy fuerte con el *dap* y h, así como correlación positiva fuerte con hd, evidenciando que, a mayor dimensión de los árboles y mejor calidad de

sitio, se espera mayor volumen de árboles; contrariamente el volumen presenta correlación negativa pero baja con la altitud, por tanto, a medida que las plantaciones de *Pinus patula* se instalen a mayores altitudes el volumen de los árboles tendera a disminuir.

Debido a la baja correlación del volumen y los datos de localización geográfica (este, norte), estas variables fueron descartadas del proceso de entrenamiento del modelo.

Figura 10

Vista tridimensional de la relación volumen, dap y altura total de los árboles cubicados



Nota: Los puntos en el grafico representa los 122 árboles cubicados rigurosamente; eje x (DAP, cm): El diámetro a la altura del pecho, eje Y (altura total, m): La altura total del árbol.

- **Eje Z (Volumen, m³):** El volumen real de los árboles determinado por cubicación rigurosa.

Tabla 8

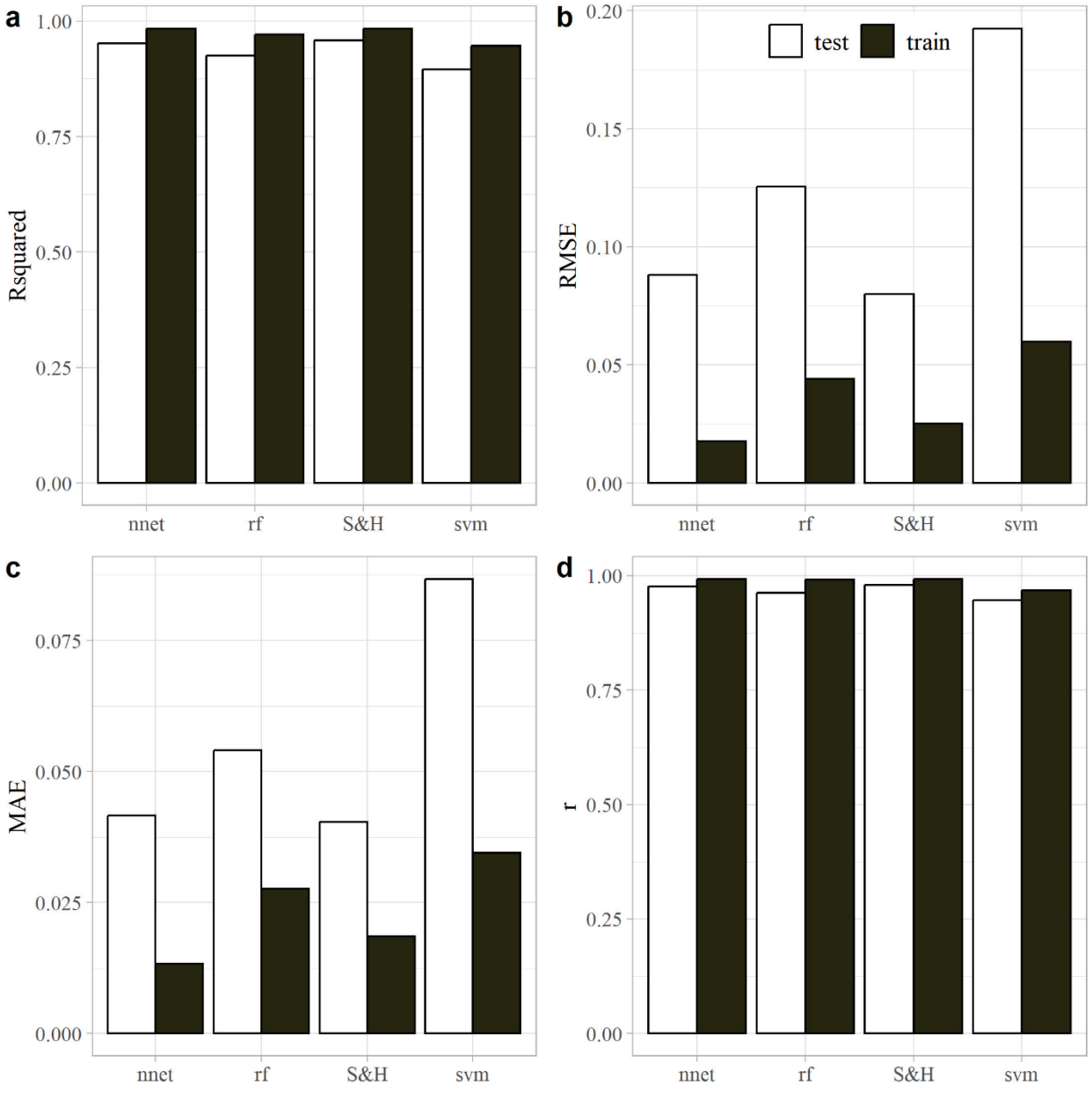
Estadísticos de bondad de ajuste de modelos entrenados

Modelo	**Coeficientes**	R^2	$RMSE$	MAE	$R_{y\hat{y}}$
S&H	$\beta_0 = 0.00004656$ $\beta_1 = 1.8158$ $\beta_2 = 1.0473$	0.9588	0.0799	0.0404	0.9794
RF		0.9253	0.1255	0.0540	0.9619
SVM		0.8956	0.1925	0.0868	0.9464
RNA		0.9523	0.0881	0.0416	0.9759

La tabla 08 muestra las estadísticas de bondad de ajuste de los modelos utilizados, SVM fue el algoritmo de menor desempeño, con un $R^2 = 0.8956$, mayor RMSE y MAE, y a su vez menor correlación entre los valores estimados y valores observados. Por otra parte, de entre los tres algoritmos entrenados y validados, RNA resultó superior, ligeramente por debajo del modelo S&H, cuyo R^2 fue de 0.9588.

Figura 11

Comparación de las estadísticas de bondad de ajuste calculados con datos de entrenamiento (barra negra) y datos de validación (barra blanca) de los cuatro modelos ajustados.



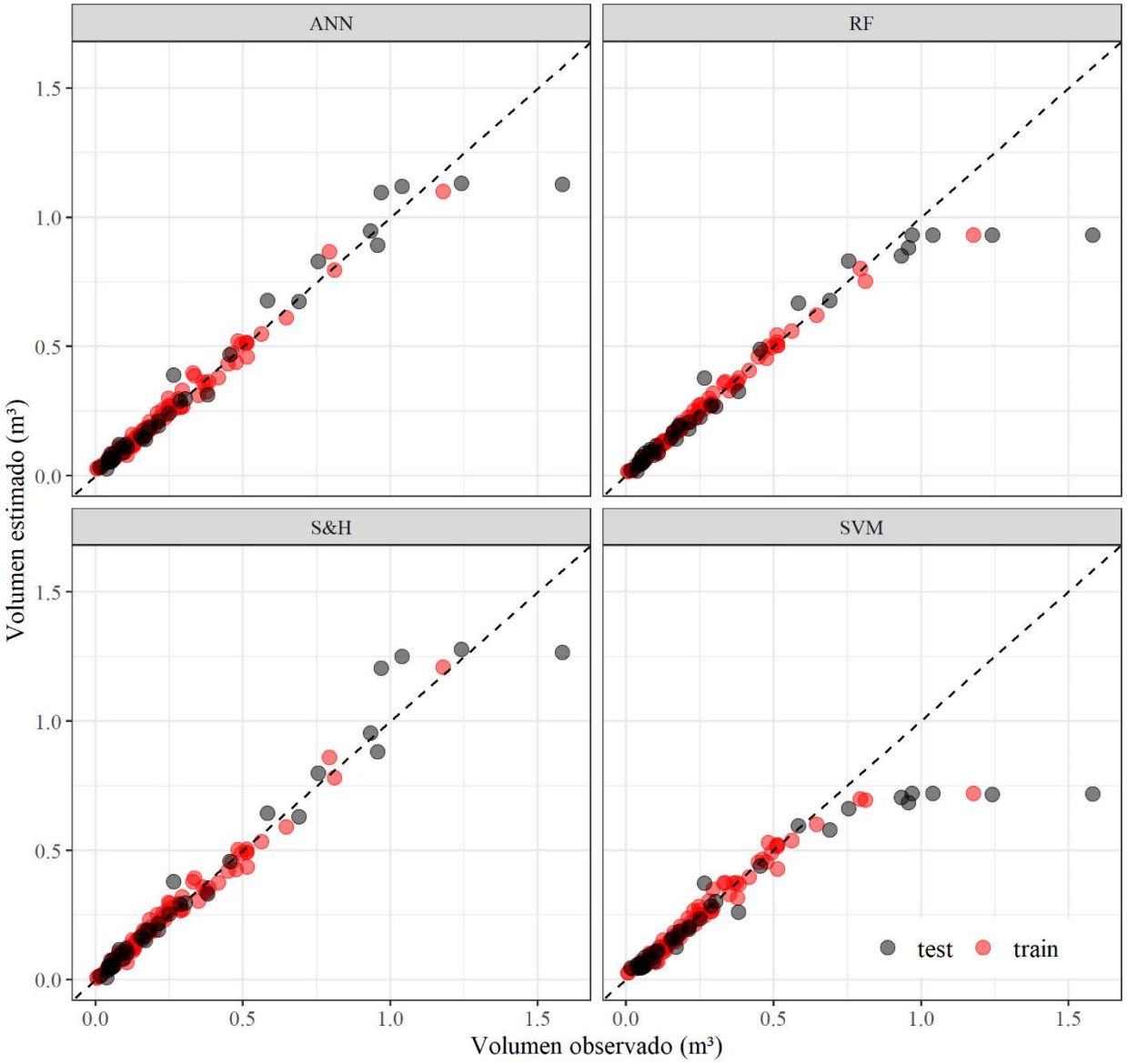
Nota: Redes Neuronales artificiales (nnet), Random Forest (rf), Schumacher & Hall (S&H) y super vector machine (svm).

En cada uno de los gráficos está compuesto por Eje Y (muestra el valor de la estadística) y en el Eje X (muestra el modelo).

Las estadísticas que muestran la eficiencia de cada modelo calculados utilizando los datos de entrenamiento y validación se muestran en la figura 11. La barra blanca corresponde a la estadística determinada con los datos de validación, iguales a los mostrados en la tabla 8; la barra negra corresponde a la estadística calculada con los datos de entrenamiento. La selección del mejor modelo fue determinada considerando las estadísticas de bondad de ajuste calculadas con los datos de validación (test), conforme a lo recomendado por Campos & Leite (2017).

Figura 12

Comparación de las estadísticas de bondad de ajuste calculados con datos de entrenamiento y datos de validación de los cuatro modelos ajustados.



En el eje X se muestra el valor el volumen observado de los árboles y en el eje Y el valor del volumen estimado por los modelos.

En esta figura se puede observar que RF y SVM tienden a subestimar el volumen de los árboles de mayor volumen, especialmente aquellos mayores a 1m3, también se observa el comportamiento altamente similar entre ANN y el modelo de S&H.

Tanto ANN y S&H subestiman el volumen de los árboles mayores a 1.5 m^3, lo cual se debe principalmente a que existen pocos datos de árboles con volumen superior a 1.3 m^3.

4.2. Contrastación de hipótesis

Al comparar las métricas de bondad de ajuste de los algoritmos utilizados, todos los modelos de inteligencia artificial son estadísticamente diferentes, por tanto, a pesar que todos los algoritmos se mostraron eficientes ($R^2 \geq 0.89$) el algoritmo de Redes Neuronales Artificiales se mostró estadísticamente superior que Random Forest y Support Vector Machine, por lo tanto, existen evidencias estadísticas para rechazar H0.

4.3. Discusión de resultados

Información general

Todas plantaciones de la provincia de Chota que formaron parte de este estudio fueron georreferenciadas e inventariadas. En estas, se evidenció homogeneidad en las parcelas, aptas para estudios posteriores; se instalaron 21 parcelas permanentes de monitoreo con área de 400 m^2 consideradas eficientes y de costo reducido (Campos & Leite, 2017). Del mismo modo, los datos de volumen obtenidos del muestreo de árboles cubicados rigurosamente, necesarios para aplicar a los algoritmos de inteligencia artificial, permitieron obtener resultados satisfactorios.

Las plantaciones de *Pinus patula* de la provincia de Chota se encuentran en un amplio rango de distribución diamétrica, verificado a través del coeficiente de variación de 33,75% el cual es considerado alto nivel de variación dentro de las masas forestales y a su vez es un indicativo de

manejo silvicultural de las plantaciones, principalmente falta de podas y raleos, cuyas prácticas tienen efectos inmediatos y a corto plazo en el crecimiento y homogeneidad de bosques y plantaciones (Bhandari et al., 2024).

El principal objetivo de determinar la distribución del dap en la planificación de manejo forestal es la caracterización de las masas forestales a nivel de árbol (Rijal & Sharma, 2024). La información a nivel de árbol es necesaria, por ejemplo, en las simulaciones de crecimiento cuando se aplican modelos a nivel de árbol (Liu et al., 2025), directamente relacionado al objetivo del estudio. La cubicación rigurosa es una técnica fundamental para determinar el volumen de la madera en los bosques, principalmente basada en ajuste de modelos de redes neuronales, así mismo, es utilizada en estudios dasométricas, que permite determinar el volumen de secciones individuales de los árboles, lo cual finalmente se obtiene el volumen total (Guzmán et al., 2020).

A pesar del amplio rango de variación observado, especialmente en el dap, los promedios de dap y h obtenidos, muestran que el crecimiento de las plantaciones en las provincias de Chota es bueno. Los resultados estadísticos confirman que, en el área de estudio, *Pinus patula* tiene un comportamiento normal en la distribución de dichas variables.

(Castillo et al., 2018) detalla que el dap y v son altamente variables, lo cual resulta fundamental para la toma de decisiones en el manejo de plantaciones forestales. Estos parámetros son clave en el desarrollo de modelos de crecimiento, la caracterización estructural del rodal, la estimación del rendimiento maderable y la determinación de la altura de los árboles, con el propósito de evaluar la calidad de sitio. La variabilidad en el diámetro a la altura del Pecho (DAP) y el volumen de los árboles en las plantaciones forestales del Perú se debe a una combinación de factores ambientales, edáficos, genéticos y de manejo forestal (Guariguata et al., 2017).

Las 21 parcelas inventariadas en este estudio permitieron obtener datos de diámetro y altura de 1005 árboles, cuya validación fue realizada mediante el muestreo destructivo de más del 10% de ese total, asegurando que el error de medición, principalmente de la altura total, ha sido mínimo, indicador de que el equipo de medición ha sido utilizado correctamente.

Carreño et al. (2023) analizaron que la relación entre DAP y la altura en 21 especies. Se identifico que la variabilidad en el DAP puede ser más sensible a factores externos, como el manejo forestal, en comparación la altura que tiene a ser una variable más conservadora en su crecimiento.

Los sitios de plantaciones tienen características similares y que a su vez serian favorables para el crecimiento de pino lo cual confirma que la alta variabilidad del dap y vol se deben principalmente al manejo y probablemente a la procedencia genética, el nivel altitudinal también tiene poca variación, lo cual indica que todo el rango altitudinal en el que están distribuidas las plantaciones y están adecuadamente representadas en la muestra colectada.

La especie *Pinus patula*, se ha convertido en uno de los principales factores de controversia en la sierra peruana y la provincia de Chota no es ajena a esa realidad; por presentar una resistencia ambiental a su expansión a nivel de plantaciones maderables. Sin embargo, los resultados aquí mostrados permiten afirmar que esta especie tiene potencial competitivo, pudiendo convertirse en una opción de desarrollo económico a nivel de pequeño productor o empresas forestales, para lo cual es crucial que toda la cadena productiva sea constantemente monitoreada y guiada por profesionales capacitados, con especial énfasis en la aplicación de podas y raleos.

Según Hernández et al. (2018) el coeficiente de correlación de Pearson ocurre cuando se da para estudiar la relación (o correlación) entre dos variables; por ejemplo, la relación entre la variable dependiente que es el volumen y la altura como independiente. Este coeficiente

proporciona un resumen numérico que va de -1 a +1, donde cada punto final representa una relación lineal perfecta, ya sea negativa o positiva.

La variación de las condiciones edáficas y climáticas a lo largo de los gradientes altitudinales influye directamente en las especies forestales, afectando sus funciones fisiológicas, disminuyendo su crecimiento y altura conforme aumenta la altitud. Las especies forestales tropicales pueden desarrollarse dentro de rangos altitudinales y edafoclimáticos amplios o restringidos. Como resultado, en zonas más elevadas predominan árboles de menor tamaño, con múltiples fustes, usualmente retorcidos, y una elevada proporción entre el diámetro del tronco y a la altura de la copa. (Murga et al., 2021).

En la tabla 8 se identifica que el modelo de S&H resulta ser el más efectivo en este caso, ya que tiene el mejor equilibrio entre las métricas (alto R^2, bajo error y alta correlación), La RNA es la segunda mejor opción, mostrando resultados muy similares a la regresión, el modelo RF tiene un desempeño razonable, pero algo inferior y el modelo SVM tiene el peor desempeño, con los mayores errores y la menor correlación; el entrenamiento del clasificador RF con solo 50 árboles pudo haber limitado su capacidad de capturar patrones complejos, el rendimiento de RF en los conjuntos de muestreo del 60 %, 80 % y 100 % no mostró diferencias significativas lo cual el modelo alcanzó su capacidad máxima de aprendizaje con el 60 % de los datos. El desempeño del modelo SVM fue superior al de RF, especialmente en conjuntos de datos equilibrados y el rendimiento de SVM se estabilizó al utilizar el 60 % del tamaño total de las muestras, lo que implica que este algoritmo puede ser más eficiente en el uso de datos y menos dependiente de tamaños grandes de muestra (Garzón et al., 2023).

La decisión de ajustar y aplicar el modelo de volumen de Schumacher-Hall a los modelos de razón de volumen fue porque este modelo se ha ajustado exitosamente. En la figura 11 a muestra

la comparación de R^2, abarcando el cuanto de la variabilidad de los datos son explicados por el modelo de Schumacher-Hall; los valores de la Tabla 8 refiere que la barra en blanco son datos de validación; se ha hecho un ajuste del modelo, se ha divido los datos en un 70 % ajuste o entrenamiento y en un 30 % de datos de validación. Con ambos conjuntos de datos se ha calculado cada una de las estadísticas y se observa que para el R^2 la barra negra es alta y más baja para las otras variables.

Finalmente decimos que los valores de las estadísticas tanto de estimados y observados se dice que R^2 es el porcentaje que el modelo explica la variabilidad de los datos y todos están sobre el 90 %. También tenemos: R^2: en este caso se analizó que mientras menor sea el valor, el modelo es mejor y MAE: se dice que más bajo el error el modelo es mejores, concluimos que la red neuronal como el modelo de Schumacher-Hall son más buenos ya que el modelo RF.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

- El volumen de los árboles se determina con valores que van desde 0.0061 m^3 hasta 1.5827 m^3. El promedio (0.2757 m^3) y la mediana (0.1826 m^3) son significativamente menores que el valor máximo, lo que indica la presencia de algunos árboles con un volumen mucho mayor que la mayoría. La desviación estándar es relativamente alta (0.2787 m^3), lo que refleja la variabilidad en los tamaños de los árboles.
- Nuestros resultados determinaron que las variables más importantes para estimar el volumen son altura, diámetro, altura dominante y altitud.
- La elección del modelo más adecuado se fundamentó principalmente en el análisis de las estadísticas de bondad de ajuste obtenidas a partir de los datos de validación. Para la estimación del volumen individual de árboles de *Pinus patula* en plantaciones forestales de la provincia de Chota, se emplearon diversos enfoques, entre ellos los algoritmos Random Forest (RF), Support Vector Machine (SVM) y Artificial Neural Network (ANN*)*, así como el modelo tradicional de Schumacher & Hall (S&H), con el propósito de comparar su desempeño predictivo.
- El algoritmo ANN mostró los mejores resultados, con R^2 = 0.9523 y RSME= 0.0881, sin embargo, no superó al modelo S&H, que obtuvo un R^2= 0.9588 y RSME= 0.0799 pero a la vez, no existió diferencias estadísticas entre estos dos, sugiriendo que, tanto el algoritmo ANN y el modelo S&H pueden ser utilizados para estimar el volumen individual de *Pinus patula* procedentes de plantaciones forestales de la provincia de Chota es el modelo S&H,

presentó un R^2 de 0.9588 y valores similares entre ellos para las demás estadísticas de ajuste calculadas y analizadas.

5.2 Recomendaciones

- Es fundamental que la Escuela Profesional de Ingeniería Forestal y Ambiental difunda información sobre la relevancia de conocer el volumen de los árboles en la provincia de Chota. Permite una gestión sostenible de los recursos forestales, optimiza la planificación de aprovechamiento y contribuye a la valoración económica de las plantaciones.
- La Escuela Profesional de Ingeniería Forestal y Ambiental debe garantizar una correcta utilización de los instrumentos de medición forestal para minimizar errores en la estimación de las principales variables dasométricas. Un manejo inadecuado puede comprometer la veracidad de los datos, generando distorsiones en la evaluación del crecimiento y sobreestimaciones en el valor comercial real de las plantaciones.
- Ampliar el número de parcelas de medición y muestras de árboles procedentes de plantaciones forestales a fin de reducir el error de muestreo, que este estudio fue considerado el 20%.
- Probar otros algoritmos de inteligencia artificial y seguir mostrando estas técnicas de regresión entre la comunidad universitaria y en general.

CAPÍTULO VI.

REFERENCIAS

Aarón, D., Cerda, A., Cueto, J., & Muñoz, J. (2012). *Determinación de carbón total y carbón orgánico en diferentes suelos usando método de combustión seca. Vol. 12,* (pp.)143-151. https://dialnet.unirioja.es/servlet/articulo?codigo=5519440

Amat, J. (2020). *Distributional regression rorest: Random forest probabilístico*. https://www.cienciadedatos.net/documentos/54_distributional_regresion_forest.html

Bhandari, S., Pinno, B., Stadt, K., & Thomas, B. (2024). La respuesta del crecimiento al aclareo precomercial del pino lodgepole es a corto plazo, pero los efectos sobre la distribución del tamaño persisten durante décadas. *Forestry Chronicle*, *100*(1), 50–58. https://doi.org/10.5558/tfc2024-002

Bonghee, L. (2023). Model Selection with Coefficient of Determination in Linear Mixed Effects Model (2023) | Bonghee Lee. - *Journal of the Korean Data Analysis Society*, *25*, 91–98. https://scispace.com/papers/model-selection-with-coefficient-of-determination-in-linear-1u00zp45

Breiman, L. (2001). Random Forests. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-62008-0_35

Campos, E. (2016). *Introducción a las máquinas de vector soporte (SVM) en aprendizaje*

supervisado. https://zaguan.unizar.es/record/59156#

Campos, J., & Leite, H. (2017). *Mensuração Florestal: perguntas e respostas*. https://www.passeidireto.com/arquivo/113750014/campos-leite-mensuracao-florestal-perguntas-e-respostas-2017-pdf

Caparrine, F. (2015). *Redes neuronales: una visión superficial*. http://www.cs.us.es/~fsancho/?e=72

Carreño, A., Mora, D., & Serpa, A. (2023). Correlación de diámetro-altura de árboles en la Universidad Francisco de Paula Santander,Cúcuta. *Revista Perspectivas*, *8*(S1), 259–268. https://doi.org/10.22463/25909215.4138

Castaneda, W., Polo, B., & Vega, F. (2023). Artificial neural networks: a measurement of forecast learnings as potential demand. *Universidad Ciencia y Tecnología*, *27*(118), 51–60. https://doi.org/10.47460/uct.v27i118.686

Castillo, E., Jarillo, J., & Escobar, R. (2018). Relación altura-diámetro en tres especies cultivadas en una plantación forestal comercial en el Este tropical de México. *Revista Chapingo, Serie Ciencias Forestales y Del Ambiente*, *24*(1), 33–48. https://doi.org/10.5154/r.rchscfa.2017.05.033

Christmann, A., & Van, A. (2008). Bouligand derivatives and robustness of support vector machines for regression. *Journal of Machine Learning Research*, *9*, 915–936.

Coello, B., Casas, L., Pérez, O., & Caballero, Y. (2015). Redes neuronales artificiales en la

producción de tecnología educativa para la enseñanza de la diagonalización. *Academia y Virtualidad*, *8*(1), 12. https://doi.org/10.18359/ravi.442

Correia, G., Ribeiro, A., Fernandes, G., Sára, S., Marques, M., & Rosa, A. (2018). Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of the Total Environment*, *619–620*, 1473–1481. https://doi.org/10.1016/j.scitotenv.2017.11.138

Costa, A., Rodríguez, J., Pereira, M., De Oliveira, G., Marimon, H., & Trondoli, E. (2023). Artificial intelligence tools and vegetation indices combined to estimate aboveground biomass in tropical forests. *Https://Doi.Org/10.1117/1.JRS.17.024512*, *17*(2), 024512. https://doi.org/10.1117/1.JRS.17.024512

Cruz, C., De los Santos, H., & Valdez, L. (2008). *Sistema compatible de ahusamiento-volumen para Pinuscooperi Blanco en Durango, México*. 473–485.

Dávila, A. (2018). Modelagem de povoamentos de eucalipto no Nordeste da Bahía utilizando regressao e redes neurais artificiais. *Advanced Optical Materials*, *10*(1), 1–9. https://doi.org/10.1103/PhysRevB.101.089902%0Ahttp://dx.doi.org/10.1016/j.nantod.2015.04.009%0Ahttp://dx.doi.org/10.1038/s41467-018-05514-9%0Ahttp://dx.doi.org/10.1038/s41467-019-13856-1%0Ahttp://dx.doi.org/10.1038/s41467-020-14365-2%0Ahttp://dx.doi.org/1

Diamantopoulou, M. (2022). Simulation of over-bark tree bole diameters, through the RFr (Random Forest Regression) algorithm. *Folia Oecologica*, *49*(2), 93–101. https://doi.org/10.2478/foecol-2022-0010

Espinoza, E. (2018). Las variables y su operacionalización en la investigación educativa. Parte I. *Conrado*, *14*, 39–49. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442018000500039&lng=es&nrm=iso&tlng=es

Flores, E., Aguirre, O., Quiñonez, G., González, M., & Jiménez, J. (2019). *Estimación del diámetro normal, altura y volumen de Pinus pseudostrobus Lindl. en función del diámetro del tocón*. *vol 10 (55*, (pp.) 1-58. https://www.locus.ufv.br/bitstream/123456789/6780/1/texto completo.pdf

Galán, R., Santos, H., & Valdez, J. (2008). Crecimiento y rendimiento de Cedrela odorata L. y Tabebuia donnell-smithii Rose en San José Chacalapa, Pochutla, Oaxaca. *Scielo*. http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-04712008000200006

García, W., Pérez, G., Quiñonez, G., De los Santos, H., & Rodríguez, G. (2020). Avances y perspectivas en la modelación aplicada a la planeación forestal en México. *Madera y Bosques*, *26*(2), 1–16. https://doi.org/10.21829/myb.2020.2622004

Garzón, J., Sánchez, N., & Londoño, D. (2023). Evaluación comparativa de los algoritmos de aprendizaje automático Support Vector Machine y Random Forest. *Ciencia e Ingeniería Neogranadina*, *33*(2), 131–148. https://doi.org/10.18359/rcin.6996

Geng, S. (2024). Analysis of the Different Statistical Metrics in Machine Learning. *Geng, S*, *88*, 350–356. https://doi.org/10.54097/jhq3tv19

Goycochea, G., Elera, D., Baselly, J., Pereira, L., & Garcia, H. (2022). Configuration of the deep

neural network hyperparameters for the hypsometric modeling of the Guazuma crinita Mart. in the Peruvian Amazon. *Forests*, *13*(5). https://doi.org/10.3390/f13050697

Guariguata, M., Arce, J., Ammour, T., & Capella, J. (2017). *Las plantaciones forestales en Perú:reflexxiones, estatus social y perspectivas a futuro* (p. 4). https://www.cifor.org/publications/pdf_files/OccPapers/OP-169.pdf

Guirado, E., Martínez-Valderrama, J., & Maestre, F. T. (2021). Potencial de la inteligencia artificial para avanzar en el estudio de la desertificación. *Ecosistemas*, *30*(3), 1–10. https://doi.org/https://doi.org/10.7818/ECOS.2250

Guzmán, J., Aguirre, O., & Vargas, B. (2020). Forest volume estimation techniques with special emphasis on the tropics. *Revista Chapingo, Serie Ciencias Forestales y Del Ambiente*, *26*(2), 291–306. https://doi.org/10.5154/R.RCHSCFA.2019.07.061

Gygi, J., Kleinstein, S., & Guan, L. (2023). Predictive overfitting in immunological applications: Pitfalls and solutions. *Human Vaccines & Immunotherapeutics*, *19*(2), 2251830–2251830. https://doi.org/10.1080/21645515.2023.2251830

Haykin, S. (2001). Kalman filtering and neural networks. *Adaptive Filters*, 501–514. https://doi.org/10.1002/9780470374122.ch46

Hernández, J., Espinoza, F., Rodríguez, J., Chacón, J., Toloza, C., Arenas, M., Cariillo, S., & Bermúdez, V. (2018). Sobre el uso adecuado del coeficiente de correlación de Pearson. *Archivos Venezolanos de Farmacología y Terapéutica*, *37*(5), 586–601.

http://saber.ucv.ve/ojs/index.php/rev_aavft/article/view/16165

Hernández, R., & Mendoza, C. (2018). Metodología de la investigación las rutas cuantitativa,cualitativa y mixta. In *Mc Graw Hill* (Vol. 66). https://doi.org/- ISBN 978-92-75-32913-9

Hernandez, S. (2014). Metodología de la investigación. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

Hildt, E., Leszczuk, A., & Somma, F. (2020). *Evaluación de un modelo de Deep Learning para la estimación de densidad en bosques implantados*. *October*. https://doi.org/10.13140/RG.2.2.32307.04647

Hodson, T. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development*, *15*(14), 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022

Hynynem, J. (2011). Conceptos básicos para la modelación del crecimiento forestal. *Recursos Naturales y Ambiente*, *64*, 1–6.

Jevšenak, J., & Skudnik, M. (2021). A random forest model for basal area increment predictions from national forest inventory data. *Forest Ecology and Management*, *479*(July 2020), 118601. https://doi.org/10.1016/j.foreco.2020.118601

Li, S., Fang, L., Sun, Y., Xia, L., & Lou, X. (2023). Development of Double-Angle Laser Measuring Device for Stand Diameter at Breast Height. *Chinese Journal of Sensors and*

Actuators, *36*(6), 990–998. https://doi.org/10.3969/j.issn.1004-1699.2023.06.021

Liu, X., Li, F., Hao, Y., Zou, X., Cheng, S., & Dong, L. (2025). Modeling tree radial growth with competition and crown characteristics from an aerial perspective. *Forest Ecology and Management*, *576*, 122387. https://doi.org/10.1016/J.FORECO.2024.122387

Llerena, C., Hermoza, R., & Llerena, L. (2016). *Plantaciones forestales , agua y gestión de cuencas*. pp.1-36. https://www.researchgate.net/publication/291979716%0APlantaciones

Marques, M., Breda, D., Garcia, H., Ribeiro, S., Ferreira, M., Rode, R., & Lopes, A. (2014). Redes neuronales artificiales para la estimación del volumen de los árboles. *Revista Arvore*, *38*(2), 283–288. https://doi.org/10.1590/S0100-67622014000200008

Martins, J., Marques, M., Ferreira, E., Fernandes, G., Ribeiro, A., Dias, C., Franca, E., Silva, J., Correia, G., Naruna, M., & Rodrigues, M. (2019). Técnicas computacionales aplicadas a la estimación del volumen y la biomasa de los árboles en la sabana brasileña. *Journal of Environmental Management*, *249*(August), 109368. https://doi.org/10.1016/j.jenvman.2019.109368

Martins, Marques, M., Ribeiro, A., Fernandes, G., Almeida, A., Ferreira, E., Otone, M., Silva, J., & Mafra, N. (2021). Prognosis of forest production using machine learning techniques. *Information Processing in Agriculture*, *xxxx*. https://doi.org/10.1016/j.inpa.2021.09.004

Mirez, Y. (2021). *Estimación de biomasa y contenido de carbono en Pinus patula de diferentes edades en la jalca del distrito de Chalamarca , Chota – Cajamarca*. 77.

Monnet, J., Chanussot, J., & Berger, F. (2011). Support vector regression for the estimation of forest stand parameters using airborne laser scanning. *IEEE Geoscience and Remote Sensing Letters*, *8*(3), 580–584. https://doi.org/10.1109/LGRS.2010.2094179

Montaño, J. J. (2002). Redes neuronales artificiales aplicadas al análisis de datos [Universidad de las Islas Baleras]. In *Network*. https://www.tesisenred.net/bitstream/handle/10803/9441/tjjmm1de1.pdf?sequence=1

Murga, H., Coronado, M., Abanto, C., & De Almeida, F. (2021). Altitudinal gradient and its influence on the edofoclimatic characteristics of tropical forests. *Madera y Bosques*, *27*(3), 1–13. https://doi.org/10.21829/myb.2021.2732271

Nunes, I., Hernane, D., & Andrade, R. (2010). Redes neurais artificiais para engenharia e ciências aplicadas. In *São Paulo: Artliber*. https://artliber.com.br/amostra/redes_neurais.pdf

Oluwajuwon, T., Ogbuka, C., Ogana, F., Hossain, M., Israel, R., & Lee, D. (2024). Describing and Modelling Stem Form of Tropical Tree Species with Form Factor: A Comprehensive Review. *Forests*, *16*(1), 29–29. https://doi.org/10.3390/F16010029

Organización de las Naciones Unidas para la Agricultura y la Alimentación. (2010). *Casos ejemplares de manejo forestal comunitario en América Latina y el Caribe*.

Orrillo, E. (2019). *Comparación de tres métodos de estimación de volumen con la medición de trozas de Pinus patula Schl. et Cham en el caserío La Apalina - La Encañada – Cajamarca* [Universidad Nacional de Cajamarca].

https://repositorio.unc.edu.pe/handle/20.500.14074/2843

Ortega, S., & Salgado, J. (2020). *Ficha bibliografíca documento de opción de cargo*.

Özçelik, R., Diamantopoulou, M. J., Brooks, J. R., & Wiant, H. V. (2010). Estimating tree bole volume using artificial neural network models for four species in Turkey. *Journal of Environmental Management*, *91*(3), 742–753. https://doi.org/10.1016/j.jenvman.2009.10.002

Peter, H. (2023). An adjusted coefficient of determination (R2) for generalized linear mixed models in one go. *Biometrical Journal*, *65*(7), 1–17. https://doi.org/10.1002/bimj.202200290

Pinno, B. D., Lieffers, V. J., & Landhäusser, S. M. (2012). Inconsistent growth response to fertilization and thinning of Lodgepole Pine in the Rocky Mountain Foothills Is Linked to Site Index. *International Journal of Forestry Research*, *2012*. https://doi.org/10.1155/2012/193975

Reaño, J. (2020). *Aplicación de inteligencia artificial en diagnósticos de enfermedades* (Vol. 2507, Issue February). Universitas Studiorum Sanctus Turibius de Mogrovejo. Chiclayo. Perú.

Revilla, J., Abanto, C., Guerra, W., García, D., Guerra, H., Domínguez, G., & Da Silva, I. (2021). Modelos alométricos para estimar el volumen de madera de Guazuma crinita en plantaciones forestales. *Scientia Agropecuaria*, *12*(1), 25–31. http://www.scielo.org.pe/pdf/agro/v12n1/2306-6741-agro-12-01-25.pdf

Rijal, B., & Sharma, M. (2024). Modelling Diameter at Breast Height Distribution for Eight Commercial Species in Natural-Origin Mixed Forests of Ontario, Canada. *Forests*, *15*(6).

https://doi.org/10.3390/f15060977

Rodríguez, F., Pérez, F., Águeda, B., Alonso, R., García, A., Olano, J., & García, M. (2021). La inteligencia artificial ya es una realidad en el sector forestal. *Foresta*, *81*, 22–23.

Sánchez, R. (2019). Valoración económica de madera en pie de una plantación de Pinus patula Schiede ex Schltdl. & Ccham. en la comunidad campesina de Cumbico - Cajamarca. *Universidad Nacional de Cajamarca*. http://repositorio.unc.edu.pe/handle/UNC/1009

Santiago, W., Ángeles, G., Quiñonez, G., Santos, H., & Rodríguez, G. (2020). Avances y perspectivas en la modelación aplicada a la planeación forestal en México. *Madera y Bosques*, *26*(2). https://doi.org/10.21829/MYB.2020.2622004

Şenyurt, M., Ercanlı, İ., Günlü, A., Bolat, F., & Bulut, S. (2020). Artificial neural network models for predicting relationships between diameter at breast height and stump diameter: Crimean pine stands at ÇAKÜ forest. *Bosque*, *41*(1), 25–34. https://doi.org/10.4067/S0717-92002020000100025

Sotomayor, L., Cracknell, M., & Musk, R. (2023). Supervised machine learning for predicting and interpreting dynamic drivers of plantation forest productivity in northern Tasmania, Australia. *Computers and Electronics in Agriculture*, *209*(December 2022), 107804. https://doi.org/10.1016/j.compag.2023.107804

Suryanto, A., & Muqtadir, A. (2019). Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi. *Saintekbu*, *11*(1), 78–83.

https://doi.org/10.32764/saintekbu.v11i1.298

Tian, H., Zhu, J., He, X., Chen, X., Jian, Z., Li, C., Ou, Q., Li, Q., Huang, G., Liu, C., & Xiao, W. (2022). Using machine learning algorithms to estimate stand volume growth of Larix and Quercus forests based on national-scale Forest Inventory data in China. *Forest Ecosystems*, *9*(March). https://doi.org/10.1016/j.fecs.2022.100037

Tronson, N. (2020). Uncertainty versus prediction error in Pavlovian fear conditioning: Commentary on Walker et al. (2019). *European Journal of Neuroscience*, *52*(5), 3485–3486. https://doi.org/10.1111/EJN.14578

Wager, S. (2019). *Cross-Validation, Risk Estimation, and Model Selection*. 1–6. http://arxiv.org/abs/1909.11696

Wang, L., Zhou, X., Zhu, X., Dong, Z., & Guo, W. (2016). Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop Journal*, *4*(3), 212–219. https://doi.org/10.1016/j.cj.2016.01.008

Ying, G., Li, Z. Y., Chen, E. X., & Qi-Sheng, H. (2011). Estimation of forest biomass using support vector machines from comprehensive remote sensing data. *2011 International Conference on Remote Sensing, Environment and Transportation Engineering, RSETE 2011 - Proceedings*, 2146–2150. https://doi.org/10.1109/RSETE.2011.5964732

Zamora, I. A., Sanz, M. C., & Pando, V. (2004). *Madera aserrada que emplea variables de árboles en pie para pinus radiata D . DON*. *140*, 135–140.

Zhu, Z., Kleinn, C., & Nölke, N. (2021). Assessing tree crown volume - A review. *Forestry*, *94*(1), 18–35. https://doi.org/10.1093/forestry/cpaa037

CAPÍTULO VII. ANEXO

Anexo 1

Obtención de coordenadas de la parcela ubicada

Anexo 2

Enumeración de los árboles en la parcela ubicada con spray

Anexo 3

Obtención de datos de dap y altura

Anexo 4

Cubicación de árboles

Anexo 5

Asesor y amigos que apoyaron en la ejecución de esta tesis

Anexo 6

Datos generales de las plantaciones inventariadas

Distrito	Lugar	Propietario	Año	Este	Norte	Altitud	Parcela
Chota	El Mirador	Teófilo Irigoin Delgado	2012	766617.43	9274716.33	3435	1
Chota	El Mirador	Teófilo Irigoin Delgado	2012	766792.96	9274871.76	3478	2
Chota	Palma	Segundo	2004	762024	9281591	2895	3
Chota	Shotorco	Alcides Sempertegui Ticlla	2004	765128	9278043	3142	4
Conchan	Conchan	MDC	1998	6°25.6887	078° 38 4081	2408	5
Chalamarca	Chalamarca	Octavio Tirado Edquen	2000	778852	9278364	2759	6
Chota	Choctapata Alto	José Rolando Bustamante Cieza	2004	765961	9274717	3303	7

Chiguirip	Vista Alegre	Ananías Torres Mejía	2007	752042	9288329	2626	8
Chota	Leoneropampa	Rosa Vásquez	2004	752698	9283988	3148	9
Chota	La Quinua	Hermilano Oblitas Barboza	2006	764709	9279865	3123	10
Chalamarca	Conga el verde	Juan Fernández Sánchez	2004	772277	9283221	3182	11
Tacabamba	Conga de chucmar	Leoncio Guevara Sánchez	2005	772698	9288547	3388	12
Chota	Bellandina	María Eugenia Diaz Vásquez	2005	766434	9276540	3422	13
Chota	Samangay	Miguel Bustamante Diaz	2006	766107	9266102	3364	14

Lajas	El Aliso	Esperanza Hoyos Diaz	2006	746955	9271625	2745	15
Chota	Cochopampa	María Aguilar Sánchez	1994	758314	9276126	2448	16
Chota	Nuevo Oriente	María Gálvez Mejía	2003	767188	9270122	3184	17
Llama	Pargo	Julio Montenegro Barboza	1999	715391	9283156	2875	18
Huambos	Llamocuna	José Jorge Diaz Arrasco	2000	721157	9286662	2451	19
Huambos	Huambos	Alicia Castro Tentalean	2005	726603	9286362	2367	20
Querocoto	Querocoto	Abel Ortiz Pérez	2007	717932	9297331	2404	21

Anexo 7

Datos de todas las plantaciones inventariadas

parcela	arbol	eje	cap	dap	h	HD
1	1		76.0	24.2	9	12.35
1	2		36.2	11.5	9.7	12.35
1	3		66.5	21.2	9.4	12.35
1	4		45.3	14.4	10.9	12.35
1	5		75.6	24.1	9.6	12.35
1	6		62.6	19.9	8	12.35
1	7		63.5	20.2	10.1	12.35
1	8		68.0	21.6	10.7	12.35
1	9		80.0	25.5	11.4	12.35
1	10		69.6	22.2	10.9	12.35
1	11		73.2	23.3	10.9	12.35
1	12		74.2	23.6	10.9	12.35
1	13		58.5	18.6	9.4	12.35
1	14		44.7	14.2	10.1	12.35
1	15		16.3	5.2	11.4	12.35
1	16		82.1	26.1	9.5	12.35
1	17		49.3	15.7	9.3	12.35
1	18		85.5	27.2	10.5	12.35
1	19		59.5	18.9	9.7	12.35
1	20		47.3	15.1	9.8	12.35
1	21		70.8	22.5	9.8	12.35
1	22		67.0	21.3	10.5	12.35
1	23		67.2	21.4	9.5	12.35
1	24		69.9	22.2	12.6	12.35
1	25		58.5	18.6	11.5	12.35
1	26		69.0	22	10.6	12.35
1	27		95.4	30.4	11.5	12.35
1	28		50.6	16.1	9.8	12.35
1	29		51.6	16.4	9.4	12.35
1	30		87.1	27.7	10.8	12.35
1	31		47.0	15	8.6	12.35
1	32		56.9	18.1	9.6	12.35
1	33		51.8	16.5	9.8	12.35
1	34		68.0	21.6	9.7	12.35
1	35		79.3	25.2	10.3	12.35
1	36		68.7	21.9	10.9	12.35
1	37		34.8	11.1	10.3	12.35
1	38		41.3	13.1	11.2	12.35
1	39		60.2	19.2	11.3	12.35
1	40		63.1	20.1	12.7	12.35
1	41		54.0	17.2	11.7	12.35
1	42		63.8	20.3	10.9	12.35
1	43		61.6	19.6	11.8	12.35
1	44		46.3	14.7	9.2	12.35
1	45		27.2	8.7	10.4	12.35
1	46		43.4	13.8	10.7	12.35
1	47		26.8	8.5	11.5	12.35
1	48		22.7	7.2	12.7	12.35
1	49		77.0	24.5	11.5	12.35
1	50		68.0	21.6	10.6	12.35
1	51		74.2	23.6	10.5	12.35
1	52		65.6	20.9	11	12.35
1	53		49.8	15.9	11.2	12.35
1	54		65.7	20.9	11.5	12.35
1	55		41.6	13.2	9.4	12.35
2	1			17.3	8.9	11.8
2	2			19	8.6	11.8
2	3			18.2	9.4	11.8
2	4	eje 1		9.6	9.1	11.8
2		eje 2		15.9	9.6	11.8
2		eje 3		16.1	9.9	11.8
2		eje 4		14.7	9.9	11.8
2	5			21	9.8	11.8
2	6			15	9	11.8
2	7			16	9.4	11.8
2	8			19.4	9.6	11.8
2	9			19	9.6	11.8
2	10			18.5	10.2	11.8
2	11			17.6	9	11.8
2	12			20.1	8.8	11.8
2	13			22.1	9.2	11.8
2	14			25.9	9.8	11.8
2	15			18.3	11.5	11.8
2	16			21.7	12.9	11.8
2	17			23.3	10.2	11.8
2	18			16.3	9.5	11.8
2	19			21	10.3	11.8
2	20			18.4	9.4	11.8
2	21			16.7	8.4	11.8
2	22			15.5	8.2	11.8
2	23			19.7	10.6	11.8
2	24			19	10.8	11.8
2	25			17.4	9.5	11.8
2	26			20.9	9.7	11.8
2	27			20.7	10	11.8
2	28			21.7	10.2	11.8
2	29			12.5	10.5	11.8
2	30			20.1	10	11.8
2	31			6	6.4	11.8
2	32			21.3	10.3	11.8
2	33			21.1	9.9	11.8
2	34			21.5	9.9	11.8
2	35			24.9	11.4	11.8
2	36			19.4	10.3	11.8
2	37			16.3	10.8	11.8
2	38			23.4	10.8	11.8
2	39	eje 1		21.8	9.6	11.8
2		eje 2		11.3	8.5	11.8
2	40			16.6	8.5	11.8

parcela	arbol	eje	cap	dap	h	HD
2	41			20.7	10.6	11.8
2	42			19.9	11.3	11.8
2	43			20.3	10.6	11.8
2	44			16.4	9.7	11.8
2	45			20.7	9.6	11.8
2	46			18.5	10.3	11.8
2	47			7.9	6.9	11.8
2	48			22.5	12.7	11.8
2	49			18.9	8.6	11.8
2	50			18	8.3	11.8
3	1		65	20.7	14.6	17.175
3	2	eje 1	56	17.8	12.7	17.175
3		eje 2	36	11.5	8.1	17.175
3	3		77	24.5	14.9	17.175
3	4	eje 1	55	17.5	12.3	17.175
3		eje 2	30	9.5	10.1	17.175
3		eje 3	60	19.1	12.5	17.175
3	5		62	19.7	12.3	17.175
3	6	eje 1	68	21.6	12.8	17.175
3		eje 2	64	20.4	13.3	17.175
3	7		100.2	31.9	18.1	17.175
3	8		46	14.6	10.6	17.175
3	9	eje 1	48	15.3	11.6	17.175
3		eje 2	44	14	11.4	17.175
3	10		18	5.7	4.6	17.175
3	11		48	15.3	10	17.175
3	12		85	27.1	13.6	17.175
3	13	eje 1	37	11.8	11.7	17.175
3		eje 2	48	15.3	15.9	17.175
3	14		100	31.8	17.3	17.175
3	15		74	23.6	17.5	17.175
3	16		72	22.9	14.8	17.175
3	17		92	29.3	15.4	17.175
3	18		53	16.9	12.1	17.175
3	19	eje 1	50	15.9	10.9	17.175
3		eje 2	56	17.8	11.4	17.175
3		eje 3	44	14	11.4	17.175
3	20		70	22.3	12.4	17.175
3	21		58	18.5	16.8	17.175
3	22		52	16.6	13	17.175
3	23		47	15	13.4	17.175
3	24		53	16.9	14.4	17.175
3	25		38	12.1	17.8	17.175
3	26		44	14	12.9	17.175
3	27		73	23.2	15.1	17.175
3	28		75	23.9	17.7	17.175
3	29	eje 1	100	31.8	16.2	17.175
3		eje 2	50	15.9	11.8	17.175
3	30		90	28.6	15.1	17.175
3	31		59	18.8	11	17.175
3	32		88	28	14.7	17.175
3	33		30	9.5	11	17.175
3	34		75	23.9	14.2	17.175
3	35		66	21	14	17.175
3	36		80	25.5	15.1	17.175
3	37		89	28.3	11.2	17.175
3	38		85	27.1	15.1	17.175
3	39		80	25.5	15.1	17.175
3	40		24	7.6	8.7	17.175
3	41		80	25.5	18.5	17.175
4	1		92	29.3	17.3	20.25
4	2		82	26.1	18.2	20.25
4	3		86	27.4	11.8	20.25
4	4		87	27.7	13.5	20.25
4	5		47	15	10.6	20.25
4	6		94	29.9	15.5	20.25
4	7		56	17.8	10.5	20.25
4	8		103	32.8	16.1	20.25
4	9		59	18.8	12.3	20.25
4	10		74	23.6	18.5	20.25
4	11		65	20.7	12.7	20.25
4	12		83	26.4	12.1	20.25
4	13		79	25.1	16.8	20.25
4	14		52	16.6	10	20.25
4	15		102	32.5	16.8	20.25
4	16		78	24.8	14.6	20.25
4	17		76	24.2	15.1	20.25
4	18		99	31.5	16.3	20.25
4	19		85	27.1	18.5	20.25
4	20		61	19.4	14.9	20.25
4	21		101	32.1	18.1	20.25
4	22		64	20.4	14.3	20.25
4	23		87	27.7	14.1	20.25
4	24		59	18.8	11.4	20.25
4	25		104	33.1	24.5	20.25
4	26		104	33.1	22.6	20.25
4	27		93	29.6	15.1	20.25
4	28		41	13.1	12.9	20.25
4	29		106	33.7	22.4	20.25
4	30		43	13.7	13	20.25
4	31		76	24.2	15.2	20.25
4	32		81	25.8	15.1	20.25
4	33		79	25.1	18.6	20.25
4	34		94	29.9	12.6	20.25
4	35		81	25.8	10.1	20.25
4	36		64	20.4	12.5	20.25
4	37		55	17.5	10.4	20.25
4	38		60	19.1	16.4	20.25
4	39		90	28.6	17.1	20.25

parcela	arbol	eje	cap	dap	h	HD
4	40		66	21	14	20.25
4	41		119	37.9	19.2	20.25
4	42		108	34.4	19.5	20.25
4	43		102	32.5	19.3	20.25
4	44		82	26.1	15.1	20.25
5	1		97	30.9	15.2	22.55
5	2		84	26.7	18.2	22.55
5	3		63	20.1	9.8	22.55
5	4		49	15.6	15.6	22.55
5	5		104	33.1	21	22.55
5	6		112	35.7	23.8	22.55
5	7		97	30.9	11.1	22.55
5	8		96	30.6	16.1	22.55
5	9		100	31.8	14.1	22.55
5	10		92	29.3	11.8	22.55
5	11		90	28.6	18.5	22.55
5	12		115	36.6	21.7	22.55
5	13		87	27.7	16.8	22.55
5	14		110	35	19	22.55
5	15		81	25.8	12.1	22.55
5	16		89	28.3	14.6	22.55
5	17		84	26.7	15.1	22.55
5	18		100	31.8	20.5	22.55
5	19		44	14	18.6	22.55
5	20		54	17.2	14.9	22.55
5	21		80	25.5	14.2	22.55
5	22		68	21.6	14.2	22.55
5	23		108	34.4	19.5	22.55
5	24		55	17.5	15.4	22.55
5	25		115	36.6	24.5	22.55
5	26		51	16.2	10.4	22.55
5	27		99	31.5	18.7	22.55
5	28		36	11.5	15.3	22.55
5	29		88	28	18.8	22.55
5	30		118	37.6	21.3	22.55
5	31		134	42.7	19.2	22.55
5	32		130	41.4	22.8	22.55
5	33		115	36.6	20	22.55
5	34		114	36.3	19.7	22.55
5	35		69	22	14.2	22.55
5	36		110	35	22.8	22.55
5	37		89	28.3	16.1	22.55
5	38		93	29.6	16	22.55
5	39		79	25.1	18.8	22.55
5	40		119	37.9	21.7	22.55
5	41		83	26.4	13.1	22.55
5	42		99	31.5	15.6	22.55
5	43		66	21	13.8	22.55
5	44		105	33.4	19.3	22.55
5	45		113	36	24.7	22.55
6	1		85	27.1	16.8	20.625
6	2		71.5	22.8	13.1	20.625
6	3	eje 1	76	24.2	17.5	20.625
6		eje 2	58	18.5	14.2	20.625
6	4	eje 1	68	21.6	20.5	20.625
6		eje 2	45	14.3	12.6	20.625
6	5		93	29.6	17	20.625
6	6		59	18.8	12.3	20.625
6	7		83	26.4	16	20.625
6	8		85	27.1	10.5	20.625
6	9		60.5	19.3	21.2	20.625
6	10		71	22.6	15	20.625
6	11	eje 1	64	20.4	12.8	20.625
6		eje 2	60	19.1	14.5	20.625
6	12		78	24.8	16.4	20.625
6	13		61	19.4	20.5	20.625
6	14		56	17.8	12.1	20.625
6	15		94	29.9	19.6	20.625
6	16	eje 1	64	20.4	15	20.625
6		eje 2	38	12.1	12.8	20.625
6	17		59	18.8	14	20.625
6	18		66	21	14.2	20.625
6	19	eje 1	56	17.8	20.4	20.625
6		eje 2	38	12.1	12.8	20.625
6	20		58	18.5	13.9	20.625
6	21	eje 1	68	21.6	16.5	20.625
6		eje 2	58	18.5	14.6	20.625
6	22		54	17.2	12.8	20.625
6	23		56	17.8	14.3	20.625
6	24		74	23.6	17.7	20.625
6	25		53	16.9	24	20.625
6	26		76	24.2	15.4	20.625
6	27		76	24.2	11.5	20.625
6	28	eje 1	60	19.1	14	20.625
6		eje 2	40	12.7	16.3	20.625
6	29		59	18.8	13.6	20.625
6	30		60	19.1	19.2	20.625
6	31		49	15.6	13.7	20.625
6	32		57	18.1	10.6	20.625
6	33		87	27.7	9.2	20.625
6	34		56	17.8	16.5	20.625
6	35		70	22.3	7.5	20.625
6	36		70	22.3	19.5	20.625
6	37		60	19.1	15.7	20.625
6	38		60	19.1	21.6	20.625
6	39		52	16.6	14.3	20.625
6	40	eje 1	57	18.1	12.5	20.625
6		eje 2	46	14.6	17.6	20.625
6	41		55	17.5	15.9	20.625

parcela	arbol	eje	cap	dap	h	HD
6	42		71	22.6	14.8	20.625
6	43		69	22	13.6	20.625
7	1		91	29	18.5	19.125
7	2		83	26.4	15.4	19.125
7	3		50	15.9	15.1	19.125
7	4		39	12.4	12.7	19.125
7	5		66	21	16.4	19.125
7	6		73	23.2	17.4	19.125
7	7		32	10.2	10.9	19.125
7	8		56	17.8	14.7	19.125
7	9		37	11.8	11.5	19.125
7	10		48	15.3	13.9	19.125
7	11		43	13.7	12.5	19.125
7	12		42	13.4	13.5	19.125
7	13		38	12.1	12	19.125
7	14		95	30.2	19.1	19.125
7	15		87	27.7	18.3	19.125
7	16		69	22	17.4	19.125
7	17		41	13.1	13.8	19.125
7	18		36	11.5	11.1	19.125
7	19		85	27.1	18.8	19.125
7	20		48	15.3	14.3	19.125
7	21		91	29	19.5	19.125
7	22		49	15.6	14.2	19.125
7	23		55	17.5	13.5	19.125
7	24		76	24.2	15.8	19.125
7	25		42	13.4	12.7	19.125
7	26		66	21	15.6	19.125
7	27		25	8	8.9	19.125
7	28		31	9.9	10.6	19.125
7	29		35	11.1	13.1	19.125
7	30		49	15.6	14.6	19.125
7	31		34	10.8	12.9	19.125
7	32		35	11.1	13.1	19.125
7	33		77	24.5	16.7	19.125
7	34		89	28.3	18.1	19.125
7	35		64	20.4	15	19.125
7	36		55	17.5	14.9	19.125
7	37		65	20.7	16.4	19.125
7	38		66	21	16	19.125
7	39		84	26.7	19.4	19.125
7	40		40	12.7	12.4	19.125
7	41		57	18.1	13.6	19.125
7	42		45	14.3	12	19.125
7	43		41	13.1	13.5	19.125
7	44		44	14	12	19.125
7	45		64	20.4	14.7	19.125
7	46		77	24.5	16.9	19.125
7	47		43	13.7	12.4	19.125
7	48		87	27.7	19.3	19.125
7	49		90	28.6	19.5	19.125
7	50		67	21.3	16.4	19.125
7	51		87	27.7	17.5	19.125
7	52		66	21	16.4	19.125
7	53		69	22	16.1	19.125
7	54		75	23.9	17.2	19.125
7	55		71	22.6	18.3	19.125
8	1		83	26.4	16.5	17.65
8	2		51	16.2	13.7	17.65
8	3		43	13.7	14.5	17.65
8	4		56	17.8	14	17.65
8	5		95.5	30.4	18.7	17.65
8	6		38.5	12.3	13.2	17.65
8	7		82	26.1	12.4	17.65
8	8		40	12.7	10.1	17.65
8	9		31.5	10	13.1	17.65
8	10		87.5	27.9	17.9	17.65
8	11		38	12.1	10.5	17.65
8	12		81	25.8	18.4	17.65
8	13		46	14.6	15.3	17.65
8	14		42.5	13.5	13.2	17.65
8	15		53	16.9	9.8	17.65
8	16		56	17.8	15.6	17.65
8	17		89	28.3	16.4	17.65
8	18		32	10.2	14.2	17.65
8	19		55	17.5	15.3	17.65
8	20		41	13.1	13.8	17.65
8	21		42	13.4	10.4	17.65
8	22		44	14	9.9	17.65
8	23		49	15.6	11.3	17.65
8	24		40	12.7	15.6	17.65
8	25		15	4.8	7.4	17.65
8	26		60	19.1	16.8	17.65
8	27		29	9.2	9.7	17.65
8	28		31	9.9	10.3	17.65
8	29		63	20.1	16.6	17.65
8	30		47	15	15.4	17.65
8	31		36	11.5	13.8	17.65
8	32		19	6	9.3	17.65
8	33		54	17.2	12.5	17.65
8	34		39	12.4	10.3	17.65
8	35		44	14	11.4	17.65
8	36		46.5	14.8	12.9	17.65
8	37		80.5	25.6	17.5	17.65
8	38		38.5	12.3	10.3	17.65
8	39		40	12.7	11.2	17.65
8	40		74.5	23.7	16.8	17.65
8	41		36	11.5	11	17.65
8	42		36.5	11.6	11.3	17.65

parcela	arbol	eje	cap	dap	h	HD
8	43		55	17.5	10.5	17.65
8	44		36	11.5	9.8	17.65
8	45		52.5	16.7	11.7	17.65
8	46		51.5	16.4	11.6	17.65
8	47		82.5	26.3	17.9	17.65
8	48		50	15.9	15.4	17.65
8	49		32	10.2	10.4	17.65
8	50		63	20.1	13.2	17.65
8	51		47	15	12.1	17.65
8	52		50	15.9	13.7	17.65
8	53		41	13.1	12	17.65
8	54		33	10.5	11.4	17.65
8	55		23	7.3	9.9	17.65
9	1		89	28.3	18.5	21.35
9	2		73	23.2	18.1	21.35
9	3		67	21.3	16.6	21.35
9	4		109	34.7	20.7	21.35
9	5		68	21.6	16.3	21.35
9	6		57	18.1	15.9	21.35
9	7		88	28	18.7	21.35
9	8		61	19.4	16.3	21.35
9	9		83	26.4	17.6	21.35
9	10		98	31.2	20.4	21.35
9	11		93	29.6	19.9	21.35
9	12		51.5	16.4	15.6	21.35
9	13		48	15.3	14.5	21.35
9	14		73	23.2	17.2	21.35
9	15		49	15.6	14.7	21.35
9	16		82.5	26.3	18.8	21.35
9	17		70	22.3	17.3	21.35
9	18		87.5	27.9	19.5	21.35
9	19		52	16.6	14.6	21.35
9	20		75	23.9	16.2	21.35
9	21		105	33.4	20.9	21.35
9	22		60	19.1	15.4	21.35
9	23		57.5	18.3	15.1	21.35
9	24		102	32.5	21.5	21.35
9	25		48	15.3	14.7	21.35
9	26		88	28	19.6	21.35
9	27		99	31.5	20.5	21.35
9	28		76	24.2	18.6	21.35
9	29		57	18.1	16.3	21.35
9	30		69	22	17.1	21.35
9	31		74	23.6	18.4	21.35
9	32		88	28	18.1	21.35
9	33		103	32.8	19.7	21.35
9	34		78	24.8	17.5	21.35
9	35		47	15	14.6	21.35
9	36		106	33.7	21.6	21.35
9	37		81	25.8	18.4	21.35
9	38		32	10.2	14.1	21.35
9	39		65	20.7	15.8	21.35
9	40		72	22.9	16.7	21.35
9	41		44	14	14.3	21.35
9	42		113	36	21.6	21.35
9	43		42	13.4	13.9	21.35
9	44		80	25.5	19.5	21.35
9	45		100	31.8	21.3	21.35
9	46		90	28.6	19.7	21.35
9	47		64	20.4	16.4	21.35
9	48		98	31.2	19.6	21.35
9	49		95	30.2	19.8	21.35
9	50		60	19.1	16.4	21.35
10	1		85	27.1	19.3	19.275
10	2		63	20.1	14.5	19.275
10	3		59	18.8	15.1	19.275
10	4		70	22.3	16.4	19.275
10	5		56	17.8	13.2	19.275
10	6		77	24.5	16.6	19.275
10	7		95	30.2	19.3	19.275
10	8		78	24.8	17.1	19.275
10	9		54	17.2	13.5	19.275
10	10		72	22.9	16.9	19.275
10	11		68	21.6	16.2	19.275
10	12		79	25.1	18.6	19.275
10	13		47	15	13.7	19.275
10	14		76	24.2	17.4	19.275
10	15		64	20.4	15.6	19.275
10	16		51	16.2	14.4	19.275
10	17		69	22	15.8	19.275
10	18		48	15.3	13.3	19.275
10	19		95	30.2	20.4	19.275
10	20		84	26.7	18.7	19.275
10	21		72	22.9	17.2	19.275
10	22		96	30.6	20.8	19.275
10	23		49	15.6	13.3	19.275
10	24		63	20.1	14.6	19.275
10	25		78	24.8	18.1	19.275
10	26		71	22.6	17.2	19.275
10	27		63	20.1	16.4	19.275
10	28		56	17.8	15.3	19.275
10	29		45	14.3	12.8	19.275
10	30		81	25.8	19.1	19.275
10	31		72	22.9	17.5	19.275
10	32		49	15.6	14.5	19.275
10	33		66	21	16.5	19.275
10	34		90	28.6	18.9	19.275
10	35		44	14	13.4	19.275
10	36		62	19.7	14.8	19.275
10	37		48	15.3	12.3	19.275
10	38		73	23.2	15.4	19.275
10	39		56	17.8	16	19.275
10	40		95	30.2	19.3	19.275
10	41		75	23.9	17.8	19.275
10	42		83	26.4	18.1	19.275
10	43		58	18.5	15.4	19.275
10	44		69	22	17.2	19.275
11	1		92	29.3	20.3	20.35
11	2		67	21.3	16.1	20.35
11	3		65.5	20.8	16.5	20.35
11	4		81	25.8	19.4	20.35
11	5		56	17.8	14.3	20.35
11	6		45	14.3	14.1	20.35
11	7		83	26.4	18.5	20.35
11	8		62	19.7	13.4	20.35
11	9		74	23.6	15.6	20.35
11	10		49	15.6	13.6	20.35
11	11		57.2	18.2	16.4	20.35
11	12		65	20.7	16.7	20.35
11	13		80	25.5	18.2	20.35
11	14		90.5	28.8	20.5	20.35
11	15		83	26.4	19.6	20.35
11	16		63	20.1	15.2	20.35
11	17		59.3	18.9	15.8	20.35
11	18		76	24.2	18.4	20.35
11	19		70	22.3	18.1	20.35
11	20		87	27.7	18.5	20.35
11	21		49	15.6	13.6	20.35
11	22		64	20.4	16.4	20.35
11	23		53	16.9	13.4	20.35
11	24		79	25.1	16.9	20.35
11	25		43	13.7	13.3	20.35
11	26		90	28.6	20.4	20.35
11	27		86	27.4	20.2	20.35
11	28		62	19.7	16.1	20.35
11	29		78	24.8	17.4	20.35
11	30		55	17.5	14.4	20.35
11	31		59	18.8	15.9	20.35
11	32		64.5	20.5	16.8	20.35
11	33		77.3	24.6	18.4	20.35
11	34		72	22.9	18	20.35
11	35		58	18.5	14.6	20.35
11	36		93	29.6	19.5	20.35
11	37		68	21.6	16.3	20.35
11	38		72	22.9	17.6	20.35
11	39		68	21.6	17.3	20.35
11	40		54	17.2	15.8	20.35
11	41		83	26.4	19.4	20.35
11	42		87	27.7	19.8	20.35
11	43		90	28.6	20.2	20.35
12	1		69.5	22.1	17.1	20.225
12	2		56	17.8	16.4	20.225
12	3		76	24.2	18.2	20.225
12	4		91	29	20.3	20.225
12	5		85	27.1	19.8	20.225
12	6		88	28	18.4	20.225
12	7		63.5	20.2	15.4	20.225
12	8		72	22.9	17.6	20.225
12	9		66	21	16.4	20.225
12	10		58	18.5	15.3	20.225
12	11		87	27.7	18.5	20.225
12	12		49	15.6	14.1	20.225
12	13		56	17.8	14.7	20.225
12	14		82	26.1	17.8	20.225
12	15		79	25.1	18.5	20.225
12	16		88	28	19.2	20.225
12	17		70	22.3	15.3	20.225
12	18		60	19.1	14.2	20.225
12	19		44	14	12.5	20.225
12	20		57.5	18.3	14.6	20.225
12	21		83	26.4	18.6	20.225
12	22		89	28.3	19.6	20.225
12	23		90	28.6	20.1	20.225
12	24		63	20.1	15.1	20.225
12	25		52.5	16.7	14.7	20.225
12	26		66	21	16.2	20.225
12	27		73	23.2	17.4	20.225
12	28		48	15.3	13.6	20.225
12	29		92	29.3	19.5	20.225
12	30		85	27.1	19.6	20.225
12	31		69.4	22.1	17.1	20.225
12	32		70	22.3	16.5	20.225
12	33		86	27.4	18.3	20.225
12	34		90	28.6	20.4	20.225
12	35		66	21	15	20.225
12	36		54	17.2	13.5	20.225
12	37		72	22.9	16.7	20.225
12	38		83	26.4	17.9	20.225
12	39		45.5	14.5	13.5	20.225
12	40		73	23.2	16.8	20.225
12	41		79	25.1	17.4	20.225
12	42		82	26.1	18.3	20.225
12	43		68	21.6	15.8	20.225
12	44		49.4	15.7	13.4	20.225
12	45		89	28.3	18.7	20.225
12	46		77	24.5	17.4	20.225
12	47		69	22	16.5	20.225
12	48		70.5	22.4	16.7	20.225

parcela	arbol	eje	cap	dap	h	HD
12	49		55	17.5	14.4	20.225
12	50		88	28	20.1	20.225
13	1		63	20.1	15.3	19.1
13	2		59.5	18.9	16.5	19.1
13	3		50.5	16.1	14.6	19.1
13	4		65.5	20.8	16.8	19.1
13	5		61	19.4	15.4	19.1
13	6		92	29.3	18.5	19.1
13	7		71	22.6	18.1	19.1
13	8		75	23.9	17.1	19.1
13	9		79	25.1	18.4	19.1
13	10		57	18.1	14.9	19.1
13	11		52.5	16.7	13.6	19.1
13	12		68.5	21.8	15.4	19.1
13	13		67	21.3	16.1	19.1
13	14		85	27.1	19.3	19.1
13	15		47	15	13.7	19.1
13	16		85	27.1	18.8	19.1
13	17		83	26.4	18	19.1
13	18		67	21.3	16.4	19.1
13	19		61	19.4	15	19.1
13	20		72	22.9	13	19.1
13	21		70	22.3	16.7	19.1
13	22		71	22.6	17.2	19.1
13	23		83.5	26.6	16.6	19.1
13	24		54	17.2	13.5	19.1
13	25		74	23.6	15.9	19.1
13	26		61	19.4	16.1	19.1
13	27		92	29.3	19.3	19.1
13	28		35	11.1	11.38	19.1
13	29		70.5	22.4	16.4	19.1
13	30		54	17.2	13.2	19.1
13	31		73	23.2	16.4	19.1
13	32		74.6	23.7	16.5	19.1
13	33		70	22.3	17.2	19.1
13	34		63	20.1	14.2	19.1
13	35		70	22.3	12.8	19.1
13	36		63	20.1	16.7	19.1
13	37		58.5	18.6	13.2	19.1
13	38		77	24.5	19.3	19.1
13	39		53	16.9	16.8	19.1
13	40		86	27.4	18.4	19.1
13	41		68	21.6	15.3	19.1
13	42		85	27.1	18.4	19.1
13	43		63	20.1	14.3	19.1
13	44		37	11.8	11.51	19.1
13	45		76	24.2	15.8	19.1
13	46		70	22.3	15.1	19.1
14	1		61	19.4	13	16.075
14	2		75	23.9	13.6	16.075
14	3		75	23.9	14.5	16.075
14	4		88	28	20	16.075
14	5		40	12.7	10.83	16.075
14	6		54	17.2	13.8	16.075
14	7		62	19.7	14.2	16.075
14	8		47	15	13.2	16.075
14	9		25	8	7.55	16.075
14	10		37	11.8	9.71	16.075
14	11		75	23.9	12.6	16.075
14	12		60	19.1	12.3	16.075
14	13		45	14.3	13.1	16.075
14	14		50	15.9	12.5	16.075
14	15		51	16.2	13.3	16.075
14	16		46	14.6	11.8	16.075
14	17		67	21.3	14.1	16.075
14	18		29	9.2	13.6	16.075
14	19		63	20.1	14.9	16.075
14	20		81	25.8	16.9	16.075
14	21		69	22	18	16.075
14	22		68	21.6	14	16.075
14	23		59	18.8	13.5	16.075
14	24		70	22.3	13	16.075
14	25		70	22.3	10.6	16.075
14	26		68	21.6	18.4	16.075
14	27		57	18.1	14.4	16.075
14	28		54	17.2	12	16.075
14	29		57	18.1	10.6	16.075
14	30		58	18.5	14.9	16.075
14	31		36	11.5	8.5	16.075
14	32		39	12.4	10.4	16.075
14	33		50	15.9	11.4	16.075
14	34		36	11.5	7.2	16.075
14	35		64	20.4	16.3	16.075
14	36		61	19.4	16.3	16.075
14	37		49	15.6	10.3	16.075
14	38		39	12.4	12.9	16.075
14	39		46	14.6	14.8	16.075
14	40		78	24.8	14.3	16.075
14	41		53	16.9	14	16.075
14	42		35	11.1	11.5	16.075
14	43		49	15.6	11.4	16.075
14	44		59	18.8	14	16.075
14	45		36	11.5	8.6	16.075
14	46		24	7.6	6.1	16.075
14	47		47	15	10.1	16.075
14	48		73	23.2	12.3	16.075
14	49		86	27.4	15.4	16.075
14	50		46	14.6	14.3	16.075
14	51		48	15.3	15.8	16.075

parcela	arbol	eje	cap	dap	h	HD
14	52		67	21.3	11.4	16.075
14	53		59	18.8	12.4	16.075
14	54		20	6.4	12.5	16.075
15	1		17.2	5.5	9	15.975
15	2		35.5	11.3	11.5	15.975
15	3		44.1	14	13.8	15.975
15	4		12.6	4	8.5	15.975
15	5		58.1	18.5	15.5	15.975
15	6		33.4	10.6	12.1	15.975
15	7		29	9.2	11.2	15.975
15	8		21.3	6.8	10	15.975
15	9		39.4	12.5	13.6	15.975
15	10		44.5	14.2	13.3	15.975
15	11		66.7	21.2	16.4	15.975
15	12		12.9	4.1	10.4	15.975
15	13		51.8	16.5	14.5	15.975
15	14		46	14.6	13.9	15.975
15	15		13.5	4.3	13	15.975
15	16		20.9	6.7	9.9	15.975
15	17		30.5	9.7	11.5	15.975
15	18		14.8	4.7	9.3	15.975
15	19		42.1	13.4	12.8	15.975
15	20		55.6	17.7	14.9	15.975
15	21		37	11.8	11.5	15.975
15	22		60.3	19.2	14.9	15.975
15	23		37.6	12	13	15.975
15	24		49.6	15.8	14.2	15.975
15	25		15.6	5	9.4	15.975
15	26		54	17.2	14.8	15.975
15	27		41.5	13.2	14.2	15.975
15	28		29.9	9.5	12.5	15.975
15	29		79.2	25.2	15.8	15.975
15	30		60.7	19.3	16.6	15.975
15	31		13	4.1	10.6	15.975
15	32		27.3	8.7	12.4	15.975
15	33		16.4	5.2	9.3	15.975
15	34		52.8	16.8	14.5	15.975
15	35		65.6	20.9	14.6	15.975
15	36		23.4	7.4	11.9	15.975
15	37		33.8	10.8	13.5	15.975
15	38		69.8	22.2	15.8	15.975
15	39		22.4	7.1	11.2	15.975
15	40		14.3	4.6	13.5	15.975
15	41		69.8	22.2	15.7	15.975
15	42		48.1	15.3	11.2	15.975
15	43		27.4	8.7	12.8	15.975
15	44		18.3	5.8	9.5	15.975
15	45		48.1	15.3	13.4	15.975
15	46		27.4	8.7	12.8	15.975
15	47		48.7	15.5	13	15.975
15	48		18.3	5.8	9.7	15.975
15	49		22.2	7.1	10.9	15.975
15	50		35.5	11.3	13	15.975
15	51		74.5	23.7	16.8	15.975
15	52		26.5	8.4	11.2	15.975
15	53		36.7	11.7	12.5	15.975
15	54		12.2	3.9	7.4	15.975
15	55		21.9	7	10.9	15.975
16	1		92	29.3	18.1	23.975
16	2		140	44.6	23.4	23.975
16	3		125.5	39.9	21.9	23.975
16	4		113	36	20.5	23.975
16	5		99	31.5	19.6	23.975
16	6		82.5	26.3	18.9	23.975
16	7		88.5	28.2	18.4	23.975
16	8		145.5	46.3	25.2	23.975
16	9		92	29.3	19.2	23.975
16	10		108	34.4	20.5	23.975
16	11		98.5	31.4	19.2	23.975
16	12		112	35.7	20.9	23.975
16	13		120	38.2	21.3	23.975
16	14		119	37.9	21	23.975
16	15		140.5	44.7	23.6	23.975
16	16		110	35	21.3	23.975
16	17		98	31.2	18.5	23.975
16	18		99	31.5	19.7	23.975
16	19		115	36.6	21.5	23.975
16	20		109	34.7	20.4	23.975
16	21		138	43.9	23.3	23.975
16	22		107	34.1	21.5	23.975
16	23		128	40.7	22.6	23.975
16	24		107	34.1	20.4	23.975
16	25		94.5	30.1	18.9	23.975
16	26		113	36	19.9	23.975
16	27		106	33.7	17.9	23.975
16	28		95	30.2	18.5	23.975
16	29		88	28	18.3	23.975
16	30		74	23.6	16.9	23.975
16	31		95	30.2	20.4	23.975
16	32		122	38.8	22.9	23.975
16	33		141	44.9	23.6	23.975
16	34		82	26.1	17.7	23.975
16	35		92.5	29.4	18.5	23.975
16	36		80	25.5	16.9	23.975
16	37		110	35	20.8	23.975
16	38		97	30.9	19.5	23.975
16	39		91	29	18.3	23.975
16	40		118.5	37.7	22.4	23.975
16	41		79.5	25.3	17.6	23.975

parcela	arbol	eje	cap	dap	h	HD
16	42		77	24.5	17.2	23.975
16	43		137	43.6	23.5	23.975
17	1		55	17.5	15.5	17.95
17	2		96.1	30.6	18.9	17.95
17	3		82	26.1	16.2	17.95
17	4		49	15.6	15.3	17.95
17	5		43.5	13.8	15	17.95
17	6		73	23.2	17.1	17.95
17	7		80	25.5	17.9	17.95
17	8		57.6	18.3	15.6	17.95
17	9		50	15.9	15.2	17.95
17	10		40.5	12.9	14.5	17.95
17	11		49	15.6	15	17.95
17	12		26	8.3	13.2	17.95
17	13		65	20.7	15.8	17.95
17	14		54.4	17.3	15.4	17.95
17	15		36	11.5	13.1	17.95
17	16		101	32.1	19	17.95
17	17		47.3	15.1	14.3	17.95
17	18		49	15.6	14.8	17.95
17	19		69.6	22.2	16.6	17.95
17	20		42	13.4	14.6	17.95
17	21		48	15.3	14.9	17.95
17	22		68.5	21.8	16.2	17.95
17	23		29	9.2	12	17.95
17	24		40	12.7	13.4	17.95
17	25		86.4	27.5	17.9	17.95
17	26		46.4	14.8	15.2	17.95
17	27		34	10.8	14	17.95
17	28		39.5	12.6	13.9	17.95
17	29		54	17.2	14.8	17.95
17	30		79.6	25.3	17	17.95
17	31		70	22.3	16.5	17.95
17	32		61.9	19.7	15.8	17.95
17	33		71	22.6	16.5	17.95
17	34		60.3	19.2	15.4	17.95
17	35		70	22.3	16.8	17.95
17	36		68	21.6	16	17.95
17	37		32	10.2	14.2	17.95
17	38		59.7	19	15.6	17.95
17	39		66	21	16.3	17.95
17	40		73	23.2	16.8	17.95
17	41		56	17.8	14.7	17.95
17	42		48.5	15.4	13.9	17.95
17	43		63	20.1	15.9	17.95
17	44		81	25.8	17.5	17.95
17	45		92.4	29.4	18.4	17.95
18	1		55	17.5	13.8	17.7
18	2		68	21.6	14.3	17.7
18	3		93.5	29.8	17.5	17.7
18	4		69	22	15.6	17.7
18	5		75	23.9	14.6	17.7
18	6		70.5	22.4	15.1	17.7
18	7		95	30.2	16.3	17.7
18	8		82	26.1	17.2	17.7
18	9		48.5	15.4	13.8	17.7
18	10		58	18.5	15.4	17.7
18	11		93.5	29.8	17.5	17.7
18	12		68	21.6	15.1	17.7
18	13		70	22.3	15.6	17.7
18	14		89	28.3	15.9	17.7
18	15		79.8	25.4	16	17.7
18	16		64	20.4	13.9	17.7
18	17		79	25.1	16.4	17.7
18	18		73	23.2	15.5	17.7
18	19		57.4	18.3	14.3	17.7
18	20		106	33.7	18.5	17.7
18	21		89	28.3	16.5	17.7
18	22		98.4	31.3	17.2	17.7
18	23		96	30.6	17.9	17.7
18	24		73	23.2	15.4	17.7
18	25		70	22.3	16	17.7
18	26		61	19.4	13.6	17.7
18	27		62	19.7	13	17.7
18	28		95	30.2	17.5	17.7
18	29		93.5	29.8	17	17.7
18	30		78.8	25.1	15.4	17.7
18	31		62	19.7	13.8	17.7
18	32		89	28.3	15.6	17.7
18	33		88.7	28.2	15	17.7
18	34		80.5	25.6	15.7	17.7
18	35		66	21	14.5	17.7
18	36		78	24.8	15.2	17.7
18	37		74.5	23.7	14.9	17.7
18	38		63	20.1	14.5	17.7
18	39		92	29.3	16.9	17.7
18	40		75.4	24	15.6	17.7
18	41		67	21.3	13.5	17.7
18	42		87.5	27.9	16.3	17.7
19	1		85	27.1	19.5	20.275
19	2		67.6	21.5	16.2	20.275
19	3		45	14.3	14.5	20.275
19	4		76	24.2	19.1	20.275
19	5		67	21.3	16.2	20.275
19	6		83	26.4	18.6	20.275
19	7		35.5	11.3	14.3	20.275
19	8		58	18.5	15.9	20.275
19	9		59	18.8	16.7	20.275
19	10		62	19.7	17	20.275

parcela	arbol	eje	cap	dap	h	HD
19	11		51.8	16.5	15.5	20.275
19	12		64	20.4	16.4	20.275
19	13		77.5	24.7	18.5	20.275
19	14		90	28.6	20.5	20.275
19	15		71	22.6	18.3	20.275
19	16		69.5	22.1	17.2	20.275
19	17		78	24.8	19.6	20.275
19	18		58	18.5	15.9	20.275
19	19		101	32.1	22.4	20.275
19	20		42.7	13.6	12.5	20.275
19	21		63	20.1	17.3	20.275
19	22		64	20.4	15.9	20.275
19	23		89	28.3	20.6	20.275
19	24		62.8	20	16.2	20.275
19	25		81.5	25.9	19.7	20.275
19	26		63	20.1	16.1	20.275
19	27		72	22.9	18	20.275
19	28		53.5	17	15.1	20.275
19	29		62.6	19.9	15.6	20.275
19	30		58	18.5	15.5	20.275
19	31		55	17.5	14.8	20.275
19	32		56.4	18	13.4	20.275
19	33		46	14.6	13	20.275
19	34		72	22.9	18.1	20.275
19	35		48	15.3	14.2	20.275
19	36		57.5	18.3	16.4	20.275
19	37		51	16.2	13.9	20.275
19	38		75.6	24.1	17.3	20.275
19	39		56	17.8	15.5	20.275
19	40		71	22.6	18.6	20.275
20	1		50.5	16.1	13.9	18.3
20	2		58	18.5	14.5	18.3
20	3		61	19.4	15.2	18.3
20	4		72	22.9	17.4	18.3
20	5		85	27.1	18.4	18.3
20	6		57	18.1	13.8	18.3
20	7		81	25.8	17.9	18.3
20	8		62.5	19.9	15.9	18.3
20	9		51	16.2	13.2	18.3
20	10		69	22	14.6	18.3
20	11		90	28.6	18.8	18.3
20	12		43.5	13.8	13.9	18.3
20	13		56	17.8	14.2	18.3
20	14		45	14.3	13.1	18.3
20	15		64	20.4	16.2	18.3
20	16		33.8	10.8	12.7	18.3
20	17		43	13.7	13.3	18.3
20	18		71	22.6	17.2	18.3
20	19		62	19.7	14.7	18.3
20	20		80	25.5	17.8	18.3
20	21		66.4	21.1	15.3	18.3
20	22		74	23.6	16.4	18.3
20	23		82	26.1	17.8	18.3
20	24		46.6	14.8	14.4	18.3
20	25		39.5	12.6	13.5	18.3
20	26		29	9.2	12.9	18.3
20	27		81	25.8	17.6	18.3
20	28		70	22.3	17.4	18.3
20	29		39	12.4	13.5	18.3
20	30		73	23.2	16.5	18.3
20	31		77	24.5	17.6	18.3
20	32		60	19.1	15.2	18.3
20	33		55.4	17.6	14.3	18.3
20	34		83	26.4	18	18.3
20	35		62	19.7	15.4	18.3
20	36		74.5	23.7	18.2	18.3
20	37		48	15.3	14.3	18.3
20	38		49.5	15.8	14.7	18.3
20	39		85.5	27.2	18.8	18.3
20	40		62	19.7	15.6	18.3
20	41		54	17.2	14.4	18.3
21	1		42	13.4	13.9	16.825
21	2		38.5	12.3	12.3	16.825
21	3		52	16.6	14.8	16.825
21	4		35	11.1	12	16.825
21	5		43.5	13.8	14.5	16.825
21	6		60	19.1	14.1	16.825
21	7		88	28	17.4	16.825
21	8		32	10.2	12	16.825
21	9		47	15	14.3	16.825
21	10		39	12.4	12.4	16.825
21	11		45	14.3	13.1	16.825
21	12		73	23.2	15.8	16.825
21	13		68	21.6	14.5	16.825
21	14		65	20.7	15	16.825
21	15		54.5	17.3	15.1	16.825
21	16		39	12.4	12.2	16.825
21	17		41	13.1	13.4	16.825
21	18		46	14.6	14	16.825
21	19		36	11.5	12.6	16.825
21	20		71	22.6	15.9	16.825
21	21		63	20.1	14.8	16.825
21	22		69.5	22.1	15.1	16.825
21	23		52	16.6	13.7	16.825
21	24		38.6	12.3	13.6	16.825
21	25		37	11.8	12.5	16.825
21	26		30	9.5	11.7	16.825
21	27		44	14	14.3	16.825
21	28		88	28	17.2	16.825
21	29		58	18.5	15	16.825
21	30		57.5	18.3	14.9	16.825
21	31		64	20.4	14.2	16.825
21	32		53	16.9	14.6	16.825
21	33		77	24.5	14.8	16.825
21	34		41	13.1	13.3	16.825
21	35		44	14	13.9	16.825
21	36		43	13.7	13	16.825
21	37		80.5	25.6	16.8	16.825
21	38		63	20.1	15	16.825
21	39		48	15.3	14.3	16.825
21	40		49	15.6	14.4	16.825
21	41		55	17.5	15.1	16.825
21	42		57	18.1	15.6	16.825
21	43		32	10.2	12.3	16.825

Anexo 8

Codigos aplicados en R

```
library(readxl)

library(ggplot2)

library(dplyr)

windowsFonts(Times = windowsFont('Times New Roman'))

dfinv <- read_excel('Excel/DATOS_PLANTACIONES_150324.xlsx',

            sheet = 'INVENTARIO', range = 'A1:O1006')

dfinv

names(dfinv)

nrow(dfinv)

dfinv

# estatistica descriptiva

summary(dfinv)

#Kolmogorov-Smirnov test (n = 50) del DAP y HT

ks.test(dfinv$dap, "pnorm", mean = mean(dfinv$dap), sd = sd(dfinv$dap))

ks.test(dfinv$h, "pnorm", mean = mean(dfinv$h), sd = sd(dfinv$h))

dfinv_qqplot <- bind_rows(tibble(variable = 'dap (cm)', valor = dfinv$dap),

                  tibble(variable = 'h', valor = dfinv$h))
```

```
tiff(filename = "Figuras/Fig.07_QQplot_dap_h_inventario.tif",
    width = 190, height = 100, units = 'mm', res = 300, compression = "jpeg")
ggplot(dfinv_qqplot, aes(sample = valor)) +
  stat_qq(size = 2.5, shape = 21) +
  stat_qq_line(size = 1, color = 'red')  +
  facet_wrap(~variable,  scale = 'free_y' ) +
  labs(x = "Quantiles teóricos", y = "Quantiles datos de inventario")+
  theme_bw() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10))
dev.off()
# Resultados de Inventario por parcela
df_parcela <- dfinv %>%
  group_by(parcela) %>%
  summarise(across(dap:hd, mean, na.rm= TRUE), across(Vol_mod_Chota, sum, na.rm= TRUE))
df_parcela <- df_parcela |> mutate(vol_ha = Vol_mod_Chota * 10000/400)
df_parcela
```

```
#Numero de arvores por parcela
Arv <- df %>% group_by(parcela) %>%
  summarise(total_count = n(),
            .groups = 'drop')
# Determinacion del Numero de Clases
summary(dfinv$dap)
k <- ceiling(1+3.322*log10(nrow(dfinv)))
k
C <- ceiling((max(dfinv$dap) - min(dfinv$dap))/k)
C
# Grafico de Distribución diamétrica
plot_dap <- ggplot(dfinv, aes(dap)) +
  geom_histogram(colour ='white', fill = '#252525', binwidth = C )+
  scale_x_continuous(limits = c(2, max(dfinv$dap)),
                     breaks = seq(2, ceiling(max(dfinv$dap)), C)) +
  labs(x = 'dap (cm)', y = 'Frecuencia') +
  theme_classic() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
```

```
      axis.text.y = element_text(size = 10),
      axis.text.x = element_text(size = 10))


#Ancho de clase de altura total
Ch <- ceiling((max(dfinv$h) - min(dfinv$h))/k)
Ch
# Grafico de Distribución diamétrica
plot_h <- ggplot(dfinv, aes(h)) +
  geom_histogram(colour ='white', fill = '#252525', binwidth = Ch ) +
  scale_x_continuous(limits = c(4.5, max(dfinv$h)-8),
                     breaks = seq(4.5, max(dfinv$h)-8, Ch)) +
  labs(x = 'h (m)', y = NULL) +
  theme_classic() +
  theme(text=element_text(family='Times', size=12),
      legend.text=element_text(size=12),
      axis.text.y = element_text(size = 10),
      axis.text.x = element_text(size = 10))
tiff(filename = "Figuras/Fig.06_Distribucion_dap_h_inventario.tif",
    width = 190, height = 120, units = 'mm', res = 300, compression = "jpeg")
```

```
ggpubr::ggarrange(plot_dap, plot_h, labels = c("a", "b"),
          ncol = 2, nrow = 1)
dev.off()
#Número de Clases
#Formula de Sturges
nc <- ceiling(1 + log2(nrow(df)))
nc
#Tamaño o ancho de clase
tc <- floor((max(df$dap) - min(df$dap))/nc)
tc
breaks = seq(floor(min(df$dap)), ceiling(max(df$dap)), by = tc)
breaks
# Histogram with Sturges method
ggplot(df, aes(dap)) +
  geom_histogram(color = 1, fill = "gray90",
                 breaks = breaks) +
  scale_x_continuous(limits=c(6,33), breaks=seq(3,33, by = 3)) +
  ggtitle("Distribución Diamétrica Plantaciones de Pinus sp. Chota")
tiff("Distribucion_diamterica.tif", width = 130, height = 100,
```

```
    units = 'mm', res = 300)
ggplot(df, aes(dap, label = ..count..)) +
  geom_histogram(binwidth=3, color = 1, fill = 'palegreen',
                 breaks = breaks) +
  xlab('DAP (cm)') + ylab('Frecuencia') +
  scale_x_continuous(limits=c(6,33), breaks=breaks) +
  scale_y_continuous(limits=c(0,51)) +
  geom_text(stat="bin", size = 4,vjust=-1, breaks = breaks, position="stack")+
  ggtitle("Distribución diamétrica plantaciones de Pinus sp. Chota") +
  theme(plot.title = element_text(size = 9, face = "bold"))
dev.off()
#Relación Hipsométrica
rh1 <- lm(ht ~ I(log(dap)) + hd, df)
summary(rh1)
12222222222222222222222222
### Import libraries
library(tidyverse)
library(modelr)
library(randomForest)
```

```
library(caret)
library(tidymodels)
windowsFonts(Times = windowsFont('Times New Roman'))
#Data
df <- readxl::read_excel("data/volume.xlsx")
df
names(df)
nrow(df)
summary(df)
#Kolmogorov-Smirnov test (n = 50) del DAP y HT
ks.test(df$dap, "pnorm", mean = mean(df$dap), sd = sd(df$dap))
ks.test(df$h, "pnorm", mean = mean(df$h), sd = sd(df$h))


df_qqplot <- bind_rows(tibble(variable = 'dap (cm)', valor = df$dap),
                       tibble(variable = 'h', valor = df$h))
tiff(filename = "Figuras/Fig.09_QQplot_dap_h_arboles_muestra.tif",
     width = 190, height = 100, units = 'mm', res = 300, compression = "jpeg")
ggplot(df_qqplot, aes(sample = valor)) +
  stat_qq(size = 2.5, shape = 21) +
```

```
  stat_qq_line(size = 1, color = 'red')  +
  facet_wrap(~variable,  scale = 'free_y' ) +
  labs(x = "Quantiles teóricos", y = "Quantiles datos de árboles muestra")+
  theme_bw() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10))
dev.off()
# Grafico de Distribución diamétrica
C <- 4
p_dap <- ggplot(df, aes(dap)) +
  geom_histogram(colour ='white', fill = '#252525', binwidth = C )+
  scale_x_continuous(limits = c(2, max(dfinv$dap)),
                     breaks = seq(2, ceiling(max(dfinv$dap)), C)) +
  labs(x = 'dap (cm)', y = 'Frecuencia') +
  theme_classic() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
```

```
      axis.text.y = element_text(size = 10),
      axis.text.x = element_text(size = 10))
#Ancho de clase de altura total
Ch <- 3
# Grafico de Distribución de alturas
p_h <- ggplot(df, aes(h)) +
  geom_histogram(colour ='white', fill = '#252525', binwidth = Ch ) +
  scale_x_continuous(limits = c(4.5, max(dfinv$h)-8),
                     breaks = seq(4.5, max(dfinv$h)-8, Ch)) +
  labs(x = 'h (m)', y = NULL) +
  theme_classic() +
  theme(text=element_text(family='Times', size=12),
      legend.text=element_text(size=12),
      axis.text.y = element_text(size = 10),
      axis.text.x = element_text(size = 10))

#Grafico de distribucion de dap y h
tiff(filename = "Figuras/Fig.07_Distribucion_dap_h_muestra.tif",
    width = 190, height = 120, units = 'mm', res = 300, compression = "jpeg")
```

```
ggpubr::ggarrange(p_dap, p_h, labels = c("a", "b"),
          ncol = 2, nrow = 1)
dev.off()
#Correlacion de variables dependientes a independientes
df <- df[7:13]
cor <- cor(df, method = "pearson")
cor
tiff("Figuras/Fig.08_Correlacion_variables.tif", units="mm", width=170, height=155, res=300)
corrplot::corrplot.mixed(cor,
                         upper = 'square',
                         lower = 'number',
                         tl.pos = 'lt',
                         addgrid.col = 'gray80',
                         tl.col = 'black',
                         number.cex = 1,
                         tl.cex=1)
dev.off()

#################################################################
```

```
###### seleccion de variables consideradas en el modelo

df <- df |> dplyr::select(vol, dap:hd, altitud)


###### 70% of the sample size

smp_size <- floor(0.70 * nrow(df))

###### set the seed to make your partition reproducible

set.seed(523)

train_ind <- sort(sample(seq_len(nrow(df)), size = smp_size))


###### Split data into training and test sets

train <- df[train_ind, ]

test <- df[-train_ind, ]

################### Modelos volumetricos ###################

###### Modelo volumetrico Schumacher & Hall

sh = vol ~ b0*(dap^b1)*(h^b2)

reg1 <- nls(sh, data = train,

        start = list(b0 = 1, b1 = 0.5, b2 = -1.025),

        control = nls.control(maxiter = 500))

tidy(reg1)
```

```
augment(reg1)

glance(reg1)

#########################################################

fitControl <- trainControl(method = "repeatedcv", number = 10, repeats = 10

###### Ajuste algoritmo linear regression

set.seed(825)

lmFit <- train(vol ~ ., data = train,

               method = "lm",

               trControl = fitControl,

               #verbose = T,

               metric = "RMSE")

lmFit$finalModel

lmFit$results

varImp(lmFit)

ggplot(varImp(lmFit)) +

  ggtitle("Variable Importances Linear Model")

###### Ajuste algoritmo random forest

tg_rf <- data.frame(mtry = seq(1, 4, by =1))
```

```
set.seed(825)

rfFit <- train(vol ~ ., data = train,

               method = "rf",

               trControl = fitControl,

               tuneGrid = tg_rf,

               verbose = FALSE,

               metric = "RMSE")

rfFit

ggplot(varImp(rfFit)) +

  ggtitle("Variable Importances Random Forest")



rfFit_imp <- caret::varImp(rfFit, scale=F)

randomForest::importance(rfFit$finalModel)

plot(rfFit_imp)

###### Ajuste algoritmo Support Vector Machine

tg_svm <- expand.grid(C = c(0.25, .5, 1),

                      sigma = c(0.1, 0.5, 1))

set.seed(825)

svmFit <- train(vol ~ ., data = train,
```

```
            method = "svmRadial",

            trControl = fitControl,

            preProcess = c("center", "scale"),

            tuneGrid = tg_svm,

            verbose = FALSE,

            metric = "RMSE")

svmFit

svmFit$results

svmFit$bestTune

###### Ajuste algoritmo nnet

# Normalizacion de variables

process <- preProcess(df, method=c("range"))

df_ann <- predict(process,df)

summary(df_ann)

## set the seed to make your partition reproducible



train_ann <- df_ann[train_ind, ]

test_ann <- df_ann[-train_ind, ]
```

```
set.seed(825)

nnetFit <- train(vol ~ ., data = train_ann,

                 method = "nnet",

                 trControl = fitControl,

                 #trControl = trainControl(method = "cv"),

                 tuneLength = 10,

                 verbose = FALSE,

                 metric = "RMSE")

nnetFit$result

#########################################################

#### Calculo volumen estimado

train_ann$vol_nnet <- predict(nnetFit, train_ann[-1])

test_ann$vol_nnet <- predict(nnetFit, test_ann[-1])

#Denormalize data

minvec <- sapply(df,min)

maxvec <- sapply(df,max)

denormalize <- function(x, minval, maxval) {

  x*(maxval-minval) + minval

}
```

```
#Volumen estimado en datos de entrenamiento
train_result <- dplyr::as_tibble(Map(denormalize, train_ann, minvec, maxvec))
train_result <- train_result |>
  mutate(vol_SH = predict(reg1, train[2:5]),
         vol_lm = predict(lmFit, train[2:5]),
         vol_rf = predict(rfFit, train[2:5]),
         vol_svm = predict(svmFit, train[2:5]),
         subset = 'train')
train_result
#Volumen estimado en datos de validacion
test_result <- dplyr::as_tibble(Map(denormalize, test_ann, minvec, maxvec))
test_result <- test_result |>
  mutate(vol_SH = predict(reg1, test[2:5]),
         vol_lm = predict(lmFit, test[2:5]),
         vol_rf = predict(rfFit, test[2:5]),
         vol_svm = predict(svmFit, test[2:5]),
         subset = 'test')
test_result
```

```
#########################################################

############### Plot datos observados vs estimados

df_results <- dplyr::bind_rows(train_result, test_result) |>

  dplyr::select(vol, vol_nnet, vol_SH, vol_rf:subset)

names(df_results) <- c('vol_obs', 'ANN', 'S&H', 'RF', 'SVM', 'subset')

df_results <- df_results |>

  pivot_longer(cols = !c(subset, vol_obs),  names_to = "model", values_to = "vol_est") |>

  select(subset, model,  vol_obs, vol_est) |>

  mutate(subset = factor(subset), model = factor(model))

df_results

#Grafico datos observados vs estimados

tiff(filename = "Figuras/Fig.10_Volumen_Observado_Estimado.tif",

     width = 190, height = 180, units = 'mm', res = 300, compression = "jpeg")

ggplot(df_results, aes(vol_obs, vol_est, colour = subset)) +

  geom_point(size = 3, alpha = 0.5) + scale_colour_manual(values=c("black", "red")) +

  lims(x = c(0,1.6), y = c(0,1.6)) +

  geom_abline(intercept = 0, slope = 1, color = 'black', linetype=2) +

  facet_wrap(~model) +

  labs(x = 'Volumen observado (m³)', y = 'Volumen estimado (m³)', color = NULL)+
```

```
  theme_bw() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        legend.direction = 'horizontal',
        legend.position = c(0.85,0.055),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10))
dev.off()
#########################################################
################ Estadisticas de los modelos
model_list <- list(lm = lmFit, rf = rfFit, svm = svmFit, nnet = nnetFit)
res <- resamples(model_list)
summary(res)


est <- summary(res)


#Tabla de estadisticas de los Modelos
estadisticas_train <- dplyr::tibble(subset = 'train',
                          modelo = c('S&H','lm', 'rf', 'svm', 'nnet'),
```

```
                    Rsquared = c(modelr::rsquare(reg1, train), est$statistics$Rsquared[,'Mean']),
                    RMSE = c(modelr::rmse(reg1, train), est$statistics$RMSE[,'Mean']),
                    MAE = c(modelr::mae(reg1, train), est$statistics$MAE[,'Mean']),
                    r = c(cor(train_result$vol, train_result$vol_SH, method = "pearson"),
                        cor(train_result$vol, train_result$vol_lm, method = "pearson"),
                        cor(train_result$vol, train_result$vol_rf, method = "pearson"),
                        cor(train_result$vol, train_result$vol_svm, method = "pearson"),
                        cor(train_result$vol, train_result$vol_nnet, method = "pearson")))
estadisticas_train
estadisticas_test <- dplyr::tibble(subset = 'test',
                        modelo = c('S&H','lm', 'rf', 'svm', 'nnet'),
                        Rsquared = c(modelr::rsquare(reg1, test),
                                postResample(test_result$vol_lm, test_result$vol)['Rsquared'],
                                postResample(test_result$vol_rf, test_result$vol)['Rsquared'],
                                postResample(test_result$vol_svm, test_result$vol)['Rsquared'],
                                postResample(test_result$vol_nnet, test_result$vol)['Rsquared']),
                        RMSE = c(modelr::rmse(reg1, test),
                                Metrics::rmse(test_result$vol, test_result$vol_lm),
                                Metrics::rmse(test_result$vol, test_result$vol_rf),
```

```
                Metrics::rmse(test_result$vol, test_result$vol_svm),
                Metrics::rmse(test_result$vol, test_result$vol_nnet)),
            MAE = c(modelr::mae(reg1, test),
                Metrics::mae(test_result$vol, test_result$vol_lm),
                Metrics::mae(test_result$vol, test_result$vol_rf),
                Metrics::mae(test_result$vol, test_result$vol_svm),
                Metrics::mae(test_result$vol, test_result$vol_nnet)),
            r = c(cor(test_result$vol, test_result$vol_SH, method = "pearson"),
                cor(test_result$vol, test_result$vol_lm, method = "pearson"),
                cor(test_result$vol, test_result$vol_rf, method = "pearson"),
                cor(test_result$vol, test_result$vol_svm, method = "pearson"),
                cor(test_result$vol, test_result$vol_nnet, method = "pearson")))
estadisticas_test
estadisticas <- dplyr::bind_rows(estadisticas_train, estadisticas_test)
estadisticas <- estadisticas |> filter(modelo != 'lm'
########### Graficos de las estadisticas
plot_r2 <- ggplot(estadisticas, aes(x = modelo, y = Rsquared, label = Rsquared, fill =
factor(subset))) +
  geom_col(position = 'dodge', color = "black",) +
  scale_fill_manual(values = c('white', '#252510')) +
```

```
  labs(x = NULL, fill = NULL) +
  theme_light() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10),
        legend.position = 'none'
plot_rmse <- ggplot(estadisticas, aes(x = modelo, y = RMSE, label = RMSE, fill = factor(subset)))
+
  geom_col(position = 'dodge', color = "black",) +
  scale_fill_manual(values = c('white', '#252510')) +
  labs(x = NULL, fill = NULL) +
  theme_light() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10),
        legend.direction = "horizontal",
        legend.position = c(0.5,0.93))
```

```
plot_mae <- ggplot(estadisticas, aes(x = modelo, y = MAE, label = MAE, fill = factor(subset))) +
  geom_col(position = 'dodge', color = "black",) +
  scale_fill_manual(values = c('white', '#252510')) +
  labs(x = NULL, fill = NULL) +
  theme_light() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10),
        legend.position = 'none')


plot_r <- ggplot(estadisticas, aes(x = modelo, y = r, label = r, fill = factor(subset))) +
  geom_col(position = 'dodge', color = "black",) +
  scale_fill_manual(values = c('white', '#252510')) +
  labs(x = NULL, fill = NULL) +
  theme_light() +
  theme(text=element_text(family='Times', size=12),
        legend.text=element_text(size=12),
        axis.text.y = element_text(size = 10),
```

```
      axis.text.x = element_text(size = 10),
      legend.position = 'none')
#Grafico de distribucion de dap y h
tiff(filename = "Figuras/Fig.09_Estadisticas_Modelos.tif",
    width = 190, height = 190, units = 'mm', res = 300, compression = "jpeg")
ggpubr::ggarrange(plot_r2, plot_rmse, plot_mae, plot_r, labels = c("a", "b", 'c', 'd'),
                  ncol = 2, nrow = 2)
dev.off()
#Comparacion de Modelos
compare_models(lmFit, rfFit)
compare_models(lmFit, svmFit)
compare_models(lmFit, nnetFit)
compare_models(rfFit, svmFit)
compare_models(rfFit, nnetFit)
compare_models(svmFit, nnetFit)
## RMSE for nnet Model
rsquare(nnetFit, train_ann)
rsquare(nnetFit, test_ann)
## R2 for ANN Model
```

```
R2_train_nnet <- 1 - sum((train_ann$vol - train_ann$vol_nnet)^2)/sum((train_ann$vol-mean(train_ann$vol))^2)

R2_train_nnet


R2_test_nnet <- 1 - sum((test_ann$vol - test_ann$vol_nnet)^2)/sum((test_ann$vol-mean(test_ann$vol))^2)

R2_test_nnet


## RMSE for ANN Model

RMSE_train_nnet = Metrics::rmse(train_ann$vol, train_ann$vol_nnet)

RMSE_train_nnet


RMSE_test_nnet = Metrics::rmse(test_ann$vol, test_ann$vol_nnet)

RMSE_test_nnet

## MAE for ANN Model

MAE_train_nnet = Metrics::mae(train_ann$vol, train_ann$vol_nnet)

MAE_train_nnet

MAE_test_nnet = Metrics::mae(test_ann$vol, test_ann$vol_nnet)

MAE_test_nnet

## Coeficiente de correlacion valores predichos y observados
```

```
ryy.nnet = cor(test_ann$vol, test_ann$vol_nnet, method = "pearson")

ryy.nnet

## Error porcentual for regression Model

Ereg1 = (test_ann$vol - test_ann$vol_nnet)/test_ann$vol

Ereg1
```