EVALUACIÓN DE LA EFECTIVIDAD DEL AJUSTE DE ALGUNOS MODELOS DE REGRESIÓN UTILIZADOS PARA ESTIMAR LA RELACIÓN ALTURA-DIÁMETRO EN PARCELAS PERMANENTES DE RENDIMIENTO Y ACLAREO EN TECA (*Tectona grandis* LINN)

Tania Zambrano[1], Maricela Suárez[1] y Mauricio Jerez[2]
Universidad de Los Andes, [1]Facultad de Ciencias. [2]Facultad de Ciencias Forestales y Ambientales,
Centro de Estudios Forestales de Postgrado, Mérida-Venezuela.
E-mail: zamtania@ciens.ula.ve.

Resumen

Se ensayaron varios modelos de regresión lineal y no lineal para estimar la relación de la altura total (*h*) en función del diámetro a la altura de pecho (*d*) en parcelas permanentes de rendimiento y aclareo establecidas en plantaciones de teca (*Tectona grandis* Linn.) ubicadas en la Unidad Experimental de la Reserva Forestal de Caparo, Barinas, Venezuela. La relación altura-diámetro es empleada por muchos modelos de predicción del rendimiento para estimar la altura promedio de los árboles de un diámetro o clase diamétrica dada. Se analizó la efectividad del ajuste de los modelos de regresión mediante el uso de estadísticos de los residuos que miden la magnitud del sesgo y la precisión del ajuste. Para determinar la habilidad predictiva de los modelos ensayados se utilizaron tres grupos de datos: 1) conjunto de desarrollo (utilizados para los ajustes); 2) datos de prueba (datos no utilizados en la construcción del modelo); y 3) una combinación de los datos de desarrollo y los datos de prueba. Asimismo se evaluó la habilidad de los modelos para estimar la altura total utilizando valores de diámetro a la altura de pecho fuera del rango de diámetros utilizados (extrapolación). Se hacen algunas recomendaciones metodológicas del ajuste de modelos altura-diámetro.

Palabras clave: Teca (*Tectona grandis*), relación altura-diámetro, análisis de regresión.

Abstract

Eleven linear and nonlinear regression models were tested for estimating height-diameter relationships in Teak (*Tectona grandis* Linn.) plantations using data obtained from permanent plots established between 1971 and 1973 at the Caparo Forest Reserve Experimental Station in Barinas, Venezuela. Height-diameter relationships are used in many growth and yield models to predict the mean height for a given diameter at breast height or diameter class. The prediction performance of each regression model was analyzed using the methodology described by Arabatzis and Burkhart (1992) which uses statistics derived from residual analysis for predicting bias and goodness of fit. Three datasets were used to determine the model's predictive ability: 1) model development dataset 2) test data (not used in developing the model) and 3) both datasets combined. In addition, the model's ability to estimate heights from diameters outside the range of the original data (extrapolation) was tested. Several recommendations in relation to height-diameter curve fitting using linear and non-linear regression are given.

Key words: Teak (*Tectona grandis*), height estimation, regression analysis.

Introducción

Uno de los principales problemas en la evaluación del rendimiento a través de parcelas permanentes de rendimiento y aclareo es la estimación del volumen de los árboles en pié. Generalmente, se trata de usar técnicas sencillas, de bajo costo y que a la vez tengan una precisión aceptable. Una de las mas utilizadas consiste en desarrollar ecuaciones del tipo $V = f(d,h)$, donde "V" es el volumen del árbol, usualmente bajo corteza, "d" es el diámetro a la altura de pecho (sobre corteza) y "h" es la altura del árbol (total, comercial, etc.).

La medición de todos los diámetros de una parcela es un proceso relativamente rápido y sencillo, pero las alturas implican un proceso mas lento y costoso

RECIBIDO 11-00 - ACEPTADO 03-01

por lo que, generalmente, sólo se toma una muestra. Esta, usualmente, consiste en una selección aleatoria entre todos los árboles de la parcela, o bien, la selección aleatoria de árboles dentro de categorías diamétricas previamente definidas. Los pares de datos (h, d) obtenidos para esa muestra son utilizados para construir una ecuación que estime la altura de todos los árboles en la parcela a través de su diámetro a la altura de pecho mediante técnicas de regresión lineal y menos frecuentemente técnicas de regresión no lineal.

En la aplicación cotidiana de estas técnicas, suelen cometerse errores en la interpretación de los resultados, debido, por una parte, a que la muestra disponible generalmente es pequeña y de dudosa representatividad; y por otra parte, a que no se toman en cuenta los supuestos básicos de los modelos de regresión.

En la Unidad Experimental de la Reserva Forestal Caparo (Barinas, Venezuela), se vienen efectuando mediciones, desde hace más de 20 años, en parcelas permanentes de rendimiento y aclareo ubicadas en ensayos piloto de teca (***Tectona grandis***) establecidos entre 1971 y 1973. Estas parcelas con superficies entre 800 y 1600 m^2 representan una variedad de regímenes de espesura[1] que incluyen diferentes espaciamientos iniciales, frecuencias, intensidades y edades de aclareo. A partir de estas mediciones se ha obtenido información sobre el rendimiento biológico y económico comparativo de la especie bajo los diferentes regímenes aplicados (Quintero, 1995, Osorio, 1997). Anualmente se miden los diámetros de todos los árboles de cada parcela y, adicionalmente, se toma una muestra de alturas de árboles seleccionados aleatoriamente dentro de categorías diamétricas de 10 cm. Estos datos son ingresados a un sistema de información 'SINFOPLAN' (Zerpa 1989) el cual produce resultados sobre diámetros promedio, densidad, área basal y volumen por hectárea, por parcela y categoría diamétrica.

Al comparar los resultados de volumen utilizando diferentes modelos, se encontró que el factor más crítico en la estimación de los volúmenes no resultó ser la ecuación de volumen, sino la ecuación que establece la relación altura-diámetro que genera las estimaciones de altura para los arboles en la parcela. Dependiendo de la ecuación utilizada, las estimaciones de volumen pueden variar hasta un 30 %, disfrazando las verdaderas diferencias entre regímenes. Se podría sugerir la utilización de una ecuación común para todas las parcelas, sin embargo, Díaz (1989) demostró que existía una importante influencia del régimen de espesura en la relación altura-diámetro para estas plantaciones. Por tanto, es posible que modelos que se ajustan bien para observaciones provenientes de un régimen dado, no sean los mas apropiados para obtener las mejores estimaciones de altura en otros regímenes.

En el presente trabajo se ensayaron varios modelos de regresión de la altura total (h) en función del diámetro a la altura de pecho (d) en parcelas permanentes de rendimiento establecidas en plantaciones de teca y se probó la efectividad del ajuste de regresión de los modelos mediante el uso de estimadores de predicción y precisión.

Antecedentes

En el campo de las plantaciones forestales se han aplicado diferentes técnicas para la selección de modelos que representen en forma adecuada la relación altura-diámetro (Strand 1959, Curtis 1967, Ek 1973, Arabatzis y Burkhart, 1992). Un modelo de altura-diámetro adecuado debe ser moderadamente flexible, es decir ajustarse a una gama de diámetros, y la curva que éste represente debe poseer como características: a) pendiente siempre positiva aproximándose a cero en el caso de diámetros grandes; b) recta que pase por el origen, y c) que pueda ajustarse por regresión lineal (Curtis 1967). Si la presencia de árboles pequeños no es de interés o no están presentes debido que se han realizado aclareos, la característica 'b' no es necesaria. La selección de un modelo en particular dependerá del objetivo del estudio, ya sea este la estimación de parámetros, predicción, extrapolación ó el simple ajuste de la nube de puntos obtenida a partir de una muestra.

Dependiendo del objetivo, los supuestos en los que se basa el análisis de regresión pueden tener mayor o menor importancia. Estos supuestos se pueden resumir en que los errores se distribuyen normal e independientemente con media cero y varianza común ($\varepsilon \sim N(0,\sigma^2)$). En caso de que el objetivo del ajuste sea la estimación, el cumplimiento del supuesto de normalidad es poco importante. Sin embargo, si la finalidad es la predicción y el contraste de hipótesis, el supuesto de normalidad es crítico (Gujarati, 1978).

La violación del supuesto de homogeneidad de varianzas, trae como consecuencia la pérdida de validéz de las pruebas "t" y "F". La presencia de autocorrelación (errores no independientes) entre los residuos trae consecuencias similares, y adicionalmente, hace que los estimadores obtenidos sean muy sensibles a ligeras variaciones en los datos de la muestra (Gujarati, 1978).

Para la comparación de modelos se han utilizado como criterios la suma de cuadrados de los residuos (MSE); los residuos PRESS (suma de cuadrados de predicción); y la significancia estadística de "R^2" entre otros, sin embargo, este último, que es el mas comunmente usado, puede algunas veces no representar adecuadamente el comportamiento del modelo (Crocker, 1972; Barret 1974). En ecuaciones con transformaciones de la variable dependiente, la significancia estadística de "R^2" no puede ser utilizada como medida de efectividad, ni tampoco para establecer comparaciones entre el modelo transformado y el no transformado (Gujarati 1978). En estos casos se utilizan coeficientes de determinación ajustados (R^2 ajus).

En los modelos de regresión múltiple se puede presentar cierto grado de colinealidad entre las variables, lo cual aumenta la probabilidad de aceptar hipótesis falsas y producir valores exagerados de "R^2".

Cuando el objetivo del estudio es obtener ecuaciones para hacer estimaciones y extrapolaciones a una población mayor o para valores de las variables independientes fuera del rango de los valores utilizados para ajustar las curvas, se debe considerar la significancia estadística de los parámetros estimados y el comportamiento general de la curva generada.

En general, el método mas efectivo para para determinar la efectividad del ajuste de regresión de los modelos altura-diámetro es a través del uso de estadísticos derivados de los residuales, ya que estos permiten tomar en cuenta posibles violaciones a los supuestos de la regresión. En el presente trabajo se utilizaron los estadísticos propuestos por Arabatzis y Burkhart (1992) para determinar la efectividad del ajuste de los modelos de regresión utilizados.

Materiales y Métodos

Datos

Los datos utilizados provienen de parcelas permanentes de rendimiento (PPR) con historial similar (edad, densidad, altura y diámetro promedio) ubicadas en plantaciones de teca establecidas en 1973 en el Area Experimental de la Reserva Forestal Caparo (Barinas, Venezuela). En cada PPR se miden anualmente el diámetro de todos los árboles y para las alturas se toma una muestra aleatoria de tres a diez árboles por categoría diamétrica con 10 cm de amplitud. Para el estudio se utilizaron mediciones altura-diámetro a los 20 años de edad provenientes de una muestra de tres parcelas permanentes de rendimiento y aclareo.

Métodos

Se utilizó la Prueba de Barttlet para determinar si la variabilidad en la altura y diámetro entre parcelas presentaba diferencias significativas. De seis parcelas seleccionadas inicialmente se encontró que tres de ellas no presentaron diferencias significativas en los promedios y varianzas de la altura-diámetro. A estas tres parcelas se les hizo un análisis de covarianza (Hinkelmann & Kempthorne, 1994) con el fin de determinar si existe una tendencia similar en la relación altura-diámetro de las diferentes parcelas. Si la relación altura-diámetro de las diferentes parcelas no difiere significativamente, entonces es posible considerarlas como provenientes de una misma población. De las observaciones que conforman la muestra, se seleccionó aleatoriamente una parte para el conjunto de desarrollo a partir del cual se ajustaron los modelos a ensayar y se dejó otro grupo de observaciones para el conjunto de prueba. En la validación de los modelos predictivos se utilizó la metodología propuesta por Bocking (1976) que compara los valores originales, valores de predicción de observaciones no incluídas en la ecuación de ajuste y la combinación de ambos. Posteriormente, para determinar la efectividad del ajuste de regresión de los modelos ensayados se utilizó la metodología propuesta por Arabatzis y Burkhart (1992), la cual utiliza como criterios de comparación los estadísticos de los residuales: media de los residuos ($\bar{e}$); varianza de los residuos (v) y error cuadrático medio (MS). El residuo se define por; donde e_i es el residuo correspondiente a el i-ésimo árbol en la parcela; h_i es la altura observada y es la altura estimada para el i-ésimo árbol a partir del modelo de regresión considerado. El residuo promedio ($\bar{e}_i$), el cual se considera como una estimación del sesgo de predicción para cada parcela se define como:

$$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n}$$

El estimador de la precisión del modelo para cada parcela (n) se calculó como

$$e = \left(\frac{\sum_{i=1}^{n} e_i - \frac{\left(\sum_{i=1}^{n} e_i\right)}{n}}{(n-1)n} \right)$$

y la media cuadrática del error (Mean Square Error), que da una idea de la exactitud del modelo, se obtuvo combinando el sesgo de predicción y la precisión del modelo de la forma siguiente

$$MS_i = \bar{e}^2 + \upsilon$$

El MS es interpretable como una medida del error debido al muestreo que puede a su vez descomponerse en una parte aleatoria expresada a través de la varianza (υ) y una componente sistemática medida por el sesgo ($\bar{e}_i$). Los promedios de $\bar{e}_i$, υ_i y MS de las tres parcelas se calcularon de la siguiente manera:

$$\bar{e} = \frac{\sum_{i=1}^{3} \bar{e}_i}{3}, \qquad \bar{\upsilon} = \frac{\sum_{i=1}^{3} \bar{\upsilon}_i}{3}, \quad y$$

$$\overline{MS} = \frac{\sum_{i=1}^{3} \overline{MS}_i}{3} \text{, respectivamente}$$

Los valores de $\bar{e}$, $\sqrt{\bar{\upsilon}}$ y $\sqrt{\overline{MS}}$ se utilizaron como criterios de comparación de la predicción y eficiencia del muestreo.

Modelos de regresión

Se seleccionó un conjunto de modelos (Cuadro 1) de uso común para la estimación de la relación altura-diámetro. Algunos se caracterizan por ser flexibles, sencillos y de fácil interpretación biológica del comportamiento de los árboles (modelos 1 al 6). Otros, como los modelos polinomiales (7 a 11) a pesar de no tener interpretación biológica son de uso común y se fundamentan en el ajuste de los datos a segmentos de curvas que logran maximizar el "R^2". Todos los modelos utilizados se ajustaron mediante cuadrados mínimos ordinarios (regresión lineal) excepto los modelos 1 y el 2 que se ajustaron mediante regresión no lineal a través del PROCNLIN[2] del SAS, método Marquardt.

De acuerdo a Arabatzis y Burkhart (1992), los modelos 1 y 3 se fundamentan en la relación alométrica de altura-diámetro de la forma funcional $h = ad^b$ (h= altura, d= diámetro a la altura de pecho) Esta función pasa por el orígen pero no es asintótica para diámetros grandes. Ambos modelos no son equivalentes y no pueden obtenerse de tomar el logaritmo o antilogaritmo del otro. La diferencia entre los dos modelos es debida al término del error, en el modelo 1 el error es aditivo y en el 3 es multiplicativo de la forma no transformada $h = ad^b(1+\varepsilon)$ (Myers 1966). Los modelos 2 y 4 están basados en las funciones de potencia de la forma $h = b_0 \, x \, exp^{b/d}$ los cuales se comportan asintóticamente para diámetros grandes; el modelo 4 implica una transformación logarítmica. Se debe tener en cuenta que los ajustes logarítmicos tienden a producir subestimaciones de la altura y la distribución de los residuos está influenciada por la transformación de la variable dependiente, lo cual puede conducir a la obtención de curvas sesgadas (Arabatzis y Burkhart, 1992). Los modelos 5 y 6 son asintóticos para diámetros grandes. El modelo 5 puede dar estimados de altura negativos para árboles con diámetros muy pequeños, pero generalmente, esto ocurre fuera del ámbito de los diámetros de interés. Los modelos 7 al 9 son de tipo polinómico, muy comunes en estudios dasométricos, pero que deben usarse con mucho cuidado, sobre todo si se desea hacer extrapolaciones, ya que ajustes con polinomios superior al grado 3 resultan en curvas serpenteantes y sin sentido (Little and Jackson Hills, 1987). El modelo 10 (Díaz, 1989) es de tipo mixto y el modelo 11 es del tipo "recíproco".

Cuadro 1. Modelos propuestos para ajustes de curvas altura-diámetro en parcelas permanantes en plantaciones de teca de la Reserva Forestal de Caparo, Barinas-Venezuela.

Nro.	Forma del modelo	Método de estimación	Estimador de la altura
1	$h = ad^{b} + \varepsilon$	No lineal M.C.	$\hat{h} = \hat{a}d^{\hat{b}}$
2	$h = ae^{b/d} + \varepsilon$	No lineal M.C	$\hat{h} = \hat{a}e^{\hat{b}/d}$
3	$\ln h = a + b\ln d + \varepsilon$	Lineal M.C.	$\hat{h} = e^{\hat{a}}d^{\hat{b}}$
4	$\ln h = a + b(1/d) + \varepsilon$	Lineal M.C.	$\hat{h} = e^{\hat{a}+\hat{b}(1/d)}$
5	$h = a + b\ln d + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}\ln d$
6	$h = a + b(1/d) + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}(1/d)$
7	$h = a + bd + cd^{2} + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}d + \hat{c}d^{2}$
8	$h = a + bd + ... + ed^{4} + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}d + ... + \hat{e}d^{4}$
9	$h = a + bd + ed^{4} + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}d + \hat{e}d^{4}$
10	$h = a + b\ln d + c(1/d) + e\sqrt{d} + \varepsilon$	Lineal M.C.	$\hat{h} = \hat{a} + \hat{b}\ln d + \hat{c}(1/d) + \hat{e}\sqrt{d}$
11	$h^{-1} = a + bd + cd^{2} + \varepsilon$	Lineal M.C.	$\hat{h} = 1/(\hat{a} + \hat{b}d + \hat{c}d^{2})$

M.C.= mínimos cuadrados

ε es el término del error; $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$

$\hat{a}$, $\hat{b}$, $\hat{c}$, $\hat{e}$ son los estimadores de a, b, c y e respectivamente.

Resultados y Discusión

En el Cuadro 2 se presentan los resultados de la prueba de Bartlett, los cuales mostraron que las varianzas en altura y diámetro de los árboles provenientes de las parcelas seleccionadas no son significativamente diferentes. Los resultados del análisis de covarianza (Cuadro 3) revelan que no hay evidencias significativas para rechazar la hipótesis de que las líneas de regresión ajustadas individualmente no difieren entre sí, por lo tanto los datos altura-diámetro entre las parcelas pueden agruparse con el fin de aumentar el tamaño de la muestra. Las estadísticas de la muestra aleatoria (60 observaciones) seleccionada para el ajuste de los modelos ensayados se presentan en el Cuadro 4. En la Figura 1 se presenta la nube de puntos formada por los pares de altura-diámetro observados.

Efectividad del Ajuste

Los modelos que incluyen más de una variable independiente (7 a 11) presentaron coeficientes de determinación ajustados (R^2ajus) prácticamente iguales a los de los modelos de una sola variable. Las estimaciones de los parámetros y coeficientes de determinación de los modelos ajustados por regresión se incluyen en el Cuadro 5. Los modelos polinómicos 8 y 10, a pesar de presentar un alto R^2ajus, muestran todos los coeficientes de regresión no significativos. Esto sugiere la presencia de alta colinealidad entre las variables. En el modelo polinómico 9 la prueba

Cuadro 2. Prueba de homogeneidad de varianza para la altura y diámetro de la muestra de árboles provenientes de las parcelas.

	Variación altura		Variación diámetro	
	valor	probabilidad	valor	probabilidad
Bartlett	1,00507	0,7969	1,01306	0,5589

Cuadro 3. Resultados del análisis de covarianza en las parcelas seleccionadas.

Coeficientes	Estimador	"t"	Probabilidad	
$\hat{\alpha}_0$	13.800850	10.742	0.0000	(**)
$\hat{\alpha}_1$	0.111780	0.056	0.9557	(ns)
$\hat{\alpha}_2$	-0.189625	-0.084	0.9336	(ns)
$\hat{\beta}_1$	0.388026	7.831	0.0000	(**)
$\hat{\beta}_2$	0.012069	0.151	0.8801	(ns)
$\hat{\beta}_3$	-0.002461	-0.027	0.9783	(ns)
F= 27.9182 R^2= 0.61335				

t: Prueba "t" de significancia
ns: No significativo
**: Significativo a 0,01

Cuadro 4. Estadísticas de la masa forestal para las parcelas seleccionadas.

Variable	Unidad	Parcela 1	Parcela 2	Parcela 3
Año[1]		1973	1973	1973
Densidad	arb/ha	542	575	500
dap	cm	25.19 ± 0.83[2]	24.3 ± 0.80	25.43 ± 0.87
Altura	m	23.74 ± 0.40	23.6 ± 0.40	23.67 ± 0.42
n		28	31	35

1 Año de plantación
2 Error estándar

"*t*" señaló altamente significativo el coeficiente lineal y no significativo el término de cuarto grado. Similar comportamiento mostró el modelo polinómico 7. El único modelo de regresión múltiple con todos los coeficientes significativos fue el modelo 11. Los modelos de regresión simple (1 a 6) presentaron R^2ajus muy similares independientemente de la utilización de variables transformadas o del método de ajuste empleado (regresión lineal vs. no lineal). Asimismo, todos los coeficientes de regresión resultaron significativos, con excepción de los interceptos de los modelos 1 y 5.

El cuadro 6 muestra los resultados estadísticos de los residuos para los tres conjuntos de datos

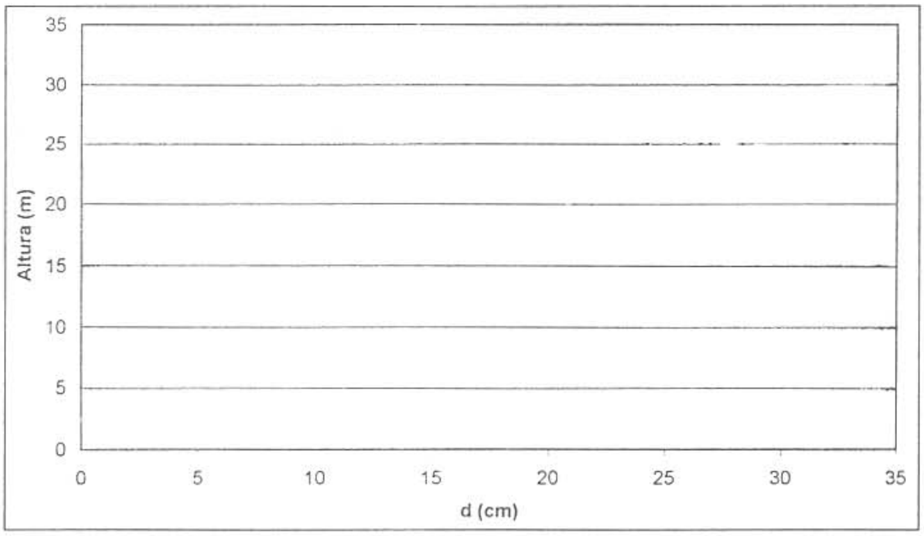

Figura 1. Relación altura-diámetro para los datos del conjunto de desarrollo.

Cuadro 5. Resultados de los análisis de regresión para los modelos seleccionados.

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
Coeficientes de regresión						
a	7,2959355 (ns)	33,6114369 (**)	1,9574320 (**)	3,5148780 (**)	-3,823680 (ns)	31,7444730 (**)
b	0,3641535 (*)	-8,7143811 (*)	0,3729850 (**)	-8,758087 (**)	8,5256130 (**)	-199,43050 (**)
R2	0,604329	0,61066	0,605861	0,610688	0,608407	0,6086071
R2 ajustado			0,5991	0,6040	0,601656	0,601859

	Modelo 7	Modelo 8	Modelo 9	Modelo 10	Modelo 11
Coeficientes de regresión					
a	6,767296 (ns)	-99,084067 (ns)	10,242204 (**)	-462,920188 (ns)	0,079625 (**)
b	1,0095410 (*)	20,277082 (ns)	0,599321 (**)	234,748344 (ns)	-0,002318 (**)
c	-0,0131981 (ns)	-1,294509 (ns)		1677,200303 (ns)	0,000033 (*)
d		0,036911 (ns)			
e		-0,000389 (ns)	-3,700857E-6 (ns)	-63,201448 (ns)	
R2	0,615155	0,637619	0,619187	0,61654	
R2 ajustado	0,601652	0,611264	0,605825	0,595997	0,5997

ns no significativo
* significativo a 0,05
** significativo a 0,01

considerados. En los modelos ajustados al grupo de datos de desarrollo se obtuvo la mayor efectividad de ajuste (menores valores de $\sqrt{MS}$). Para el conjunto de datos de desarrollo más los datos de prueba se mejoró notablemente la precisión debido a que arrojaron los menores valores de $\sqrt{v}$. Sin embargo, este aumento de la precisión ($\sqrt{v}$) combinado con un aumento del sesgo de predicción ($\overline{e}$), produjo una disminución de la efectividad del ajuste (mayores valores de $\sqrt{MS}$) para la combinación de datos de desarrollo más los datos de prueba. Los modelos ajustados al conjunto de datos de prueba arrojaron

Cuadro 6. Estadísticas de los residuales indicadores del ajuste de los modelos de altura-diámetro en parcelas permanentes de teca para los tres grupos de datos.

Modelo #	Datos de desarrollo			Datos de desarrollo + Datos de Prueba			Datos de prueba		
	$\bar{e}$	$\sqrt{v}$	$\sqrt{MS}$	$\bar{e}$	$\sqrt{v}$	$\sqrt{MS}$	$\bar{e}$	$\sqrt{v}$	$\sqrt{MS}$
1	0,00	0,17	0,17	0,22	0,15	0,27	0,43	0,23	0,49
2	0,00	0,17	0,17	0,21	0,15	0,25	0,40	0,23	0,46
3	0,04	0,17	0,18	0,26	0,15	0,30	0,47	0,23	0,52
4	0,04	0,17	0,18	0,25	0,15	0,29	0,44	0,23	0,50
5	0,00	0,17	0,17	0,21	0,15	0,26	0,42	0,23	0,48
6	0,00	0,17	0,17	0,21	0,15	0,25	0,40	0,23	0,46
7	0,00	0,17	0,17	0,21	0,14	0,25	0,39	0,23	0,45
8	0,01	0,17	0,17	0,19	0,14	0,23	0,38	0,22	0,44
9	0,00	0,17	0,17	0,20	0,14	0,25	0,38	0,23	0,45
10	0,00	0,17	0,17	0,20	0,14	0,25	0,38	0,23	0,44
11	0,10	0,17	0,20	0,31	0,14	0,34	0,52	0,13	0,53

el mayor sesgo de predicción ($\bar{e}$), menor precisión ($\sqrt{v}$) y menor efectividad de ajuste (mayor valor de $\sqrt{MS}$).

El análisis de los estadísticos de los residuales (Cuadro 6) mostró que los modelos de regresión polinómica presentaron la mejor efectividad de ajuste en la muestra, registrando los menores valores de $\bar{e}$, $\sqrt{v}$ y $\sqrt{MS}$. En orden de importancia con respecto al $\sqrt{MS}$ sobresalen los modelos 8, 10, 7 y 9. Los modelos 3 y 4 con transformación logarítmica de la variable dependiente y el modelo 11 arrojaron el mayor sesgo de predicción ($\bar{e}$). Esto puede deberse a que la distribución de los residuos se ve influenciada por cualquier transformación de la variable dependiente y puede conducir a la obtención de curvas sesgadas. El modelo de regresión no lineal 2 y el de regresión lineal 6, mostraron un comportamiento aceptable y similar en cuanto a la efectividad del ajuste. Los estadísticos de los residuos de los modelos para las parcelas 1, 2 y 3 consideradas individualmente (Cuadro 7) siguieron el mismo patrón de comportamiento. Estos resultados indican, que para las parcelas consideradas, los modelos polinómicos (7 a 10) parecen ajustarse mejor y presentar menos sesgo que modelos de regresión con mas simples o con variables transformadas, aunque se utilizen observaciones diferentes a áquellas utilizadas para construir el modelo. Esto pudiera ser un artificio debido a que ambos tipos de datos (desarrollo y prueba) provienen de la selección completamente aleatoria dentro de una muestra bastante homogénea. Cabe destacar que los modelos que utilizaron como variable independiente el recíproco de *d* presentan valores de los estadísticos de los residuales muy similares a los de los modelos polinómicos con la ventaja de que los primeros son mas simples. Los modelos que usaron transformaciones de la variable dependiente (3, 4 y 11) fueron los que mayor sesgo y variabilidad presentaron.

Extrapolación a valores de diámetros fuera del rango de las observaciones

Usualmente, las ecuaciones de relación altura-diámetro son ajustadas a partir de un conjunto de observaciones que abarca un intervalo limitado de diámetros. Sin embargo, estas ecuaciones son empleadas para estimar alturas de árboles cuyos diámetros están fuera de ese intervalo (extrapolación). Los resultados generados pueden utilizarse para estimar volúmenes de rodales, o para estimación de alturas y volúmenes en modelos de crecimiento. Para analizar el comportamiento de los modelos ajustados cuando el objetivo es la extrapolación, se calcularon los valores de altura estimados para diámetros entre 0 y 50 cm. Con fines de comparación los modelos se agruparon en modelos

Cuadro 7. Estadísticas de los residuales indicadores del ajuste de los modelos de altura-diámetro en parcelas permanentes de teca para las observaciones discriminadas por parcela.

Modelo #	Parcela 1			Parcela 2			Parcela 3		
	$\bar{e}$	$\sqrt{\bar{v}}$	$\sqrt{MS}$	$\bar{e}$	$\sqrt{\bar{v}}$	$\sqrt{MS}$	$\bar{e}$	$\sqrt{\bar{v}}$	$\sqrt{MS}$
1	0.20	0.28	**0.34**	0.40	0.29	**0.50**	0.08	0.25	**0.26**
2	0.16	0.28	**0.32**	0.38	0.28	**0.47**	0.10	0.24	**0.26**
3	0.23	0.27	**0.36**	0.45	0.29	**0.53**	0.11	0.25	**0.27**
4	0.20	0.28	**0.34**	0.42	0.28	**0.51**	0.14	0.24	**0.28**
5	0.18	0.28	**0.33**	0.39	0.29	**0.49**	0.08	0.24	**0.26**
6	0.15	0.28	**0.31**	0.37	0.28	**0.47**	0.11	0.24	**0.26**
7	0.16	0.27	**0.32**	0.37	0.28	**0.46**	0.10	0.24	**0.26**
8	0.14	0.26	**0.30**	0.35	0.28	**0.45**	0.07	0.24	**0.25**
9	0.16	0.27	**0.31**	0.36	0.28	**0.46**	0.10	0.24	**0.26**
10	0.18	0.27	**0.32**	0.36	0.28	**0.45**	0.09	0.24	**0.25**
11	0.27	0.27	**0.38**	0.47	0.28	**0.55**	0.19	0.24	**0.31**

no-lineales (1 y 2) y lineales con transformación de la variable independiente (5 y 6). Las curvas de estos dos grupos se presentan en la Figura 2. Asimismo, en la Figura 3 se muestran las curvas correspondientes a los modelos con transformación de la variable dependiente (3, 4 y 11). Finalmente en la Figura 4 se presentan las curvas de los modelos polinómicos (modelos 7 a 10).

En la Figura 2 se observa que los modelos 2 y 6 (no lineal y lineal considerando el recíproco del diámetro como variable independiente) se comportan asintóticamente para valores altos de diámetro. Por otro lado, el modelo 2 tiene un buen comportamiento para diámetros bajos, mientras que el modelo 6 puede presentar valores negativos por debajo de 10 cm de diámetro a la altura de pecho. Los modelos 1 y 5 por otro lado no se comportan asintóticamente, por lo que tenderían a producir sobreestimaciones de la altura, si la extrapolación se hace para diámetros grandes fuera del recorrido de las observaciones. En la Figura

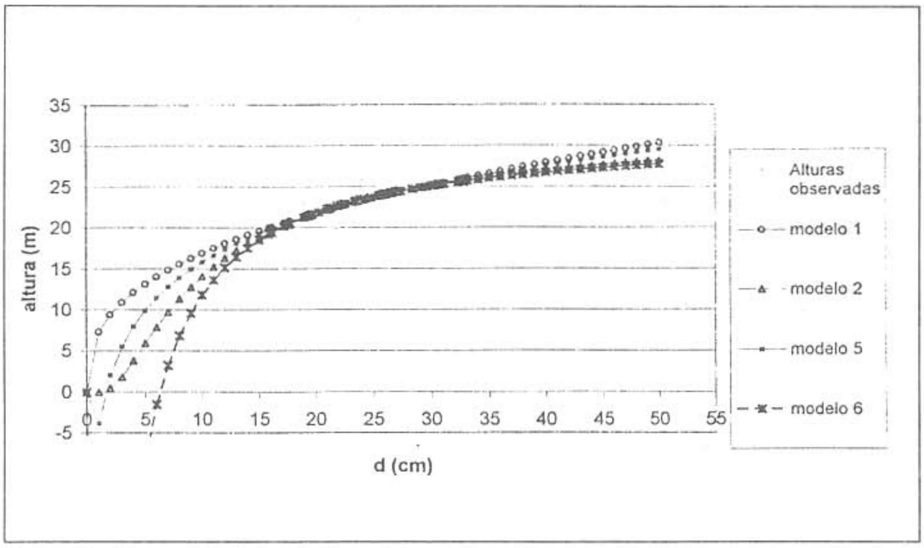

Figura 2. Curvas ajustadas de los modelos 1 y 2 (no lineales) y 5 y 6 (lineales) para diámetros entre 0 y 50 cm.

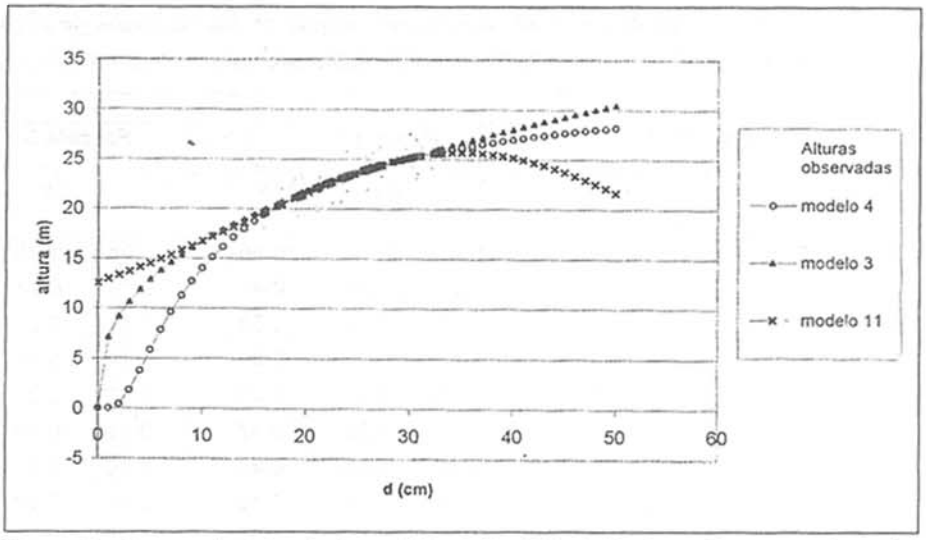

Figura 3. Curvas ajustadas de los modelos 4, 5 y 6 (variable dependiante transformada) para diámetros entre 0 y 50 cm.

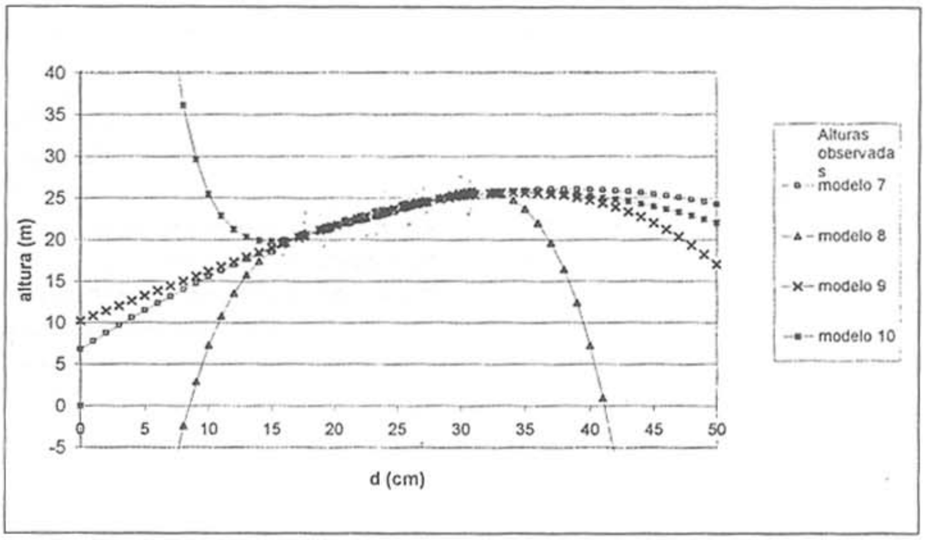

Figura 4. Curvas ajustadas de los modelos 7 a 11 (polinómicos) para diámetros entre 0 y 50 cm.

3, los modelos 3 y 4 presentan un comportamiento bastante aceptable para diámetros pequeños, pero para diámetros grandes podrían ocurrir sobreestimaciones. Por otro lado, el modelo 11 tiende a sobreestimar la altura cuando los diámetros son pequeños y produce un decrecimiento en las alturas para diámetros grandes. Recuérdese que estos modelos presentaron el mayor sesgo de predicción. Por último, en la Figura 4 se observa como el comportamiento de los modelos polinómicos puede producir resultados absurdos cuando se extrapola fuera del rango de diámetros observados (por ejemplo modelos 8 y 10).

Conclusiones y Recomendaciones

Se concluye que los estadísticos de los residuos, $\bar{e}$ (sesgo de predicción), ($\sqrt{\bar{v}}$) (sesgo de precisión) y $\sqrt{MS}$ (combinación de los dos anteriores), son criterios apropiados para comparar la efectividad del ajuste de modelos lineales, no lineales, transformados y no transformados, pero no son indicativos del posible comportamiento de los modelos cuando se quieren hacer estimaciones fuera del recorrido de las variables independientes.

Los estadísticos de los residuos permitieron detectar problemas de sesgamiento y precisión reducida en modelos donde se usaron tranformaciones de las variables dependiente e independientes.

El modelo 2 de regresión no lineal y el modelo 6 de regresión lineal tomando el inverso del diámetro como variable independiente parecen ser los más apropiados para el ajuste de curvas de relación altura-diámetro en estas parcelas. Ambos presentan un sesgo reducido y precisión alta, a la vez que presentan comportamientos aceptables cuando se extrapola a valores de diámetro fuera de los utilizados para el ajuste.

En el presente caso, no se observaron ventajas de los modelos no lineales sobre los modelos lineales. Si la efectividad de ambas clases de modelos es similar, entonces es preferible usar modelos lineales los cuales son mas fáciles de ajustar.

Los modelos polinómicos son eficientes si el objetivo es simplemente ajustar una curva a la nube de puntos observada. Pero, si el objetivo del ajuste de curvas de regresión altura diámetro es hacer predicciones y extrapolaciones (fuera del rango de los puntos usados en el ajuste), no es conveniente el uso de estos modelos, pues pueden producir estimaciones absurdas. En todo caso, se deben hacer estimaciones de altura utilizando valores extremos de las variables independientes o dibujar la curva ajustada para los valores de interés a fin de conocer el comportamiento de la curva de regresión.

El análisis de covarianza y la prueba de homogeneidad de varianzas puede usarse para determinar sí observaciones provenientes de diferentes parcelas o regímenes de espesura presentan una relación altura diámetro común. Esto permitiría combinar datos de diferentes parcelas para formar una muestra más grande, lo cual conduciria a la obtención de curvas mas precisas con un espectro mas amplio en su aplicación.

Referencias bibliográficas

ARABATZIS A. y BURKHART H. 1992. An evaluation of sampling methods and model forms estimating height-diameter relationships in Loblolly pine plantations. *Forest. Science* 38:192-198.

BARRET, J.P.1974. The coefficient of determination some limitations. *The American Statistician* 28:19-20.

BOCKING R. 1976. The analysis and selection of variables in linear regression. *Biometrics* 32:1-49.

CROCKER, D.C. 1972. Some interpretations of the multiple correlation coefficient. *The American Statistician* 26: 31-35.

CURTIS, R. O. 1967. Height - diameter and height - diameter - age equations for secon growth Douglas fir. *Forest. Science* 13: 365-375.

DÍAZ, A. 1989. Influencia de la Espesura en la Relación Altura-Diámetro de la teca en Caparo, Barinas, Venezuela. Proyecto CC2-7. Comodato ULA-MARNR. Mérida.

EK A.R. 1973. Performance of regression models for tree heiht estimation with small sample sizes. p. 67-80 In: Statistics in forestry research, Proc. 4 th Conf. Advisory Group of For. Stat., Vancouver. BC.

GUJARATI, K. 1978. Econometría Básica. Mc Graw-Hill.

HINKELMANN K. y O. KEMPTHORNE. 1994. Design and Analysis of Experiments. Vol. I. Introduction to Experimental Design. Jhon Wiley and Sons, INC. New York, pp.485.

MYERS, C. A. 1966.Height-diameter curves for tree species subject to stagnation. R. Mt. For. And Range Expt. Sta. U.S. For. S. Res Note RM-69.2 pp.

LITTLE, T.M. y F. JACKSON HILLS. 1987. Métodos Estadísticos para la Investigación en la Agricultura. Editorial Trillas, México.

OSORIO, O. 1997. Regímenes de espesura y sus efectos sobre la rentabilidad en plantaciones de teca (*Tectona grandis* L.f) en Caparo, Venezuela. Universidad de Los Andes, Facultad de Ciencias Forestales y Ambientales, Centro de Estudios Forestales y Ambientales de Postgrado. Mérida, Venezuela. Tesis M.Sc.

QUINTERO, J. 1995. Comparación de regímenes de espesura en plantaciones de teca (*Tectona grandis* L.) en la Unidad Experimental de la Reserva Forestal de Caparo. Universidad de Los Andes, Facultad de Ciencias Forestales, Centro de Estudios Forestales de Postgrado, Mérida, Venezuela. Tesis M.Sc.

STRAND, L. 1959. The accuracy of some methods for estimating volume and increment sample plots. Medd. Norske Skogforse 15(4): 284-392.

ZERPA F. 1989. Sistema de Información sobre rendimiento de plantaciones forestales en la Unidad Experimental de la Reserva Forestal de Caparo (SINFOPLAN). Tesis de Grado, Universidad de Los Andes, Facultad de Ingeniería, Escuela de Sistemas, Mérida, Venezuela.