

# Predicción del promedio de altura total de árboles de *Tabebuia rosea* según factores que influyen en la especie

Jose Cascante Solis<sup>1</sup>, Fatima Daniela Saborío Rivera<sup>1</sup>

[jose.cascantesolis@ucr.ac.cr](mailto:jose.cascantesolis@ucr.ac.cr), [fatima.saborio@ucr.ac.cr](mailto:fatima.saborio@ucr.ac.cr)

## RESUMEN

La medición de la altura total de los árboles puede ser muy compleja en la silvicultura, pero es muy necesaria para este campo y esto presenta un gran desafío en la actualidad. La necesidad de crear una ecuación que estime de forma precisa esta altura total de los árboles en la especie *Tabebuia rosea*, es el foco de la investigación. La investigación se centró en la Ciudad Universitaria Rodrigo Facio de la Universidad de Costa Rica en el año 2023, donde se recolectó los datos de un total de 50 árboles de *Tabebuia rosea* que se encontraban dispersos por el campus. Las variables a las que se le recolectaron los datos fueron; diámetro, altura total, altura a la primera rama, diámetro de la primera copa, total de ramas principales, curvatura, sanidad del árbol, dominancia apical. Además, se determinó que la variable respuesta es la altura total. Con base en las variables determinadas se creó el modelo general y se observaron los supuestos importantes; en este caso el de normalidad, homocedasticidad, autocorrelación, relación lineal e independencia. Posteriormente, se realizó la selección de variables del modelo con base en el método del Criterio de Información de Akaike (AIC) donde se realizó la selección hacia atrás. El modelo obtenido se comparó con el general para determinar diferencias significativas, y se prefirió el modelo obtenido. Por último, en relación con la explicación y con la predicción, se obtuvo que el modelo obtenido explica y predice mucho mejor que el modelo general.

**PALABRAS CLAVE:** modelo lineal, residuales, homocedasticidad, predicción, Criterio de Información de Akaike (AIC).

## INTRODUCCIÓN

El desarrollo de ciudades y la alta concentración de población ha provocado un apresurado deterioro ambiental generando un aumento en las emisiones de carbono, temperaturas más altas, una reducción en la cobertura vegetal y pérdida de la biodiversidad, por motivo de la reducción de árboles al punto de desaparecer en gran parte del paisaje urbano (Jiménez, 2013). Una alternativa para disminuir esta problemática es la implementación de la arboricultura urbana (presencia de árboles en zonas urbanas), siendo efectivos en la fijación de dióxido de carbono (en su biomasa); su sombra reduce la temperatura, aporta en la regulación del clima, la fuerza de la precipitación y los vientos, por lo que, los árboles mejoran la calidad de vida de las personas y generan un valor agregado al ambiente urbano (Avellán, 2020). Por consiguiente, conocer la altura de los árboles es fundamental para comprender el estado de estos y, a su vez, para la toma de decisiones silviculturales como lo son podas, desrames, descopes o la corta total del árbol. Sin embargo, el obtener la altura es un dato muy difícil de recolectar, especialmente en árboles adultos muy desarrollados; además la realización de la medición puede estar sujeta a sesgos considerables, y puede estar ligada a altos costos económicos (Arias, 2004; Zúñiga, 2017), por ello está la necesidad de una ecuación que evalúe algunas

---

<sup>1</sup> Estudiantes de Estadística de la Universidad de Costa Rica

características de fácil acceso que den un acercamiento a la altura real de los árboles que se observen, esto se puede obtener mediante un modelo de regresión lineal.

La existencia de relación entre las variables diámetro y la altura, se podría aprovechar para generar modelos alométricos que puedan predecir la altura midiendo únicamente el diámetro del árbol (Guzman et al., 2019; Arias, 2004; Castillo et al., 2018). En cuanto al diámetro de la copa, él está directamente relacionado con la capacidad de crecimiento debido a que un mayor diámetro de copa, el árbol puede generar más fotosíntesis permitiendo un mayor desarrollo y, a su vez, una mayor altura total, al igual esto relaciona la altura de la primera rama viva, ya que, según en la posición en la que se encuentre permite tener un mayor largo de copa y permite generar más fotosíntesis provocando un mayor desarrollo. Por estas razones, las variables, en especial el diámetro de copa, se han utilizado como indicador de vigor y como variable predictora en modelos para evaluar el grado de competencia entre individuos (Arias, 2005).

Estas características pueden ser bastante útiles para la especie *Tabebuia rosea* (Roble sabana), ya que en la arboricultura urbana se puede tener la dificultad de medir la altura de los árboles para la planificación del manejo silvicultural, por este motivo se tiene como objetivo general desarrollar una ecuación que prediga la altura total promedio de los árboles de la especie *Tabebuia rosea* utilizando las variables con mayor grado de predicción. Además, de forma más específica, analizar el comportamiento de la altura total promedio de los árboles de la especie *Tabebuia rosea* en relación con factores que influyen en la misma.

Estos objetivos dan paso a la estructuración de la investigación, sin embargo, se genera una hipótesis que se debe de tener en cuenta, y es que existe al menos una variable predictora que permita predecir el promedio de la altura total del árbol de forma confiable. Si ninguna de las variables que se definieron tienen un alto poder predictivo, esto va a generar que no se pueda crear un modelo con los datos que se obtienen. Por otra parte, si la hipótesis se cumple, se puede dar como resultado una ecuación que estime el promedio de la altura total del árbol, intuyendo que esas variables tienen un alto poder predictivo.

Además, se tiene como hipótesis que entre mayor sea el diámetro de un árbol mayor será su altura total, como menciona Arias (2004) estas dos variables tienen una alta relación, por lo que árboles muy desarrollados tendrán diámetros muy grandes al igual que su altura total será muy alta.

## METODOLOGÍA

El trabajo se realizó en la Ciudad Universitaria Rodrigo Facio, de la Universidad de Costa Rica, específicamente en Finca 1, ubicada en San Pedro de Montes de Oca, en la provincia de San José, en el año 2023, con un área total de 31.50 hectáreas a una altitud aproximada de 1210 msnm (Centro de Investigaciones Geofísicas, 2023).

La recolección de los datos se realizó el día 3 de octubre del año 2023, mediante la medición de 50 árboles de *Tabebuia rosea* (Roble sabana) dispersos en el campus de Ciudad Universitaria Rodrigo Facio, a los cuales se les midió: el diámetro (en centímetros) a 1.30 metros de altura de la base del árbol, mediante la utilización de una cinta diamétrica, además, por medio de la utilización de un hipsómetro óptico de marca Suunto (PM-5/1520 PC), se procedió a tomar la medición de la altura total (en metros) y la altura a la primera rama viva del árbol (en metros), estas mediciones se obtuvieron; primeramente midiendo con una cinta métrica 15 metros horizontales desde la base del árbol, continuamente desde la posición de los 15 metros se procedió a observar por el instrumento el valor obtenido en la base del árbol y el final vertical de la copa, de igual forma se observó el valor

obtenido en la primer rama viva. Se midió el diámetro de la copa (en metros) utilizando una cinta métrica, y se contó el total de ramas principales presentes en el árbol. Además, se evaluó la sanidad del árbol definiendo la ausencia de enfermedades o la presencia de estas, de la misma manera se evaluó la curvatura presente en el árbol principalmente si presenta una curvatura leve o una curvatura muy pronunciada. Finalmente, se evaluó la dominancia apical de forma que un árbol tenía dominancia apical si presentaba crecimiento vertical de un único brote que predominará y reprimiera el crecimiento de brotes laterales, y la ausencia de dominancia apical es la ausencia de este crecimiento vertical de un único brote.

Para el análisis de los datos, dado que se obtuvo la medición de únicamente un árbol con dominancia apical y solo dos árboles presentaban enfermedades, al ser tan pocos datos no se podía crear la regresión lineal para esos casos, por lo que se tomó la decisión de eliminar esos tres datos obteniendo al final una muestra total de 47 árboles, además, se eliminó las variables categóricas dominancia apical y sanidad porque ya no tenían significado tenerlas en el modelo inicial.

Para la creación del modelo se tomaron en cuenta 2 tipos de variables: La variable respuesta que corresponde a la que se quiere estimar haciendo uso de las demás variables llamadas variables predictoras. En este caso, la variable respuesta fue la altura total (en metros) que se planeó estimar a partir de las demás variables (otras características del árbol) usándolas como predictoras.

Las variables predictoras correspondieron a el diámetro a la altura de 1.30 metros de la base del árbol (centímetros), la altura de la primera rama viva del árbol (metros), el diámetro de la copa (centímetros), la cantidad total de ramas principales presentes en el árbol y la curvatura del tronco del árbol.

Además, se usó un modelo lineal para modelar la media condicional de la siguiente manera:

$$\mu_{Y|X} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_5 * X_1 + \beta_7 X_5 * X_2 + \beta_8 X_5 * X_3 + \beta_9 X_5 * X_4$$

X1: El diámetro (en centímetros) a 1.30 metros de altura de la base del árbol.

X2: la altura a la primera rama viva del árbol (en metros).

X3: El diámetro de la copa (en centímetros).

X4: Cantidad total de ramas principales.

X5: Curvatura (Curvatura leve y curvatura pronunciada).

Por otro lado, se muestra en la Figura 6 (Ver anexo) el diagrama de flujo de la investigación, que explica de forma visual los pasos seguidos para llevar a cabo los objetivos planteados, además facilita identificar posibles problemas como la omisión de pruebas o supuestos, a su vez permite tener un control de los pasos realizado permitiendo concluir de forma correcta los resultados obtenidos.

Con respecto al diagrama de flujo; En primer lugar, se definió el tema de la investigación, así como los objetivos y las variables a observar a las cuales se les realizó un análisis exploratorio con el fin de observar su comportamiento. Luego se realizó la selección de variables (50 árboles de la especie *Tabebuia rosea*). Ya teniendo las variables se definió el modelo con las variables ya establecidas anteriormente, a este modelo se le probaron los diferentes supuestos (Normalidad, homoscedasticidad, linealidad, multicolinealidad e independencia) y se analizó si existían valores extremos o de influencia. En el caso de que no se cumplieran los supuestos se aplican medidas remediales y al cumplir los supuestos se realiza una selección de variables por el método de Criterio

de Información de Akaike (AIC) con el cual se hace una selección hacia atrás. Posteriormente, al modelo seleccionado se le probaron los supuestos de normalidad y homoscedasticidad, los cuales si no se cumplían se realizan medidas remediales y al cumplirse se procede a realizar una verificación de este modelo seleccionado mediante la suma de cuadrados de predicción (PRESS), el coeficiente de determinación (R2) y el coeficiente de poder predictivo (P2). Por último, al modelo final se le calcularon los intervalos de confianza y la ecuación estimada.

Como experta y colaboradora de la investigación se contó con la MSR. Ingeniería Forestal Lupita Vargas Fonseca, Profesora, extensionista e investigadora de la Escuela de Ingeniería Forestal, del Instituto de Costa Rica.

Para el análisis de la investigación; gráficos y diferentes pruebas se utilizó el lenguaje de R con la versión 4.2.2 (R Core Team, 2022), utilizando como principales paquetes; *car* (Fox, Weisberg, 2019), *corrgram* (Wright, 2021), *lattice* (Sarkar, Deepayan, 2008), *lmtest* (Achim, Torsten, 2002), además, la significancia utilizada fue de 0.05.

## RESULTADOS

En primer lugar, se procedió a hacer un análisis exploratorio donde se busca conocer, a partir de visualizaciones, el comportamiento de los datos recolectados; Por lo tanto, se presenta un resumen de los estadísticos descriptivos.

### Cuadro 1.

*Estadísticos descriptivos para los árboles de la especie Tabebuia rosea, Ciudad Universitaria Rodrigo Facio.*

Variables	Estadísticos descriptivos			
	Media	Máximo	Mínimo	Desviación estándar
Altura total	13.158	29.000	3.300	6.684
Diámetro	29.855	57.900	6.900	11.765
Altura 1º rama	4.487	11.100	1.200	2.320
Diámetro de copa	9.266	16.730	2.400	3.389
Cantidad de ramas	6.383	15.000	2.000	2.863

Con base en la dispersión de los datos, el Cuadro 1 refleja que el diámetro es la variable con mayor dispersión de sus datos, además se observa que su media es la más alta entre las demás variables observadas. Asimismo, la variable que presenta menor dispersión de sus datos es la altura a la primera rama viva, manteniendo la menor media en relación con las demás variables.

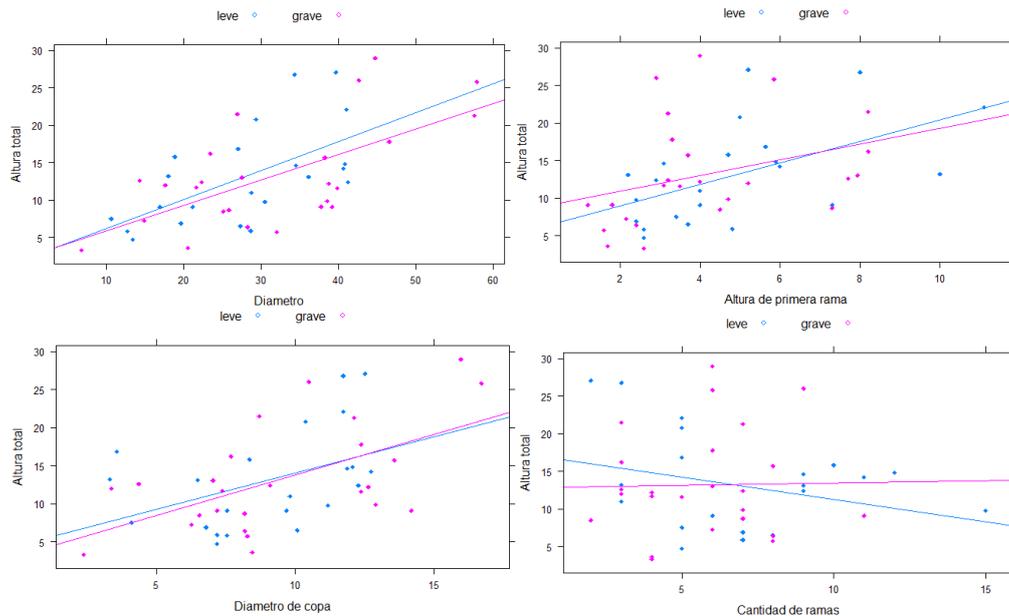
Dentro de los rangos; la variable que presenta un mayor rango es el diámetro y esto se evidencia al tener una desviación estándar alta, mientras que la altura a la primera rama viva es la que genera el menor rango de sus datos.

La relación de la variable respuesta y las predictoras se pueden observar en la Figura 9 (ver Anexo) donde; la relación de la altura total dado el diámetro muestra que conforme aumenta el diámetro, la altura total aumenta en promedio. Esta misma observación sucede en Altura total dado Altura a la primera rama viva y Altura total dado el diámetro de copa, sin embargo, en la relación de Altura total dado la cantidad de ramas muestra que conforme aumenta la cantidad de ramas principales, disminuye la Altura total en promedio.

La variable predictora curva es categórica, por lo que se necesita visualizar si esta variable presenta interacción con las variables predictoras.

**Figura 2.**

*Interacción de la variable curva contra las demás variables predictoras.*



Se observa, en la Figura 2, que la variable curva podría tener interacción contra la variable de cantidad de ramas, sin embargo, más adelante se da un mejor análisis con la comparación de modelos mediante la prueba Fisher, ya que, si realmente existe interacción, esto podría generar complejidad en el estudio. En cambio, con las otras variables se muestra que las líneas presentan una relación muy cercana, por lo que se podría intuir que no va a existir interacción.

Luego de observar la posible presencia de interacción de curva contra las demás variables predictoras, se procedió a evaluar la presencia valores extremos, primeramente, al observar el comportamiento de los datos, en la Figura 10 (ver Anexo) se podría suponer valores extremos, sin embargo, al obtener los residuales estudentizados e identificar si está la existencia de residuales tan altos para que haya valores extremos, se obtiene que no hay existencia de valores extremos.

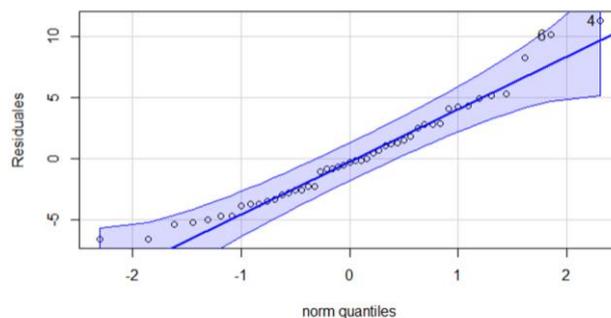
En una regresión lineal, se debe de cumplir ciertos supuestos o reglas los cuales ayudan a que las inferencias que se hagan a partir de una regresión lineal sean válidas y debidamente fundamentadas de forma teórica.

Luego del análisis exploratorio de las variables recolectadas, el análisis de los valores extremos y de influencia, se procedió a verificar los supuestos necesarios.

Con respecto a la verificación de los supuestos, se realizó un gráfico de los residuales contra los cuantiles teóricos para visualizar si a partir del modelo escogido, la distribución condicional de la variable respuesta, que en este caso corresponde a la altura total del árbol dado las variables predictoras es normal, este gráfico se llevó a cabo con los residuales del modelo.

**Figura 3.**

*Distribución de los residuales del modelo general contra los cuantiles teóricos.*

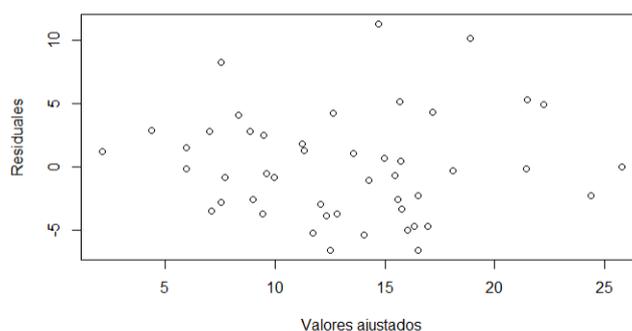


Lo que se busca en la Figura 3 es que los residuales estén dentro de las bandas de confianza, lo que en este caso se satisface, se puede suponer que el modelo cumple el supuesto de normalidad. Además, mediante una prueba formal de Shapiro Wilks con un nivel de significancia de 0.05 se obtuvo como resultado una probabilidad asociada de 0.103, por lo que no hay suficiente evidencia estadística para rechazar la hipótesis nula de que existe normalidad en los residuales y, además, se confirma que se cumple el supuesto de normalidad.

Posteriormente, para verificar si el modelo posee homoscedasticidad se realizó primero una comprobación visual de los residuales del modelo general contra los valores ajustados para así poder detectar algún patrón que indique que las varianzas no son constantes.

**Figura 4.**

*Residuales del modelo general contra los valores ajustados del mismo modelo.*



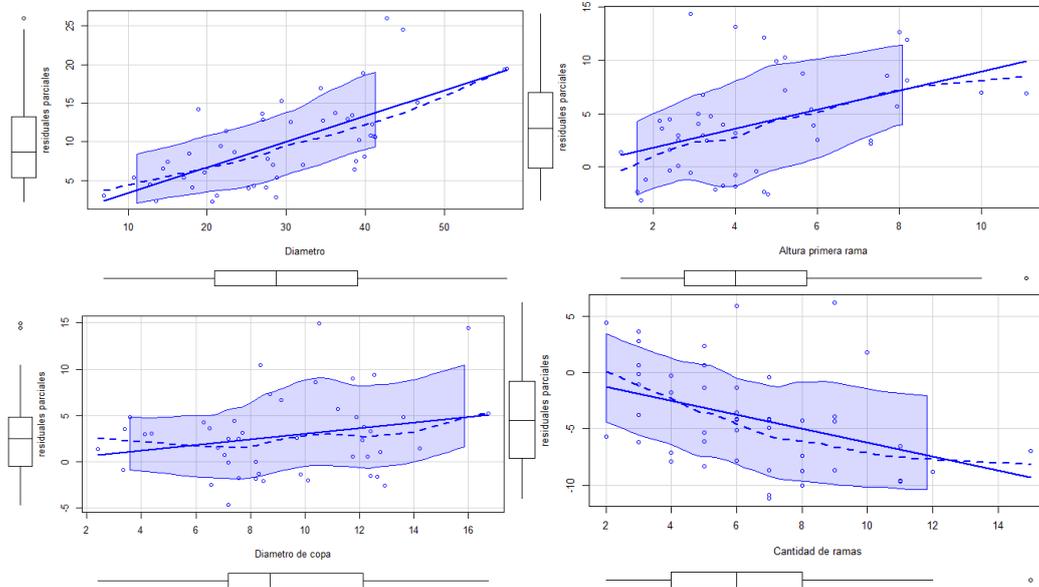
Observando la Figura 4 se aprecia como los residuales parecen seguir un comportamiento aleatorio sin presentar ningún patrón que indique que las varianzas son constantes. Al realizar la prueba formal de Breusch Pagan con un nivel de significancia de 0.05, bajo la hipótesis nula de que la varianza condicional de la respuesta es la misma para cada valor de los predictores, se obtiene una probabilidad asociada de 0.844, por lo que no hay suficiente evidencia estadística para rechazar la hipótesis nula, por lo cual se asume que hay homoscedasticidad, es decir, que las varianzas de cada distribución de la variable respuesta dado las variables predictoras son constantes.

Respecto a la no-multicolinealidad, se analizó si existía correlación de las variables predictoras; primeramente mediante la matriz de correlación lineal de Pearson se observó que las variables diámetro y diámetro de copa tenían una relación lineal de Pearson de 0.810 la cual es alta y positiva, lo que podría dar indicios de presencia de correlación entre esas variables, pero mediante la prueba de factor de inflación de la varianza ningún valor obtenido superó el valor de 10 siendo el máximo

obtenido de 3.190 correspondiente al coeficiente de diámetro, por lo que la varianza de los coeficientes estimados no se ve inflada.

Para la detección de linealidad del modelo, se establecieron los gráficos de linealidad, donde las variables predictoras se evaluaron contra sus residuales parciales;

**Figura 5.**  
*Residuales parciales para observar linealidad*



En la Figura 5 se logra apreciar como en las variables predictoras diámetro, altura primera rama, diámetro de copa mantiene una tendencia lineal creciente, sin embargo, en la variable cantidad de ramas se observa una tendencia lineal negativa. A pesar de que en todas las variables predictoras se observa un comportamiento muy lineal con los residuales parciales, el diámetro es la variable que tiene una línea de tendencia más pronunciada.

Es fundamental mencionar que a pesar de no evaluar si las variables eran independientes entre sí, los casos más comunes de dependencia de variables son los datos de tiempo, datos espaciales y datos conglomerados, por lo que al suponer independencia se obtiene gran variedad de opciones de comprender la incidencia de las variables predictoras con respecto a la variable respuesta.

Luego de poder confirmar los supuestos, se procedió a realizar la selección de las variables del modelo por el método de Criterio de Información de Akaike, donde se hace la selección hacia atrás, como resultado de esta selección se obtuvieron las variables predictoras que ajustan mejor al modelo.

$$\mu_{Y|X} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4$$

Donde:

- X1: El diámetro (en centímetros) a 1.30 metros de altura de la base del árbol.
- X2: la altura a la primera rama viva del árbol (en metros).
- X4: Cantidad total de ramas principales.

Con base en esta selección de variables se comparó el modelo mediante un análisis de varianza del modelo general respecto al modelo obtenido por la selección de variables, se determina que no hubo diferencias significativas ya que el valor obtenido de la prueba Fisher fue de 0.417 mayor al nivel de significancia de 0.05, por lo que no hay suficiente evidencia estadística para rechazar la hipótesis nula de que el aporte marginal del modelo obtenido no presenta diferencias significativas contra el modelo general, por lo que se prefiere el modelo obtenido con el método de Criterio de Información de Akaike.

Es importante mencionar que a pesar de que en la figura 2 parecía que existía interacción principalmente de la variable curva con respecto a las variables diámetro y cantidad de ramas, realmente no existía ya que, con la selección de variables en el modelo obtenido todas las interacciones fueron eliminadas y al realizar la comparación del modelo general que presentaba la interacción y el de las variables seleccionadas, resultó en que la interacción no agrega nueva información al modelo.

Al modelo obtenido por la selección de variables se le evaluó si cumplía con los supuestos de normalidad y homocedasticidad. Primeramente, para la verificación del supuesto de normalidad se realizó un gráfico de los residuales contra los cuantiles teóricos para poder visualizar si la distribución condicional de la variable respuesta, que en este caso corresponde a la altura total del árbol dado las variables predictoras diámetro, altura de primera rama y la cantidad de ramas es normal, este gráfico se llevó a cabo con los residuales del modelo.

Al analizar la figura 7 (Ver anexo) se aprecia como los residuales están dentro de las bandas de confianza, además de tener cercanía con la línea, por lo que se puede afirmar que el modelo cumple el supuesto de normalidad.

Por otro lado, se realizó una comprobación visual de los residuos del modelo obtenido por la selección de variables contra los valores ajustados, con la finalidad de verificar si se tenía homocedasticidad.

Observando la Figura 8 (Ver anexo) se nota que los residuales parecen seguir un comportamiento aleatorio sin presentar ningún patrón, por lo que se puede esperar que se cumpla el supuesto de homocedasticidad. Después de realizar la prueba formal de Breusch Pagan con nivel de significancia de 0.05, bajo la hipótesis nula de que existe homocedasticidad, se obtiene una probabilidad asociada de 0.444, por lo que no hay suficiente evidencia estadística para rechazar la hipótesis nula, por lo cual se asume que hay homocedasticidad, es decir, que las varianzas de cada distribución de la variable respuesta dado las variables predictoras son constantes.

Dado que es necesario conocer cuánto predice y explica el modelo obtenido. Es esencial conocer el poder de predicción, mediante la suma de cuadrados de predicción (PRESS), obteniendo que tan bien se usa los valores ajustados para predecir la respuesta observada es este caso la altura total, y como el valor obtenido del PRESS fue de 983.486 para el modelo con selección de variables, mientras que el modelo general presentó un PRESS de 1214.375, y como se prefiere el modelo con menor suma de cuadrados de predicción, el mejor modelo es el obtenido de la selección de variables por el método AIC.

Además de hacer la comparación de la suma de cuadrados de predicción, se analizó el coeficiente de poder predictivo, dando como resultado 54%, por lo que muestra que el modelo

obtenido tiene calidad predictiva media-alta de la altura de los árboles promedio estimada para cada uno de los valores observados.

Por otra parte, el coeficiente de determinación mostró que las variables diámetro, altura de la primera rama, cantidad de ramas del modelo obtenido explican un 59% de la variabilidad de la respuesta.

Posterior a la predicción del modelo, se realizó una validación del modelo seleccionado por el Criterio de Información de Akaike (AIC), esta resultó de la siguiente manera:

$$\hat{y} = 0.553 + 0.388X_1 + 0.959X_2 - 0.512X_4$$

Donde:

X1: El diámetro (en centímetros) a 1.30 metros de altura de la base del árbol.

X2: la altura a la primera rama viva del árbol (en metros).

X4: Cantidad total de ramas principales.

Para dar una interpretación con un aumento significativo se decide multiplicar cada coeficiente estimado por su desviación estándar, las cuáles son 12, 2 y 3 para diámetro, altura a la primera rama y cantidad de ramas, respectivamente. A partir de esto se puede interpretar que, por cada aumento de 12 centímetros en el diámetro de los árboles, se espera que la altura en promedio aumente 4.559 metros, además, por cada aumento de 2 metros en la altura a la primera rama viva de los árboles, se espera que la altura en promedio aumente 2.225 metros, y también, por cada aumento de 3 en la cantidad de ramas principales de los árboles, se espera que la altura promedio disminuya 1.466 metros.

Asimismo, se obtuvieron los intervalos de confianza para los coeficientes estimados con el mismo aumento en desviaciones estándar, interpretando únicamente los que poseen el mismo signo, los cuales corresponden a las variables diámetro y altura de la primera rama. Se puede mencionar que, se espera que con 95% de confianza que la altura total promedio aumente entre 3.225 y 6.076 metros por un aumento de 12 centímetros de diámetro, manteniendo las demás variables constantes. Además, con el mismo porcentaje de confianza, se obtiene que la altura total promedio aumente entre 0.668 y 3.169 metros por un aumento de 2 metros en la altura a la primera rama viva de los árboles, también manteniendo las demás variables constantes. Respecto a la cantidad de ramas principales, esta variable contiene el cero en su intervalo es por este motivo que no se realizó su respectiva interpretación.

Finalmente, para el intervalo de confianza para la altura total promedio, es importante mencionar que para obtener el intervalo se define un caso al cual se le asigna valores al diámetro, a la altura de la primera rama y a la cantidad de ramas. Por tal razón, se obtiene que con un 95% de confianza, la altura total promedio de los árboles de *Tabebuia rosea* que tengan 20 centímetros de diámetro, una altura a la primera rama viva de 10 metros y una cantidad de 5 ramas está entre 11.602 y 19.066 metros.

## CONCLUSIONES

El modelo resultante contiene tres variables predictoras de un total de 5 variables iniciales, quedando solamente el diámetro, la altura de la primera rama y la cantidad de ramas, resaltando que

no quedaron interacciones en el modelo. Las variables seleccionadas concuerdan con las mencionadas por Arias (2005), donde indica que la altura total puede ser predecida por el diámetro, la altura a la primera rama, ya que, la posición de la primera rama permite un mayor largo de copa relacionándose con un mayor poder fotosintético permitiendo un mayor desarrollo del árbol y por ende una mayor altura.

Estas variables pueden estar muy relacionadas a la altura total promedio del árbol, lo que quiere decir que cuando alguna de ellas aumenta la altura total promedio del árbol también lo hace, lo cual concuerda con las hipótesis realizadas al inicio donde se dijo que cuando aumentaban las variables en especial el diámetro se esperaba que también lo hiciera la altura total.

Es importante mencionar que el comportamiento descrito de las variables por medio de los gráficos, en casi todos los casos fue acertado, una excepción corresponde a que la variable curva con respecto a las variables diámetro y cantidad de ramas, parecían tener interacción según la Figura 2, la cual realmente no existía, ya que, con la selección de variables en el modelo obtenido todas las interacciones se eliminaron.

También se debe resaltar que hay más variables que pueden influir en la altura total promedio del árbol que no pudieron ser analizadas en el estudio, un ejemplo de ellas es la edad, dado que si no se lleva un control desde que se plantó el árbol va a ser difícil obtenerla.

A pesar de lo dicho anteriormente este modelo logra explicar el 59% de la variabilidad de la altura total promedio de los árboles de la especie *Tabebuia rosea*, lo cual es bueno para un modelo que contiene solamente tres variables; además el poder de predicción de este modelo es de 54%, siendo medio-alto. Así este coeficiente está diciendo que el modelo tiene un poder predictivo destacable, por esta razón se puede concluir que se cumplió con el objetivo de este trabajo el cual era predecir el comportamiento de la altura total promedio de los árboles de la especie *Tabebuia rosea* en relación con factores que influyen en la misma.

## BIBLIOGRAFÍA

- Achim, Z., Torsten H. (2002). Diagnostic Checking in Regression Relationships. *R News* 2(3), 7-10. URL. <https://CRAN.R-project.org/doc/Rnews/>
- Arias, D. (2005). Morfometría del árbol en plantaciones forestales tropicales. *Revista Forestal Mesoamericana Kurú*. 2(5), 19-32. <https://revistas.tec.ac.cr/index.php/kuru/article/view/543>
- Arias, D. (2004). Estudio de las relaciones altura-diámetro para seis especies maderables utilizadas en programas de reforestación en la Zona Sur de Costa Rica. *Kurú: Revista Forestal*. 1(2): 1-11.
- Avellán, M. (2020). Gestión del arbolado urbano en Costa Rica: desafíos y oportunidades. *Ambientico*, 275(5), 30-35. [https://www.ambientico.una.ac.cr/wp-content/uploads/tainacan-items/5/31132/275\\_30-35.pdf](https://www.ambientico.una.ac.cr/wp-content/uploads/tainacan-items/5/31132/275_30-35.pdf)
- 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) [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-40182018000100033&lng=es&nrm=iso&tlng=es](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-40182018000100033&lng=es&nrm=iso&tlng=es)
- Centro de Investigaciones Geofísicas. (2023). *Climatología de la estación meteorológica ubicada en el Centro de Investigaciones Geofísicas (CIGEFI)*, Universidad de Costa Rica. <https://cigefi.ucr.ac.cr/climatologia-estacion-cigefi/>
- Fox, J. Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>.
- Guzman, J. Aguirre, O. Gonzalez, M. Treviño, E. Jiménez, J. Vargas, B. Posadas, H. (2019). Relación altura-diámetro para *Abies religiosa* Kunth Schltdl. & Cham. en el centro y sur de México. *Revista mexicana de ciencias forestales*. 10(52), 99- 120. <https://doi.org/10.29298/rmcf.v10i52.483>
- Jiménez, Q. (2013). Arbolado urbano: beneficios, desaciertos y realidad en la Gran Área Metropolitana. *Ambientico*, 232(1), 4-12. [https://www.ambientico.una.ac.cr/wp-content/uploads/tainacan-items/5/24098/232-233\\_4-12.pdf](https://www.ambientico.una.ac.cr/wp-content/uploads/tainacan-items/5/24098/232-233_4-12.pdf)
- Rcore Team. (2022). *R: A Language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Sarkar, Deepayan. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York. ISBN, 978-0-387-75968-5, <http://lmdvr.r-forge.r-project.org>.

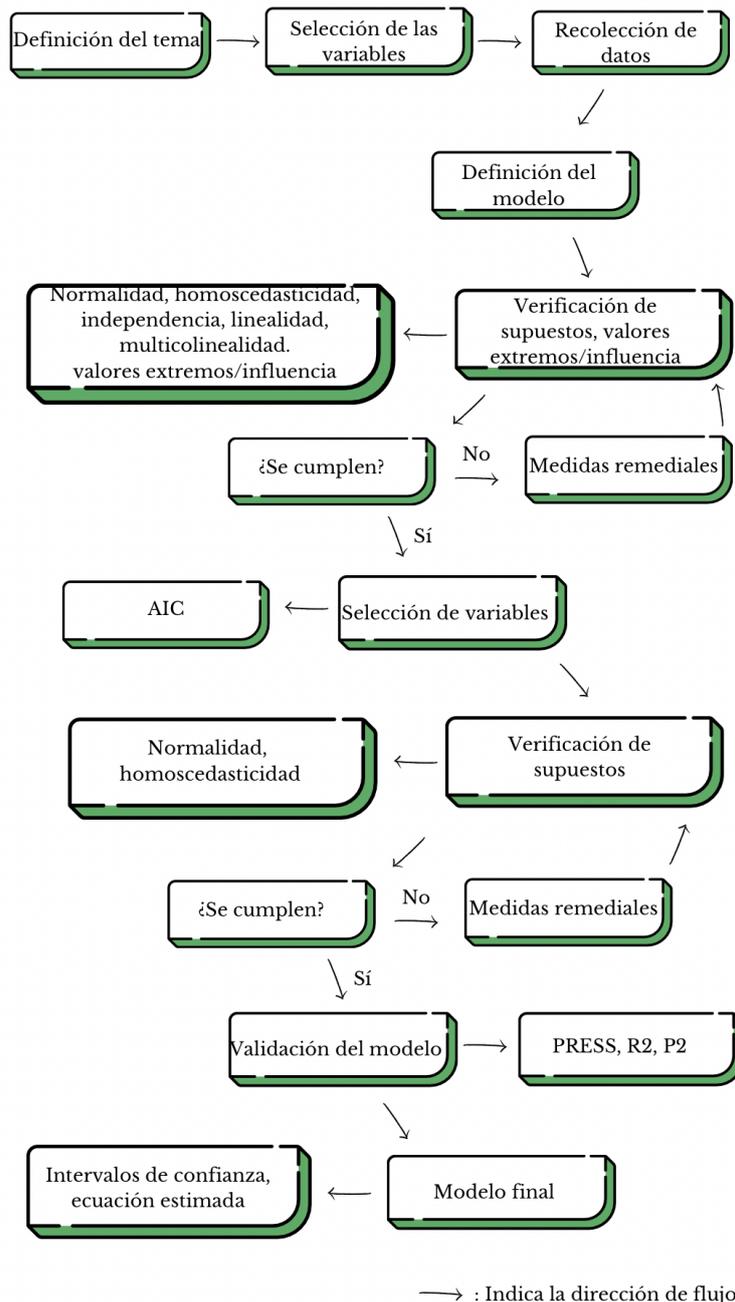
Wright, K. (2021). Corrrgram: Plot a Correlogram. R package version 1.14, <https://CRAN.R-project.org/package=corrgram>.

Zúñiga, B. (2017). *Valoración de la arboricultura y lineamientos para el manejo en 12 parques del cantón de Desamparados, San José, Costa Rica* [Tesis de licenciatura, Instituto Tecnológico de Costa Rica]. Repositorio TEC. <https://repositoriotec.tec.ac.cr/handle/2238/9382>

## ANEXOS

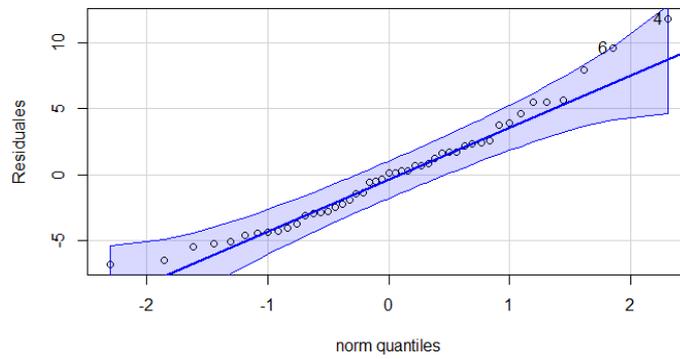
**Figura 6.**

*Diagrama de flujo para el análisis de los datos.*



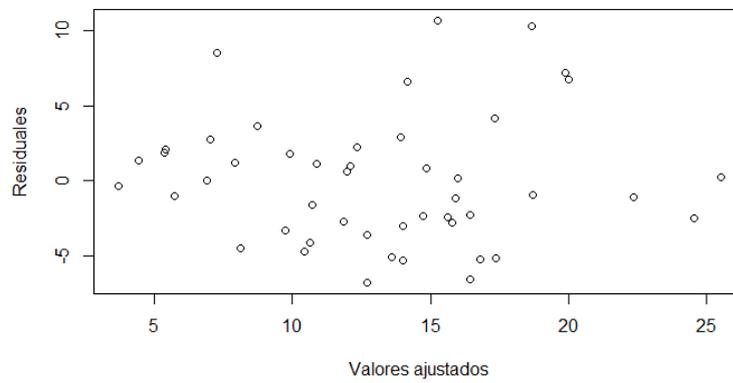
**Figura 7.**

*Distribución de los residuales del modelo obtenido de la selección de variables por el método de Criterio de Información de Akaike contra los cuantiles teóricos.*



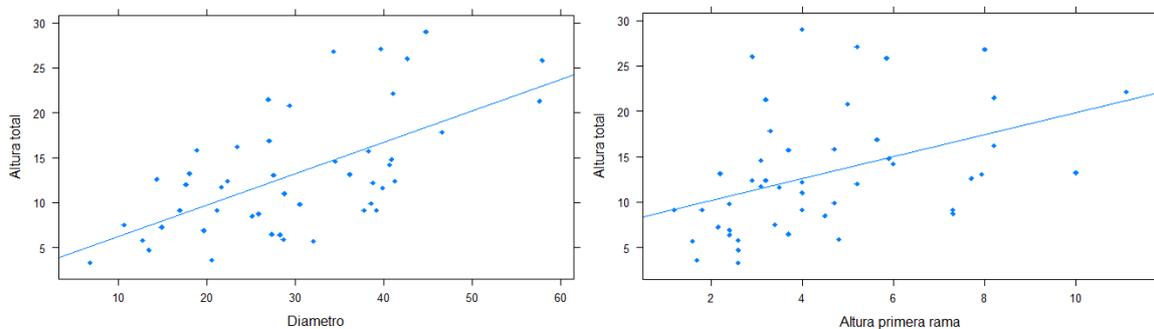
**Figura 8.**

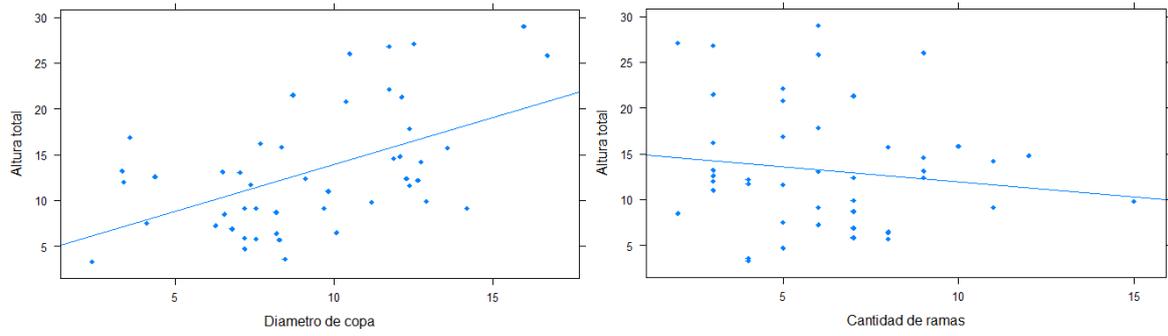
*Residuales del modelo obtenido de la selección de variables por el método de Criterio de Información de Akaike contra los valores ajustados del mismo modelo.*



**Figura 9.**

*Relación gráfica entre la respuesta promedio y las predictoras.*





**Figura 10.**  
*Análisis gráfico de posibles valores extremos*

