

## Compatible dominant height - site index model for juniper (*Juniperus deppeana* Steud.)

Modelo compatible de altura dominante - índice de sitio para táscate (*Juniperus deppeana* Steud.)

Antonio Rodríguez-Carrillo<sup>\*</sup>; Francisco Cruz-Cobos; Benedicto Vargas-Larreta; Francisco J. Hernández.

Instituto Tecnológico de El Salto. Mesa del Tecnológico s/n. C. P. 34942. El Salto, Pueblo Nuevo, Durango, MÉXICO. Correo-e: ebrat\_3@hotmail.com, tel.: +52 (675) 87 104 52 (\*Autor para correspondencia).

### Abstract

The aim of this study was to determine the site quality of juniper (*Juniperus deppeana* Steud.) in the San Dimas region of the state of Durango, Mexico, using the site index method. The database comes from stem analysis of 43 trees felled in harvesting activities. The Chapman-Richards and Schumacher models, by means of the algebraic difference and generalized algebraic difference approaches, were tested to determine the site index; in addition, the error structure was modeled with a second-order autoregressive model to remedy the dependency of existing longitudinal errors. The results showed that the Chapman-Richards model in generalized algebraic difference form provided the best fit according to the adjusted coefficient of determination ( $R^2_{adj} = 0.98$ ) and root mean square error (RMSE = 0.46 m). Plotting of the quality curves generated with this model, superimposed on the observed heights, corroborated the goodness of fit of the model selected. The equation obtained with the generalized algebraic difference approach directly estimates the dominant height and site index at any height and base age.

**Keywords:** Algebraic difference approach, generalized algebraic difference approach, Chapman-Richards.

### Resumen

El objetivo de este trabajo fue determinar la calidad de estación de táscate (*Juniperus deppeana* Steud.), a través del método de índice de sitio, en la región de San Dimas, Durango. La base de datos proviene de 43 análisis troncales de árboles derribados en actividades de aprovechamiento. Los modelos de Chapman-Richards y de Schumacher, mediante los procedimientos de diferencias algebraicas y diferencias algebraicas generalizadas, se probaron para determinar el índice de sitio; además, se modeló la estructura del error con un modelo autorregresivo de segundo orden para corregir la dependencia de errores longitudinales presentes. Los resultados mostraron que el modelo de Chapman-Richards en diferencias algebraicas generalizadas presentó el mejor ajuste de acuerdo con el coeficiente de determinación ajustado ( $R^2_{adj} = 0.98$ ) y la raíz del cuadrado medio del error (RCME = 0.46 m). La representación gráfica de las curvas de calidad generadas con este modelo, sobrepuestas a las alturas observadas, corroboraron la bondad de ajuste del modelo seleccionado. La ecuación obtenida con la metodología de diferencias algebraicas generalizadas estima directamente la altura dominante y el índice de sitio a cualquier altura y edad de referencia.

**Palabras clave:** Diferencias algebraicas, diferencias algebraicas generalizadas, Chapman-Richards.

## Introduction

Evaluation of site quality, estimated through productivity indicators, is an important element in the planning and implementation of sustainable forest management programs. In this regard, several indicators have been tested in different forest types, but the most accepted for specific and homogeneous forests has been the site index, defined as the height growth at a base age (Avery & Burkhart, 2002; Payandeh & Wang, 1994; Skovsgaard & Vanclay, 2008).

Various methods have been developed for constructing site index functions; however, the most commonly used ones are the guide curve method and the algebraic difference approach. In the first, the original model is fitted to estimate the parameters applied in the functional equations to generate families of anamorphic or polymorphic curves. This type of model is the most common way to model the height-age relationship in unispecific and homogeneous stands through static equations (Cieszewski, 2001). The algebraic difference method involves fitting the parameters of the functional equation and inevitably requires periodic measurements (Torres-Rojo, 2001), forming dynamic equations (Cieszewski, 2001). Dynamic site index equations have the property of base-age invariance, so their selection has no effect on the prediction. Bailey and Clutter (1974) presented the Algebraic Difference Approach (ADA), expressing the independent variable as a function of age and the model's parameters ( $X = f[t, B]$ ). This methodology is based on the assumption that the parameters of the base equation are dependent on site productivity, thereby generating families of anamorphic or polymorphic growth curves, depending on the parameter isolated from the equation. Cieszewski and Bailey (2000) proposed the Generalized Algebraic Difference Approach (GADA) from the ADA methodology. The major advantage of this method is that dynamic equations can be obtained from a base growth model, ensuring that more than one parameter depends on the specific conditions of the forest site (Cieszewski, 2001, 2002, 2003; Cieszewski & Bailey, 2000). The parameters must be expressed as functions of the site defined by a variable  $X$  (unobservable independent variable) that describes site productivity as a result of a combination of factors such as management regimes, soil conditions and ecological and climatic factors, thereby obtaining families of more flexible curves. With this generalization, families that are both polymorphic and with multiple asymptotes are obtained (Cieszewski, 2002). Selecting any of the methods described above has generated much controversy; however, in recent years greater importance has been given to developing dynamic site index functions with polymorphic curves dependent on a parameter, since it is believed that the trends of the curves model the height growth of trees with

## Introducción

La evaluación de la calidad de sitio, estimada a través de indicadores de productividad, es un elemento importante en la planeación e implementación de programas de manejo forestal sustentable (sostenido). En este sentido, diversos indicadores se han probado en diferentes tipos de bosques, pero el más aceptado para bosques específicos y homogéneos ha sido el índice de sitio, definido como el crecimiento en altura a una edad de referencia (Avery & Burkhart, 2002; Payandeh & Wang, 1994; Skovsgaard & Vanclay, 2008).

Diversos métodos se han desarrollado para la construcción de funciones de índices de sitio; sin embargo, los más comúnmente utilizados son los conocidos como método de la curva guía y método de la diferencia algebraica. En el primero, el modelo original se ajusta para estimar los parámetros que se aplican en las ecuaciones funcionales para generar familias de curvas de tipo anamórfico o polimórfico. Este tipo de modelo es la forma más común para modelar la relación altura-edad en rodales uniespecíficos y homogéneos a través de ecuaciones estáticas (Cieszewski, 2001). El método de la diferencia algebraica implica ajustar los parámetros de la ecuación funcional y requiere forzosamente de mediciones periódicas (Torres-Rojo, 2001), formando ecuaciones dinámicas (Cieszewski, 2001). Las ecuaciones dinámicas de índice de sitio tienen la propiedad de que la edad base es invariante, por lo que su selección no tiene ningún efecto sobre la predicción. Bailey y Clutter (1974) presentaron el procedimiento de la diferencia algebraica (Algebraic Difference Approach [ADA]), expresando la variable independiente en función de la edad y de los parámetros del modelo ( $X = f[t, B]$ ). Esta metodología se basa en el supuesto de que los parámetros de la ecuación base son dependientes de la productividad del sitio, generando con ello familias de curvas de crecimiento anamórficas o polimórficas, dependiendo del parámetro que se aísla de la ecuación. Cieszewski y Bailey (2000) propusieron el procedimiento de la diferencia algebraica generalizada (Generalized Algebraic Difference Approach [GADA]), a partir de la metodología ADA. Este método tiene la principal ventaja de que se pueden obtener ecuaciones dinámicas a partir de un modelo de crecimiento base permitiendo que más de un parámetro dependa de las condiciones específicas de la estación forestal (Cieszewski, 2001, 2002, 2003; Cieszewski & Bailey, 2000). Los parámetros deben expresarse como funciones de la estación definida por una variable  $X$  (variable no observable e independiente) que describe la productividad del sitio como resultado de la suma de factores tales como regímenes de manejo, condiciones de suelo y factores ecológicos y climáticos, obteniendo así familias de curvas más flexibles. Con esta generalización se obtienen familias que sean a la vez polimórficas y con asíntotas múltiples (Cieszewski, 2002). La selección de

greater accuracy (Goelz & Burk, 1992; Ker & Bowling, 1991; Payandeh & Wang, 1994).

In the San Dimas region of Durango, *Juniperus deppeana* Steud. accounts for about 3.55 % of the authorized volume (Soto, 2009). It is widely used in the furniture and handicraft industries, and in rural areas it is used for poles and as fuel for domestic use. Since this forest region does not have scientifically-validated site index equations, the aim of this work was to construct site quality curves of *J. deppeana* using the generalized algebraic difference approach.

## Materials and methods

### Study area

The work was conducted in the San Dimas region comprising the municipality of the same name and a portion of the municipalities of Tamazula, Canatlán, Otáez and Durango. The study area covers about 586,000 ha in which *J. deppeana* is distributed in a scattered manner. The central coordinates of the area are 24° 25' 27" NL and 105° 35' 33" WL; the predominant altitude ranges between 2,000 and 2,800 m. The region is located in three hydrologic regions: Sinaloa, Presidio-San Pedro and Nazas Aguanaval.

### Study data

A total of 273 data pairs from stem analysis of 43 dominant trees were used. The trees were selected so that they would represent all topographic, edaphic and heights above sea level conditions of the study area. Each tree was felled at the minimum possible stump height, extracting three sections from the base of the tree to diameter at breast height (1.30 m); later sections varying in length between 1.0 and 2.54 m were obtained until reaching the top of the tree. A slice was obtained from each section in which the trunk was divided, in order to count and measure the growth rings. Following the field portion of the study, the age of each section was determined and true heights were estimated with the Carmean (1972) method as modified by Newberry (1991).

### Models to determine the *J. deppeana* site index

In the present study we fitted two models that, due to their flexibility, have been widely used in other conifers; however, this work represents the first such analysis of *J. deppeana*. The dominant height-age relationship was described using models based on those of Schumacher (1939) and Chapman-Richards (Richards, 1959) that have proved to be efficient for determining the site indexes of various conifer species (Castillo et al., 2013; Diéguez, Álvarez, Barrio, & Rojo, 2005; Vargas-Larreta, Álvarez-González, Corral-Rivas, & Aguirre-Calderón,

cualquiera de los métodos descritos anteriormente ha generado gran polémica; sin embargo, en los últimos años se le ha dado mayor importancia al desarrollo de funciones de índices de sitio dinámicas con curvas polimórficas dependientes de un parámetro, pues se considera que las tendencias de las curvas modelan con mayor precisión el crecimiento en altura de los árboles (Goelz & Burk, 1992; Ker & Bowling, 1991; Payandeh & Wang, 1994).

En la región de San Dimas, Durango, *Juniperus deppeana* Steud. ocupa aproximadamente 3.55 % del volumen autorizado (Soto, 2009). Su uso es amplio en la industria mueblera y artesanal, y en el medio rural se emplea para postes y como combustible para uso doméstico. Dado que en esta región forestal no se cuenta con ecuaciones de índice de sitio validadas científicamente, el objetivo de este trabajo fue construir curvas de calidad de estación de *J. deppeana* mediante el método de diferencias algebraicas generalizado.

## Materiales y métodos

### Área de estudio

El trabajo se realizó en la región de San Dimas que comprende el municipio del mismo nombre y una porción de los municipios de Tamazula, Canatlán, Otáez y Durango. El área de estudio abarca aproximadamente 586,000 ha en las que *J. deppeana* está distribuida de manera dispersa. Las coordenadas centrales del área son 24° 25' 27" LN y 105° 35' 33" LO; la altitud predominante oscila entre 2,000 y 2,800 m. La región se ubica en tres regiones hidrológicas: Sinaloa, Presidio-San Pedro y Nazas Aguanaval.

### Datos de estudio

Se utilizaron 273 pares de datos provenientes de 43 análisis troncales de árboles dominantes. Los árboles se seleccionaron de tal manera que representaran todas las condiciones topográficas, edáficas y alturas sobre el nivel del mar del área estudiada. Cada árbol se derribó a la altura mínima posible del tocón extrayendo tres secciones desde la base del árbol hasta el diámetro normal (1.30 m); posteriormente se obtuvieron secciones de longitud variable entre 1.0 y 2.54 m hasta llegar a la punta del árbol. De cada sección en que el tronco fue dividido se obtuvo una rodaja, para contar y medir los anillos de crecimiento. La edad de cada sección se determinó en gabinete y las alturas verdaderas se estimaron con el método de Carmean (1972) modificado por Newberry (1991).

### Modelos para determinar el índice de sitio de *J. deppeana*

En el presente estudio se ajustaron dos modelos que por su flexibilidad han sido utilizados ampliamente en

**Table 1. Base models and their ADA and GADA expressions, to estimate the site index of *Juniperus deppeana***  
**Cuadro 1. Modelos base y sus expresiones en diferencias algebraicas (ADA) y diferencias algebraicas generalizadas (GADA), para estimar el índice de sitio de *Juniperus deppeana*.**

Name / Nombre	Base model / Modelo base	ADA expression / Expresión ADA	GADA expression / Expresión GADA
Schumacher	$H = a_1 e^{a_2(t/t_1)}$	$H_2 = H_1 e^{a_2(t/t_1)}$ Anamorphic (1) / Anamórfico (1)	
		$H_2 = a_1 \left( \frac{H_1}{a_1} \right)^{(t/t_1)}$ Polymorphic (2) / Polimórfico (2)	
Chapman-Richards	$H = a_1 (1 - e^{-a_2 t})^{a_3}$	$H_2 = H_1 \left( \frac{1 - e^{-a_2 t_2}}{1 - e^{-a_2 t_1}} \right)^{a_3}$ Anamorphic (3) / Anamórfico (3)	$H = H_0 \left[ \frac{1 - e^{-b_1 t_1}}{1 - e^{-b_1 t_0}} \right]^{b_2 + b_3 / X_0}$ (6)
		$H_2 = a_1 (1 - [1 - (H_1/a_1)^{1/a_3}]^{t_2/t_1})^{a_3}$ Polymorphic I (4) / Polimórfico I (4)	
		$H_2 = a_1 \left( \frac{H_1}{a_1} \right)^{\left( \frac{\ln(1 - e^{-a_2 t_2})}{\ln(1 - e^{-a_2 t_1})} \right)}$ Polymorphic II (5) / Polimórfico II (5)	

$H$  = Dominant height (m),  $t$  = Base age (years),  $a_i$  = Parameters within the base equations,  $b_i$  = Global parameters for the GADA formulations,  $X_0$  = Site-dependent parameter.

$H$  = Altura dominante (m),  $t$  = Edad de referencia o edad base (años),  $a_i$  = Parámetros dentro de las ecuaciones base,  $b_i$  = Parámetros globales para las formulaciones GADA,  $X_0$  = Parámetro dependiente de la estación.

2010). Table 1 presents these models tested in their ADA and GADA expressions, from the base equations.

### Analysis and comparison of models

The best model was selected on the basis of fit statistics in terms of accuracy, bias and graphical comparisons. The statistics calculated were: the adjusted coefficient of determination ( $R^2_{adj}$ ) that reflects the total variability explained by the model, considering all the parameters to be estimated, the root mean square error (RMSE) that analyzes the accuracy of the estimates, and the mean bias ( $\bar{E}$ ) that evaluates the deviation of the model with respect to the observed values. Although there are debates regarding the use of  $R^2$  in nonlinear regression, the general usefulness of an overall measure of the predictive ability of the model appears to override some of these limitations (Ryan, 1997). The expression of these statistics is as follows:

$$R^2_{adj} = 1 - \left[ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \right] * \left( \frac{n-1}{n-p} \right)$$

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

$$\bar{E} = \left[ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \right]$$

otras coníferas; no obstante, este trabajo representa la primera aproximación de análisis de *J. deppeana*. La relación altura dominante-edad se describió empleando los modelos base de Schumacher (1939) y el de Chapman-Richards (Richards, 1959) que han demostrado ser eficientes para la determinación de índices de sitio de diversas especies de coníferas (Castillo et al., 2013; Diéguez, Álvarez, Barrio, & Rojo, 2005; Vargas-Larreta, Álvarez-González, Corral-Rivas, & Aguirre-Calderón, 2010). El Cuadro 1 presenta dichos modelos probados en sus expresiones ADA y GADA, a partir de las ecuaciones base.

### Análisis y comparación de modelos

El mejor modelo se seleccionó en función de estadísticos de ajuste en términos de precisión, sesgo y comparaciones gráficas. El coeficiente de determinación ajustado ( $R^2_{adj}$ ) que refleja la variabilidad total explicada por el modelo, considerando todos los parámetros a estimar; la raíz del cuadrado medio del error (RCME) que analiza la precisión de las estimaciones; y el sesgo promedio ( $\bar{E}$ ) que evalúa la desviación del modelo respecto de los valores observados, fueron los estadísticos calculados. Aunque existen discusiones respecto al uso de la  $R^2$  en regresión no lineal, la utilidad general de una medida global de la capacidad predictiva del modelo parece anular algunas de esas limitaciones (Ryan, 1997). La expresión de dichos estadísticos es la siguiente:

Where:

$y_i$  = Observed value

$\hat{y}_i$  = Predicted value

$\bar{y}$  = Average height value

$n$  = Number of observations

$p$  = Number of model parameters

In addition to the statistics described, the fit of the model-generated curves superimposed with the observed values, the residues versus the predicted values, and the graphs of the root mean square error and mean bias were analyzed.

### Correcting for the autocorrelation of longitudinal errors

The possible autocorrelation of the longitudinal data was corrected using a continuous autoregressive (CAR [x]) model, which can even be used with data obtained at intervals not constant in time or space and with a different number of measurements per sampling unit (Zimmerman & Núñez, 2001). In the second-order continuous autoregressive structure (CAR [2]), the error term is expanded as:

$$e_{ij} = d_1 \rho_1^{t_{ij}-t_{ij-1}} e_{ij-1} + d_2 \rho_2^{t_{ij}-t_{ij-2}} e_{ij-2} + \varepsilon_{ij}$$

Where:

$e_{ij}$  =  $j$ -th residue of the sampling unit  $i$

$d_k = 1$  for  $j > k$  and  $0$  for  $j = k$  ( $k = 1, 2$ )

$\rho_k$  =  $k$ -order autoregressive parameter to estimate

$t_{ij} - t_{ij-k}$  = Time or distance separating the measurement  $j$ -th from the measurement  $j$ -th- $k$  in each sampling unit  $t_{ij} > t_{ij-k}$  and is the random error.

For the site index curves, the autocorrelation is spatial ( $t_{ij}$  is the age of the measurement  $j$  in the tree  $i$ ).

### Methodology for fitting the site index equations

The fitting technique for estimating the parameters was a base-age invariant method called the nested iterative procedure (Tait, Cieszewski, & Bella, 1988), which was performed in accordance with the procedure of Vargas-Larreta et al. (2010). The method estimates the specific effects of the site and assumes that there are always random and measurement errors in the data, which must be modeled (Cieszewski, 2003). The simultaneous fitting of the growth equation and the error structure (given by the autoregressive model) was made with the MODEL procedure of the statistical package SAS/ETS™ (Statistical Analysis System [SAS], 2004), allowing a dynamic updating of the residues (Vargas-Larreta et al., 2010).

$$R^2_{adj} = 1 - \left[ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] * \left( \frac{n-1}{n-p} \right)$$

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

$$\bar{E} = \left[ \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \right]$$

Donde:

$y_i$  = Valor observado

$\hat{y}_i$  = Valor predicho

$\bar{y}$  = Valor promedio de la altura

$n$  = Número de observaciones

$p$  = Número de parámetros del modelo

Además de los estadísticos descritos, se analizaron el ajuste de las curvas generadas con el modelo sobrepuestos con los valores observados, los residuos frente a los valores predichos, y los gráficos de la raíz del cuadrado medio del error y del sesgo promedio.

### Corrección de la autocorrelación de los errores longitudinales

La posible autocorrelación de los datos longitudinales se corrigió utilizando un modelo que representa una estructura autorregresiva (en inglés, continuous autoregressive model CAR [x]), que incluso puede ser utilizado con datos obtenidos a intervalos no constantes en el tiempo o en el espacio y con un número diferente de mediciones en cada unidad muestral (Zimmerman & Núñez, 2001). En la estructura autorregresiva de segundo orden (CAR [2]), el término del error se expande como:

$$e_{ij} = d_1 \rho_1^{t_{ij}-t_{ij-1}} e_{ij-1} + d_2 \rho_2^{t_{ij}-t_{ij-2}} e_{ij-2} + \varepsilon_{ij}$$

Donde:

$e_{ij}$  =  $j$ -ésimo residuo de la unidad muestral  $i$

$d_k = 1$  para  $j > k$  y  $0$  para  $j = k$  ( $k = 1, 2$ )

$\rho_k$  = Parámetro autorregresivo de orden  $k$  a estimar

$t_{ij} - t_{ij-k}$  = Tiempo o distancia que separa la medición  $j$ -ésima de la medición  $j$ -ésima- $k$  en cada unidad muestral  $t_{ij} > t_{ij-k}$  y es el error aleatorio.

En el caso de las curvas de índice de sitio, la autocorrelación es espacial ( $t_{ij}$  es la edad de la medición  $j$  en el árbol  $i$ ).



## Results and discussion

### Fitting equations using the algebraic difference approach

The mean errors (RMSE) obtained with the algebraic difference method ranged between 0.66 and 1.06 m. Table 2 presents the parameters estimated for the Schumacher and Chapman-Richards models. In this table it can be seen that the Chapman-Richards model presented the best values, showing a better fit in the polymorphic II modality (equation 5) and explaining more than 96 % of the variance in dominant height growth as a function of age. The parameters were highly significant for all equations ( $P < 0.01$ ).

### Metodología de ajuste de las ecuaciones de índice de sitio

La técnica de ajuste para la estimación de los parámetros fue un método invariante con respecto a la edad de referencia (base-age invariance), denominado método iterativo (nested iterative procedure) (Tait, Cieszewski, & Bella, 1988), que se hizo de acuerdo con el procedimiento de Vargas-Larreta et al. (2010). El método estima los efectos específicos del sitio y asume que siempre existen errores aleatorios y de medición en los datos, los cuales deben ser modelados (Cieszewski, 2003). El ajuste simultáneo de la ecuación de crecimiento y de la estructura del error (dada por el modelo autorregresivo) se hizo con el

**Table 2. Estimated parameters for the Schumacher and Chapman-Richards models.**  
**Cuadro 2. Parámetros estimados para los modelos de Schumacher y Chapman-Richard.**

Number/ Número	Model/Modelo	Parameter/ Parámetro	Estimated/ Estimado	Standard error/ Error estándar	Value-p/ Valor-p	RMSE (m)/ RCME (m)	R <sup>2</sup> <sub>adj</sub>
1	$H_2 = H_1 e^{a_2 \left( \frac{1}{t_2} - \frac{1}{t_1} \right)}$	$a_2$	-11.323	0.360	<.0001	1.06	0.91
2	$H_2 = a_1 \left( \frac{H_1}{a_1} \right)^{\left( \frac{t_1}{t_2} \right)}$	$a_1$	11.392	0.328	<.0001	0.97	0.93
3	$H_2 = H_1 \left( \frac{1 - e^{-a_2 t_2}}{1 - e^{-a_2 t_1}} \right)^{a_3}$	$a_2$	-0.034	0.002	<.0001	0.73	0.96
		$a_3$	1.278	0.049	<.0001		
4	$H_2 = a_1 (1 - [1 - (H_1/a_1)^{1/a_3}]^{t_2/t_1})^{a_3}$	$a_1$	14.101	0.441	<.0001	0.70	0.96
		$a_3$	1.160	0.041	<.0001		
5	$H_2 = a_1 \left( \frac{H_1}{a_1} \right)^{\left( \frac{\ln(1 - e^{-a_2 t_2})}{\ln(1 - e^{-a_2 t_1})} \right)}$	$a_1$	15.153	0.703	<.0001	0.66	0.96
		$a_2$	0.018	0.001	<.0001		
		$b_1$	0.033	0.002	<.0001		
		$b_2$	1.968	0.286	<.0001		
6	$H = H_0 \left[ \frac{1 - e^{-b_1 t_1}}{1 - e^{-b_1 t_0}} \right]^{b_2 + b_3 X_0}$	$b_3$	-1.553	0.640	0.0158	0.46	0.98
		$\rho_1$	0.882	0.039	<.0001		
		$\rho_2$	0.772	0.052	<.0001		

$H$  = Dominant height (m),  $t$  = Base age (years),  $a_i$  = Parameters within the base equations,  $b_i$  = Global parameters for the GADA formulations,  $\rho_k$  = k-order autoregressive parameter to estimate,  $X_0$  = Site-dependent parameter. RMSE: Root mean square error. R<sup>2</sup><sub>adj</sub> = Adjusted coefficient of determination.

$H$  = Altura dominante (m),  $t$  = Edad de referencia o edad base (años),  $a_i$  = Parámetros dentro de las ecuaciones base,  $b_i$  = Parámetros globales para las formulaciones de diferencias algebraicas generalizadas,  $\rho_k$  = Parámetro autorregresivo de orden  $k$  a estimar,  $X_0$  = Parámetro dependiente de la estación. RCME: Raíz del cuadrado medio del error. R<sup>2</sup><sub>adj</sub> = Coeficiente de determinación ajustado.

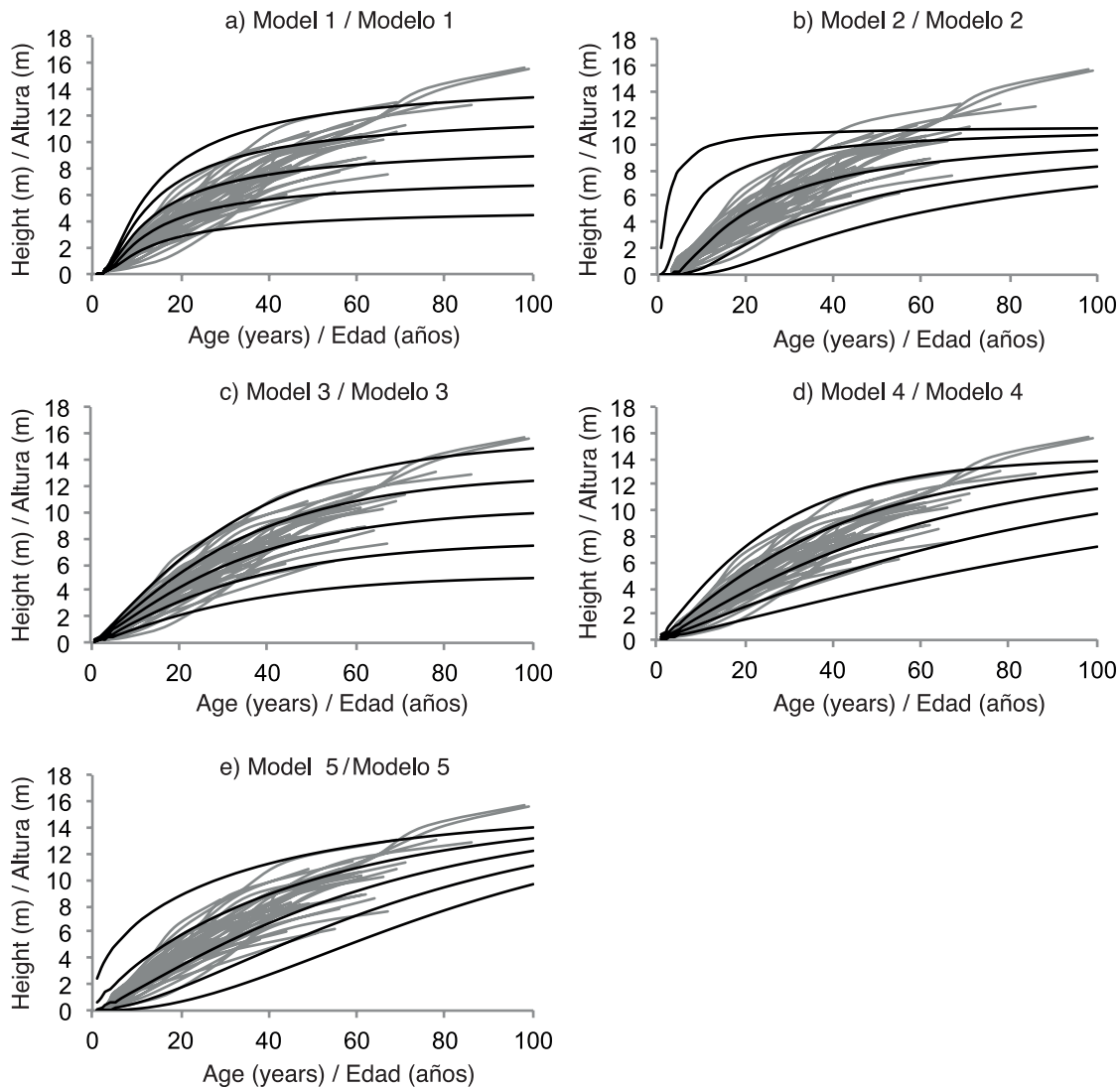
Differences in the goodness of fit statistics of the equations were minimal, so the graphical analysis was useful for selecting the best model, by overlapping the curves generated onto the observed data, and thus comparing the ability of each model to describe the dominant height-age relationship. Figure 1 shows the graphical analysis of the site index curves, which allowed comparing the individual height growth trends of *J. deppeana*. Figure 1 shows that equations 1 and 2, anamorphic and polymorphic respectively, derived from the Schumacher model, present bad fits especially at the sites of greatest productivity. On the

procedimiento MODEL del paquete estadístico SAS/ETS™ (Statistical Analysis System [SAS], 2004), permitiendo una actualización dinámica de los residuos (Vargas-Larreta et al., 2010).

## Resultados y discusión

### Ajuste de las ecuaciones por el método de la diferencia algebraica

Los errores medios (REMC) obtenidos con el método de la diferencia algebraica oscilaron entre 0.66 y 1.06 m.



**Figure 1. Family of site index curves (SI 4, 6, 8, 10 and 12 m) of *Juniperus deppeana*, generated with models by the algebraic difference approach (see Table 1) at a base age of 50 years old.**

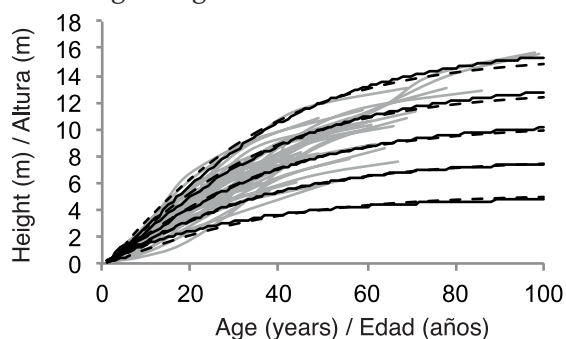
**Figura 1. Familia de curvas de índice de sitio (IS 4, 6, 8, 10 y 12 m) de *Juniperus deppeana*, generadas con los modelos por el método de diferencias algebraicas (ver Cuadro 1) a una edad de referencia de 50 años.**

other hand, equations 3, 4 and 5, derived from the Chapman-Richards models, in general, best describe the dominant height growth trend with a marked improvement in the anamorphic model (equation 3). The curves generated with equation 3 have more plausible asymptotic values than those generated with the rest. Although the two base models are biologically realistic because they consider a point of maximum growth (asymptote) and a growth rate, the Chapman-Richards model also considers an inflection point on the growth curve, making it more flexible (Alder, 1980).

### Fitting the site index equation using the generalized algebraic difference approach

Equation 5 of Chapman-Richards (Table 1) was selected as the best model to describe the dominant height growth trend of *J. deppeana*, so we proceeded to fit it with the generalized algebraic difference approach, integrating a second-order autoregressive model to correct potential autocorrelation in the errors. The model explains over 98 % of the variance in dominant height growth with a mean error of 0.46 m, which is slightly lower than the values obtained by the algebraic difference method. The Durbin-Watson statistic had a value of 1.86 (very close to 2), thereby demonstrating correction of residue dependency.

Figure 2 shows the curves obtained for site indices 4, 6, 8, 10 and 12 m at the age of 50 superimposed onto the observed data that were used in the fit. The graph shows that the curves follow the data trajectories over the entire age range and describe dominant height



**Figure 2.** Site index curves (SI 4, 6, 8, 10 and 12 m) of *Juniperus deppeana*, generated with model 3 (see Table 1) by the algebraic difference approach (dotted line) and model 6 (see Table 1) by the generalized algebraic difference approach (solid line), at a base age of 50 years.

**Figura 2.** Curvas de índice de sitio (IS 4, 6, 8, 10 y 12 m) de *Juniperus deppeana*, generadas con el modelo 3 (ver Cuadro 1) por el método de diferencias algebraicas (línea punteada) y modelo 6 (ver Cuadro 1) por el método de diferencias algebraicas generalizadas (línea continua), a una edad referencial de 50 años.

El Cuadro 2 presenta los parámetros estimados para los modelos de Schumacher y Chapman-Richards. En dicho cuadro se puede observar que el modelo de Chapman-Richards presentó los mejores valores, mostrando mejor ajuste en la modalidad polimórfico II (ecuación 5) y explicando más de 96 % de la varianza del crecimiento en altura dominante en función de la edad. Los parámetros fueron altamente significativos para todas las ecuaciones ( $P < 0.01$ ).

Las diferencias en los estadísticos de bondad de ajuste de las ecuaciones fueron mínimas, por lo que el análisis gráfico fue útil para la selección del mejor modelo, sobreponiendo las curvas generadas sobre los datos observados, y así comparar la capacidad de cada modelo para describir la relación altura dominante-edad. La Figura 1 muestra el análisis gráfico de las curvas de índice de sitio, el cual permitió comparar las tendencias individuales del crecimiento en altura de *J. deppeana*. En la Figura 1 se observa que las ecuaciones 1 y 2, anamórfica y polimórfica, respectivamente, derivadas del modelo Schumacher, presentan desajustes sobre todo en los sitios de mayor productividad. Por otra parte, las ecuaciones 3, 4 y 5, derivadas del modelo de Chapman-Richards, en lo general, describen mejor la tendencia del crecimiento en altura dominante con notoria mejoría en el modelo anamórfico (ecuación 3). Las curvas generadas con la ecuación 3 presentan valores asintóticos más plausibles que las generadas con el resto. Aunque los dos modelos base son biológicamente realistas, ya que consideran un punto de máximo crecimiento (asíntota) y una tasa de crecimiento, el modelo de Chapman Richards considera además un punto de inflexión de la curva de crecimiento, lo que lo hace más flexible (Alder, 1980).

### Ajuste de la ecuación de índice de sitio por el método de la diferencia algebraica generalizada

La ecuación 5 de Chapman-Richards (Cuadro 1) fue seleccionada como el mejor modelo para describir la tendencia del crecimiento en altura dominante de *J. deppeana*, por lo que se procedió a ajustarlo con el método de la diferencia algebraica generalizada, integrando un modelo autorregresivo de segundo orden para corregir la autocorrelación potencial de los errores. El modelo explica más de 98 % de la varianza del crecimiento en altura dominante con error medio de 0.46 m, el cual es ligeramente menor que los valores obtenidos por el método de la diferencia algebraica. El estadístico Durbin-Watson tuvo valor de 1.86 (muy cercano a 2), con lo que se demuestra corrección de la dependencia de residuos.

En la Figura 2 se muestran las curvas obtenidas para los índices de sitio 4, 6, 8, 10 y 12 m a la edad de 50 años superpuestas sobre los datos observados que se



growth plausibly for ages over 30 years. The ADA (Table 1, equation 5) and GADA (Table 1, equation 8) equations generate similar curves, as shown in Figure 2; however, the ADA method has better fits at early ages than the GADA method. Despite this, use of the generalized algebraic difference approach is recommended since a significant gain is obtained in the goodness of fit statistics, as well as more accurate parameters.

It is important to note that, from a practical standpoint, parameters  $\rho_1$  and  $\rho_2$ , used to model the error structure, are generally not used (Castillo et al., 2013; Diéguez-Aranda, Burkhart, & Amateis, 2006). The main purpose of modeling the error structure is to obtain consistent estimates of the parameters and their standard errors, while the specific parameter of site quality estimated for each individual is removed in a way similar to that of the autocorrelation process (Cieszewski, 2001); thus, the estimated height as a function of age of a new individual is directly incorporated to make predictions.

### Residual analysis

Figure 3a shows the trend of the fitted model's residues without considering the autoregressive parameters. After fitting the second-order autoregressive model, the residual correlation indicator trend disappears (Figure 3c). Figure 4 shows that considering the autocorrelation correction provides the random pattern of the residues around the line 0 with homogeneous variance and without detecting a definite trend; these results agree with those obtained by Castillo et al. (2013), Vargas-Larreta et al. (2010) and Diéguez-Aranda et al. (2005).

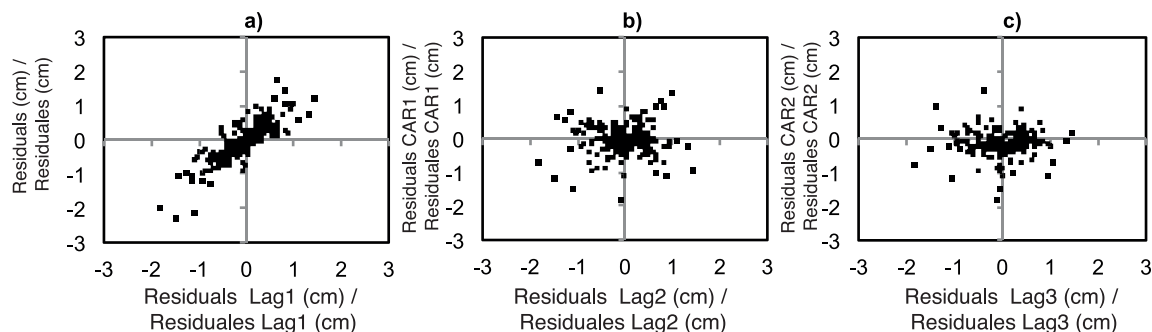
The evolution of the bias of equation 6 is analyzed in Figure 5a. This graph shows the distribution of the

emplearon en el ajuste. El gráfico muestra que las curvas siguen las trayectorias de los datos en todo el rango de edades y describen el crecimiento en altura dominante de manera plausible para edades mayores de 30 años. Las ecuaciones ADA (Cuadro 1, ecuación 5) y GADA (Cuadro 1, ecuación 8) generan curvas similares, tal como lo muestra la Figura 2; no obstante, el método ADA presenta mejores ajustes a edades tempranas que el método GADA. A pesar de lo anterior, se recomienda el uso del método de diferencias algebraicas generalizadas ya que se obtiene ganancia significativa en los estadísticos de bondad de ajuste, además de parámetros más precisos.

Es importante señalar que desde el punto de vista práctico, los parámetros  $\rho_1$  y  $\rho_2$ , utilizados para modelar la estructura del error, generalmente no se utilizan (Castillo et al., 2013; Diéguez-Aranda, Burkhart, & Amateis, 2006). El propósito principal de modelar la estructura del error es obtener estimaciones consistentes de los parámetros y de sus errores estándar, mientras que el parámetro específico de la calidad de estación estimado para cada individuo se suprime de forma similar como en el proceso de autocorrelación (Cieszewski, 2001); de esta forma, la altura estimada en función de la edad de un nuevo individuo se incorpora directamente para realizar predicciones.

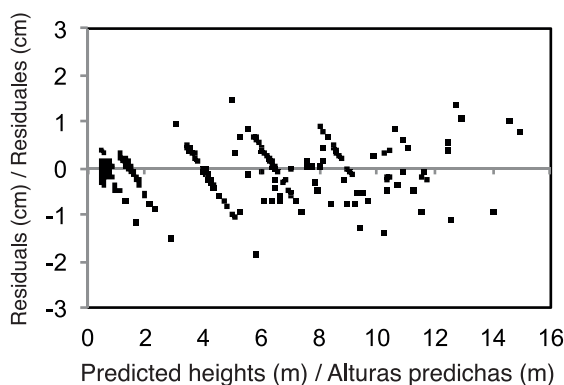
### Análisis de residuales

La Figura 3a muestra la tendencia de los residuos del modelo ajustado sin considerar los parámetros autorregresivos. Después de ajustar el modelo autorregresivo de segundo orden, la tendencia de indicios de correlación de los residuales desaparece (Figura 3c). La Figura 4 muestra que al considerar la



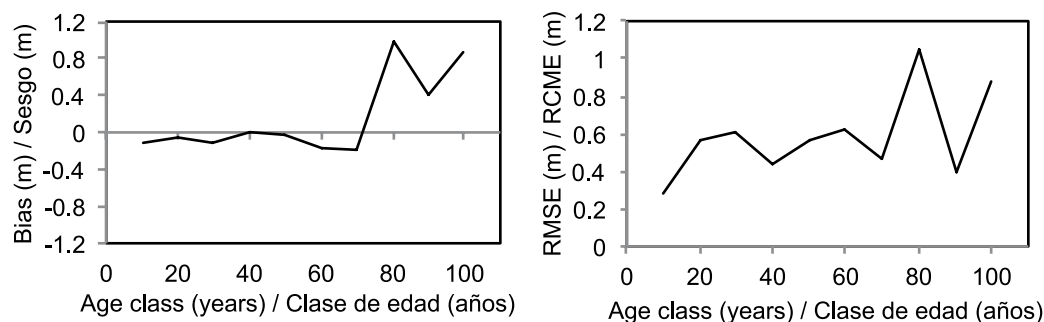
**Figure 3. Fitted residues of model 6 (Table 1): a) without considering the error structure, b) considering an order-1 continuous autoregressive model (CAR1) and through an order-2 continuous autoregressive model (CAR2).**

**Figura 3. Residuos ajustados del modelo 6 (Cuadro 1): a) sin considerar la estructura del error, b) considerando un modelo autorregresivo de orden 1 (CAR1) y a través de un modelo autorregresivo de orden 2 (CAR2).**



**Figure 4. Correlation between residues and predicted heights with model 6 (see Table 1), considering the error structure.**

**Figura 4. Correlación entre los residuos y alturas predichas con el modelo 6 (ver Cuadro 1), considerando la estructura del error.**



**Figure 5. Evolution of bias and root mean square error (RMSE) by age category, according to model 6 (see Table 2), considering the error structure.**

**Figura 5. Evolución del sesgo y raíz del cuadrado medio del error (RCME) por categoría de edad, de acuerdo con el modelo 6 (ver Cuadro 2), considerando la estructura del error.**

predicted values around the line 0 in almost all the age classes; however, the lack of data in categories over 70 years is evident, causing a considerable increase in the bias. On the other hand, the evolution of the mean error (RMSE) by height category supports the model's presumed good prediction (Figure 5b). These results are similar to those reported by Vargas-Larreta et al. (2010).

## Conclusions

The goodness of fit statistics with the ADA methodology showed significant differences between models, so it is concluded that the Chapman-Richards model presented better fits in the three modalities (anamorphic, polymorphic I and polymorphic II). When the site index curves are superimposed onto the observed data, the anamorphic model best represents the growth in all site qualities. The equation obtained with the GADA methodology is polymorphic with multiple asymptotes and base-age invariant; also, the equation directly estimates the dominant height and site index at any height and base age. Using the Chapman-Richards equation in GADA form will improve the estimation

corrección de la autocorrelación se provee el patrón aleatorio de los residuos alrededor de la línea 0 con varianza homogénea y sin detectarse tendencia definida; estos resultados coinciden con los obtenidos por Castillo et al. (2013), Vargas-Larreta et al. (2010) y Diéguez-Aranda et al. (2005).

En la Figura 5a se analiza la evolución del sesgo de la ecuación 6. Este gráfico muestra la distribución de los valores predichos alrededor de la línea 0 en casi todas las clases de edad; sin embargo, la falta de datos en categorías mayores de 70 años es evidente, lo que ocasiona aumento considerable del sesgo. Por otra parte, la evolución del error medio (REMC) por categoría de altura avala la buena predicción del modelo (Figura 5b). Estos resultados son similares a los reportados por Vargas-Larreta et al. (2010).

## Conclusiones

Los estadísticos de bondad de ajuste con la metodología ADA mostraron diferencias significativas entre modelos, con lo que se concluye que el modelo de Chapman-

of the production potential of *J. deppeana*, through its incorporation into existing forest management programs for the forest region of San Dimas, Durango.

### End of English Version

## References / Referencias

- Alder, D. E. (1980). *Estimación del volumen forestal y predicción del rendimiento, con referencia especial a los trópicos*. Roma, Italia: FAO.
- Avery, T. E. & Burkhart, H. E. (2002). *Forest measurement* (5th ed.). New York, USA: McGraw-Hill.
- Bailey, R. L. & Clutter, J. L. (1974). Base-age invariant polymorphic site curves. *Forest Science*, 20, 155–159. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/1974/00000020/00000002/art00013>
- Carmean, W. H. (1972). Site index curves for upland oaks in the Central States. *Forest Science*, 18, 109–120. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/1972/00000018/00000002/art00004>
- Castillo, L. A., Vargas-Larreta, E., Corral, R. J. J., Nájera, L. J. A., Cruz, C. F., & Hernández, F. J. (2013). Modelo compatible altura-índice de sitio para cuatro especies de pino en Santiago Papasquiaro, Durango. *Revista Mexicana de Ciencias Forestales*, 4(18), 86–103. Obtenido de <http://revistas.inifap.gob.mx/index.php/Forestales/article/view/3003/2483>
- Cieszewski, C. J. (2001). Three methods of deriving advanced dynamic site equations demonstrated on inland Douglas-fir site curves. *Canadian Journal of Forest Research*, 31, 165–173. doi: 10.1139/x00-132
- Cieszewski, C. J. (2002). Comparing fixed- and variable-base-age site equations having single versus multiple asymptotes. *Forest Science*, 48, 7–23. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/2002/00000048/00000001/art00002>
- Cieszewski, C. J. (2003). Developing a well-behaved dynamic site equation using a modified Hossfeld IV function  $Y_3 = (ax^m)/(c + x^{m-1})$ , a simplified mixed-model and scant subalpine fir data. *Forest Science*, 49, 539–554. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/2003/00000049/00000004/art00006>
- Cieszewski, C. J., & Bailey, R. L. (2000). Generalized algebraic difference approach: Theory based derivation of dynamic site equations with polymorphism and variable asymptotes. *Forest Science*, 46, 116–126. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/2000/00000046/00000001/art00015>
- Diéguez, A. U., Álvarez, G. J., Barrio, A. M., & Rojo, A. A. (2005). Site quality equations for *Pinus sylvestris* L. plantations in Galicia (north-west Spain). *Annals of Forest Science*, 62, 143–152. doi: 10.1051/forest:2005006
- Richards presentó mejores ajustes en las tres modalidades (anamórfico, polimórfico I y polimórfico II). Al sobreponer las curvas de índice de sitio sobre los datos observados, el modelo anamórfico representa mejor el crecimiento en todas las calidades de estación. La ecuación obtenida con la metodología GADA es polimórfica con múltiples asíntotas e invariante con respecto a la edad de referencia; además, la ecuación estima directamente la altura dominante y el índice de sitio a cualquier altura y edad de referencia. El uso de la ecuación de Chapman-Richards en forma GADA permitirá mejorar la estimación del potencial productivo de *J. deppeana*, a través de su incorporación a los programas de manejo forestal vigentes para la región forestal de San Dimas, Durango.

### Fin de versión en español

- Diéguez-Aranda, U., Burkhart, H. E., & Amateis, R. L. (2006). Dynamic site model for Loblolly pine (*Pinus taeda* L.) plantations in the United States. *Forest Science*, 52(3), 262–272. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/2006/00000052/00000003/art00006>
- Goelz, J. C. & Burk, T. E. (1992). Development of a well-behaved site index equation: Jack pine in north-central Ontario. *Canadian Journal of Forestry Research*, 22, 776–784. doi: 10.1139/x92-106
- Ker, M. F. & Bowling, C. (1991). Polimorphic site index equations for four New Brunswick softwood species. *Canadian Journal of Forestry Research*, 21, 728–732. doi: 10.1139/x91-103
- Newberry, J. D. (1991). A note on Carmean's estimate of height from stem analysis data. *Forest Science*, 37(1), 368–369. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/1991/00000037/00000001/art00029>
- Payandeh, B. & Wang, Y. (1994). Relative accuracy of a new base-age invariant site index model. *Forest Science*, 40(2), 341–348. Obtenido de <http://www.ingentaconnect.com/content/saf/fs/1994/00000040/00000002/art00011>
- Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, 10, 290–301. doi: 10.1093/jxb/10.2.290
- Ryan, T. P. (1997). *Modern regression methods*. New York, USA: John Wiley and Sons, Inc.
- Statistical Analysis System (SAS Institute). (2004). *SAS/ETS User's Guide, Version 9.1*. Cary, NC, USA: Author.
- Schumacher, F. X. (1939). A new growth curve and its application to timber-yield studies. *Journal of Forestry*, 37, 819–820.
- Skovsgaard, J. P. & Vanclay, J. K. (2008). Forest site productivity: A review of the evolution of dendrometric concepts

- for even-aged stands. *Forestry*, 81(1), 13–31. doi: 10.1093/forestry/cpm041
- Soto, R. J. (2009). Estudio regional forestal UMAFOR, núm. 1006, “San Dimas”. Obtenido de [http://www.conafor.gob.mx:8080/documentos/docs/9/1133ERF\\_UMAFOR1006.pdf](http://www.conafor.gob.mx:8080/documentos/docs/9/1133ERF_UMAFOR1006.pdf)
- Tait, D. E., Cieszewski, C. J., & Bella, I. E. (1988). The stand dynamics of lodge pole pine. *Canadian Journal of Forestry Research*, 18, 1255–1260. doi: 10.1139/x88-193
- Torres-Rojo J. M. (2001). Curvas de índice de sitio de forma y escala variables en investigación forestal. *Agrociencia*, 35, 87–98. Obtenido de <http://www.redalyc.org/articulo.oa?id=30235108>
- Vargas-Larreta, B., Álvarez-González, J. G., Corral-Rivas, J. J., & Aguirre-Calderón, O. (2010). Construcción de curvas dinámicas de índice de sitio para *Pinus cooperi* Blanco. El Salto P. N. Durango. *Revista Fitotecnia Mexicana*, 33(4), 343–351. Obtenido de <http://www.redalyc.org/articulo.oa?id=61015520008>
- Zimmerman, D. L. & Núñez, A. V. (2001). Parametric modelling of growth curve data: An overview (with discussion). *Test*, 10, 1–73. doi:10.1007/BF0259582