

Minimum sample size for fitting compatible taper-volume functions for three pine species in Chihuahua

Tamaño mínimo de muestra para el ajuste de funciones compatibles de ahusamiento-volumen para tres especies de pino en Chihuahua

Juan M. Villela-Suárez¹; Oscar A. Aguirre-Calderón¹; Eduardo J. Treviño-Garza¹; Marco A. González-Tagle¹; Israel Yerena-Yamalle¹; Benedicto Vargas-Larreta^{2*}

¹Universidad Autónoma de Nuevo León, Facultad de Ciencias Forestales. Carretera Nacional núm. 85, km 145. C. P. 67700. Linares, Nuevo León, México.

²Tecnológico Nacional de México/Instituto Tecnológico de El Salto. Tecnológico 101, La Forestal. C. P. 34942. El Salto, Pueblo Nuevo, Durango, México.

*Corresponding author: bvargas@itelsalto.edu.mx; tel.: +52 675 105 6568.

Abstract

Introduction: The choice of sample size is an important decision in the development of volume models and taper functions.

Objective: To calculate the minimum sample size required for fitting compatible taper-volume functions for *Pinus arizonica* Engelm., *P. durangensis* Martínez and *P. engelmannii* Carr. in Chihuahua.

Materials and methods: The methodology was divided into three phases: (i) fitting of a linear regression model to the diameter-height data of 50 trees of each species in the three forest regions; (ii) calculation of the minimum sample size required, and (iii) comparison of the goodness of fit of the taper-volume function using both sample sizes.

Results and discussion: The minimum number of trees calculated ranged from 53 (*Pinus durangensis*) to 88 (*P. engelmannii*) and it is located in the interval reported in studies carried out to estimate the optimal sample size for the development of taper functions. No significant differences were observed in the goodness of fit ($\alpha = 0.05$) in terms of the R^2 and the root mean square error, using the full sample size and the calculated minimum sample size; no significant effect was observed in the stem volume estimates.

Conclusion: The use of small samples in the fit of taper-volume models generates accurate estimates if adequate representation of the study population is ensured.

Keywords: taper function; stem volume; merchantable volume; *Pinus arizonica*, *Pinus durangensis*, *Pinus engelmannii*.

Resumen

Introducción: La elección del tamaño de muestra es una decisión importante en el desarrollo de modelos de volumen y funciones de perfil.

Objetivo: Calcular el tamaño de muestra mínimo requerido para el ajuste de funciones compatibles de ahusamiento-volumen para *Pinus arizonica* Engelm., *P. durangensis* Martínez y *P. engelmannii* Carr. en Chihuahua.

Materiales y métodos: La metodología se dividió en tres fases: (i) ajuste de un modelo de regresión lineal a los datos diámetro-altura de 50 árboles de cada especie en tres regiones forestales; (ii) cálculo del tamaño de muestra mínimo requerido, y (iii) comparación de la bondad de ajuste de la función de ahusamiento-volumen empleando ambos tamaños de muestra.

Resultados y discusión: El número de árboles mínimo calculado osciló entre 53 (*Pinus durangensis*) y 88 (*P. engelmannii*) y se ubica en el intervalo reportado en estudios de estimación del tamaño de muestra óptimo para el desarrollo de funciones de perfil. No se observaron diferencias significativas en la bondad de ajuste ($\alpha = 0.05$) en términos de la R^2 y de la raíz del error medio cuadrático, utilizando el tamaño de muestra completa y el tamaño de muestra mínimo calculado; tampoco se observó efecto significativo en las estimaciones del volumen fustal.

Conclusión: El uso de muestras pequeñas en el ajuste de modelos de ahusamiento-volumen puede generar estimaciones precisas si se garantiza la representación adecuada de la población objeto de estudio.

Palabras clave: Función de perfil; volumen fustal; volumen comercial; *Pinus arizonica*, *Pinus durangensis*, *Pinus engelmannii*.

Introduction

One of the main components of timber forest management programs are total or merchantable volume equations. In Mexico, forest management has diversified rapidly in recent years and the need for models to accurately estimate the stem volume, at different heights, is increasingly evident. The achievement of accurate estimates requires the use of models that can be applied with a certain degree of reliability in the forest population for which they were developed (Burkhart & Tomé, 2012; Crecente-Campo, Rojo-Alboreca, & Diéguez-Aranda, 2009).

In several studies, models for estimating the volume and distribution of forest species products have been compared with different statistical modeling approaches (Diéguez-Aranda, Castedo-Dorado, Álvarez-González, & Rojo, 2006; Lee, Seo, Son, Lee, & von Gadow, 2003; Rojo, Perales, Sánchez-Rodríguez, Álvarez-González, & von Gadow, 2005). One of these consists in the sectioning of the stem, and the fit of a nonlinear regression function to each section (Li & Weiskittel, 2010), which describes the variation in diameter at any point along the stem, from the ground to the tip. The prediction of the diameter along the stem enables the calculation of the total volume or the volume between two points on the stem through the integration of the model, considering that it is not enough to know the total volume of the tree, but the way in which it is distributed in the products demanded by the industry (Heidarsson & Pukkala, 2011).

The origin of taper-volume equations in Mexico dates to the late nineties, when the development of systems of equations to estimate merchantable volume at the individual tree level began (Vargas-Larreta et al., 2017). In this context, the studies of Corral, Návar, and Fernández (1999); Jiménez, Aguirre, Treviño, and Domínguez (1998); Návar and Domínguez (1997); and Navarro, Borja, Musálem, Ramírez, and Granados (2002) stand out.

A feature of taper-volume systems is that, for their development, data from the profile of the stem are required, for which, in most cases, the felling of trees is necessary for the registration of the diameter along the stem. This methodology for obtaining data from the stem profile involves considerable efforts from a technical and economic point of view (Kitikidou & Chatzilazarou, 2008). Therefore, the appropriate choice of sample size is an important decision in the development of volume models and taper functions.

In most studies in which taper functions have been developed, the sample size used has been determined arbitrarily (Vargas-Larreta et al., 2017); for example,

Introducción

Uno de los principales componentes de los programas de manejo forestal maderable son las ecuaciones de volumen total o comercial. En México, el manejo forestal se ha diversificado rápidamente en los últimos años y la necesidad de contar con modelos para la estimación precisa del volumen del fuste, a diferentes alturas, es cada vez más evidente. El logro de estimaciones precisas requiere el uso de modelos que puedan ser aplicados con cierto grado de confiabilidad en la población forestal para la cual fueron desarrollados (Burkhart & Tomé, 2012; Crecente-Campo, Rojo-Alboreca, & Diéguez-Aranda, 2009).

En diversos estudios, los modelos para la estimación del volumen y la distribución de productos de especies forestales se han comparado con distintos enfoques de modelaje estadístico (Diéguez-Aranda, Castedo-Dorado, Álvarez-González, & Rojo, 2006; Lee, Seo, Son, Lee, & von Gadow, 2003; Rojo, Perales, Sánchez-Rodríguez, Álvarez-González, & von Gadow, 2005). Uno de estos consiste en el seccionado del fuste, y el ajuste de una función de regresión no lineal a cada sección (Li & Weiskittel, 2010), que describa la variación del diámetro en cualquier punto a lo largo del tronco, desde el suelo hasta la punta. La predicción del diámetro a lo largo del fuste posibilita el cálculo del volumen total o el volumen entre dos puntos sobre el tronco mediante la integración del modelo, considerando que no es suficiente conocer el volumen total del árbol, sino la manera en que se distribuye en los productos demandados por la industria (Heidarsson & Pukkala, 2011).

El origen de las ecuaciones de ahusamiento-volumen en México se remonta a finales de la década de los noventa, cuando dio inicio el desarrollo de los sistemas de ecuaciones para estimar el volumen comercial a nivel de árbol individual (Vargas-Larreta et al., 2017). En este contexto destacan los trabajos de Corral, Návar, y Fernández (1999); Jiménez, Aguirre, Treviño, y Domínguez (1998); Návar y Domínguez (1997); y Navarro, Borja, Musálem, Ramírez, y Granados (2002).

Una característica de los sistemas de ahusamiento-volumen es que para su desarrollo se requieren datos del perfil del fuste, para lo cual, en la gran mayoría de los casos, es necesario el derribo de árboles para el registro del diámetro a lo largo del fuste. Esta metodología de obtención de datos del perfil fustal implica esfuerzos considerables desde el punto de vista técnico y económico (Kitikidou & Chatzilazarou, 2008). Por lo anterior, la elección apropiada del tamaño de muestra es una decisión importante en el desarrollo de modelos de volumen y de funciones de perfil.

En la mayoría de los estudios en los que se han

Tapia and Nívar (2011) adjusted the Kozak's function (1988) using 110 *Pinus pseudostrobus* Lindl. trees in Nuevo León; while Uranga-Valencia, De los Santos-Posadas, Valdez-Lazalde, López-Upton, and Navarro-Garza (2015) developed a compatible system of total volume and taper-volume from the model by Fang, Borders, and Bailey (2000) for three populations of *P. patula* Schiede ex Schltdl. et Cham. in Oaxaca, with sample sizes of 91, 78 and 70 trees. Taper functions have also been developed for non-*Pinus* species using variable sample sizes; for example, Tlaxcala-Méndez, De los Santos-Posadas, Hernández-de la Rosa, and López-Ayala (2015) evaluated the taper of *Cedrela odorata* L. in Veracruz using 211 trees, while Tamarit-Urías et al. (2014) used data from 144 trees to fit six models to describe the stem profile of *Tectona grandis* L. f. in Campeche, Tabasco, and Chiapas.

Based on the above, authors such as Demaerschalk and Kozak (1974), Kitikidou and Chatzilazarou (2008), Kitikidou (2010) and Saarinen et al. (2019) have proposed some methods for calculating a sample size that ensures an adequate level of accuracy in estimates of relative diameter or merchantable volume, from taper-volume functions. Many of these methods are based on the evaluation of the linear relationship between two variables measured on a preliminary sample, in which, for each value of the independent variable a sample size is estimated; in this way, the maximum value obtained from this estimation is selected as the required minimum sample size (Kitikidou, 2010).

The objective of the present study was to estimate the minimum sample size required for the fit of compatible taper-volume functions for *P. arizonica* Engelm., *P. durangensis* Martínez and *P. engelmannii* Carr. in three forest regions of Chihuahua, Mexico, and evaluate the effect of such sample size in the model behavior.

Materials and methods

Study area

The area comprises the temperate climate forests of the regions of Guadalupe y Calvo, Casas Grandes and Baja Tarahumara in the state of Chihuahua (Figure 1); the area of these three regions represents 11.4 % of the state territory (2 816 897 ha). The entity distinguishes three physiographic provinces: the Sierra Madre Occidental (SMO); the Sierras, Lomeríos and Valles Centrales (SLVC), and the Altiplano and Sierras de Oriente (ASO), characterized by the Chihuahua Desert. The region has 10 climatic units ranging from semi-warm (Acw) and semi-dry (Bs) climates, to wet temperate (Cw) climates; the predominant climate is subhumid temperate with rains in summer (Cw₁). Elevations range from 1 200 to

desarrollado funciones de perfil, el tamaño de muestra utilizado se ha determinado de manera arbitraria (Vargas-Larreta et al., 2017); por ejemplo, Tapia y Nívar (2011) ajustaron la función de Kozak (1988) utilizando 110 árboles de *Pinus pseudostrobus* Lindl. en Nuevo León, mientras que Uranga-Valencia, De los Santos-Posadas, Valdez-Lazalde, López-Upton, y Navarro-Garza (2015) desarrollaron un sistema compatible de volumen total y ahusamiento a partir del modelo de Fang, Borders, y Bailey (2000) para tres poblaciones de *P. patula* Schiede ex Schltdl. et Cham. en Oaxaca, con tamaños de muestra de 91, 78 y 70 árboles. También se han desarrollado funciones de perfil para especies que no pertenecen al género *Pinus* empleando tamaños de muestra variables; por ejemplo, Tlaxcala-Méndez, De los Santos-Posadas, Hernández-de la Rosa, y López-Ayala (2015) evaluaron el ahusamiento de *Cedrela odorata* L. en Veracruz utilizando 211 árboles, mientras que Tamarit-Urías et al. (2014) utilizaron datos de 144 árboles en el ajuste de seis modelos para describir el perfil fustal de *Tectona grandis* L. f. en Campeche, Tabasco y Chiapas.

Con base en lo anterior, autores como Demaerschalk y Kozak (1974), Kitikidou y Chatzilazarou (2008), Kitikidou (2010) y Saarinen et al. (2019) han propuesto algunos métodos para el cálculo de un tamaño de muestra que garantice un nivel de precisión adecuado en las estimaciones del diámetro relativo o el volumen comercial, a partir de funciones de ahusamiento-volumen. Muchos de estos métodos se basan en la evaluación de la relación lineal entre dos variables medidas en una muestra preliminar, en la cual, para cada valor de la variable independiente se estima un tamaño de muestra; de esta forma, el valor máximo obtenido de dicha estimación es seleccionado como el tamaño de muestra mínimo requerido (Kitikidou, 2010).

El objetivo del presente estudio fue estimar el tamaño de muestra mínimo requerido para el ajuste de funciones compatibles de ahusamiento-volumen para *P. arizonica* Engelm., *P. durangensis* Martínez y *P. engelmannii* Carr. en tres regiones forestales de Chihuahua, México, y evaluar el efecto de dicho tamaño de muestra en el comportamiento del modelo.

Materiales y métodos

Área de estudio

El área comprende los bosques de clima templado de las regiones de Guadalupe y Calvo, Casas Grandes y Baja Tarahumara en el estado de Chihuahua (Figura 1); la superficie de estas tres regiones representa 11.4 % del territorio estatal (2 816 897 ha). En la entidad se distinguen tres provincias fisiográficas: la Sierra Madre Occidental (SMO); las Sierras, Lomeríos y Valles Centrales (SLVC),

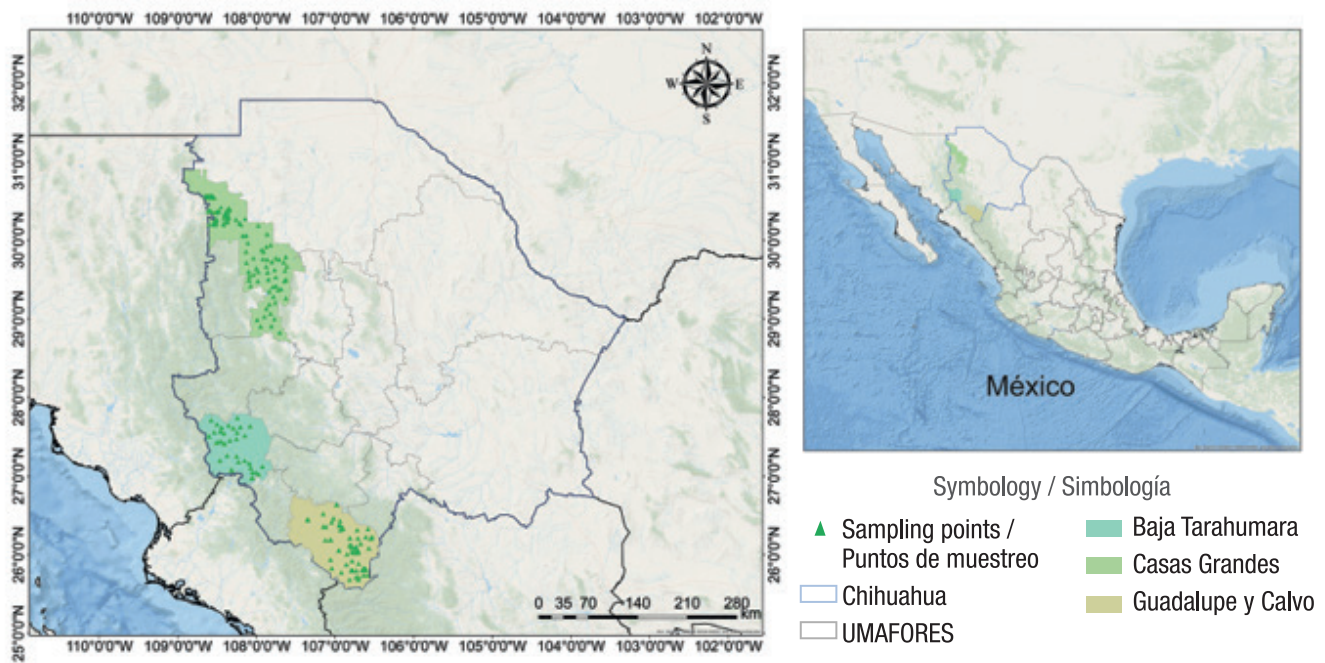


Figure 1. Geographical location of sampling sites in three forest regions of Chihuahua.

Figura 1. Ubicación geográfica de los sitios de muestreo en tres regiones forestales de Chihuahua.

3 200 m, with an average of 1 700 m and precipitation fluctuates between 700 and 1 200 mm, with an average of 933 mm (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad [CONABIO], 2014).

Data

The database was obtained from a sample of 134, 302 and 464 trees of *P. arizonica*, *P. durangensis* and *P. engelmannii*, respectively. The trees were selected through sampling targeted to include the three productivity conditions (high, medium, and low) identified by Vargas-Larreta et al. (2017) in the study area, and the distribution of trees in terms of diameter and height classes. The trees were felled and sectioned to obtain logs at 0.3 m, 0.6 m, and 1.3 m in height; from this point, the sections were obtained at intervals of 2.54 m to the total height (Vargas-Larreta et al., 2017). In addition to the normal diameter with bark (D , cm) and the total height (H , m), in each section the diameter with bark (d_i , cm) and the height of each section with respect to the ground (h_i , m) were measured. The volume of each section was calculated by the Smalian formula, while the upper section (tip) was cubed by the cone formula.

To detect outlier values in the data, the loess (Bi, 2000) procedure of the SAS/STAT™ statistical software (2004) was used with a smoothing parameter of 0.3 (Corral-Rivas, Diéguez-Aranda, Corral-Rivas, & Castedo, 2007). The residues of each loess fit were divided into 10 relative height intervals in which their distribution was examined and the limit values for discrimination of

el Altiplano y Sierras de Oriente (ASO), caracterizado por el Desierto Chihuahuense. La región presenta 10 unidades climáticas que van desde los climas de tipo semicálidos (Ac_w) y semisecos (B_s), hasta los de tipo templado húmedo (C_w); el clima predominante es templado subhúmedo con lluvias en verano (C_w). Las elevaciones varían desde 1 200 hasta 3 200 m, con un promedio de 1 700 m y la precipitación fluctúa entre 700 y 1 200 mm, con un promedio de 933 mm (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad [CONABIO], 2014).

Datos

La base de datos se obtuvo de una muestra de 134, 302 y 464 árboles de *P. arizonica*, *P. durangensis* y *P. engelmannii*, respectivamente. Los árboles se seleccionaron a través de un muestreo dirigido para incluir las tres condiciones de productividad (alta, media y baja) identificadas por Vargas-Larreta et al. (2017) en el área de estudio, y la distribución de los árboles en términos de clases de diámetro y altura. Los árboles se derribaron y seccionaron para obtener trozas a 0.3 m, 0.6 m y 1.3 m de altura; a partir de este punto, las secciones se obtuvieron en intervalos de 2.54 m hasta la altura total (Vargas-Larreta et al., 2017). Además del diámetro normal con corteza (D , cm) y la altura total (H , m), en cada sección se midió el diámetro con corteza (d_i , cm) y la altura de cada sección con respecto al suelo (h_i , m). El volumen de cada sección se calculó mediante la fórmula de Smalian, mientras que la sección superior (punta) se cubió con la fórmula

outliers were calculated, i.e., the lower quartile minus twice the interquartile range and the upper quartile plus twice the interquartile range (Bi, 2000).

Preliminary sample

The methodology developed by Demaerschalk and Kozak (1974) was used in estimating the minimum sample size for the development of taper functions. The method consists in finding two linearly correlated variables; in this case, the relative height (hi/H) and the relative diameter (di/D), being di/D the dependent variable (y) and hi/H the independent variable (x). According to this methodology, the first step is to select a preliminary number of trees, from which the final sample size is calculated. Kitikidou and Chatzilazarou (2008) used 20 trees as an initial sample; however, because the minimum number of trees reported in other studies ranges from 40 (Saarinen et al., 2019) to 60 (Cormier, Reich, Czaplowski, & Bechtold, 1992), in this study, the preliminary sample was 50. The preliminary sample was selected with the “random” function of the Excel program, whereby the trees in the total sample were combined 1 000 times and then 50 trees from each species in each forest region were extracted randomly. Subsequently, a linear regression analysis was performed between hi/H and di/D of the preliminary sample of each species.

Sample size

The next step was to define the reliability criterion which indicates the acceptable error of the final sample size estimate. Such error is defined by the confidence interval of the means of the predicted values for each value of the independent variable, so Kitikidou and Chatzilazarou (2008) predefined such acceptable error as $w_i = 0.05 \cdot X$ or $w_i = 0.10 \cdot X$ based on the desired level of accuracy (5 or 10 %), where X is the standard error of the prediction. Since the obtained value is the acceptable error, it is necessary to calculate the width of the required confidence interval (W_i) that satisfies the ratio $w_i = W_i$, which is obtained from the predicted values calculated from the linear regression between hi/H and di/D . The width of such confidence interval is given by the following equation:

$$W_i = 2t_{Ni-2, \alpha/2} \hat{\sigma} \sqrt{\frac{1}{N_i} + \frac{(X_i - \bar{X})^2}{SSX}}; SSX = VarX (Ni-1)$$

where,

W_i = width of the widest confidence interval of the estimated values for each value of the dependent variable

$t_{Ni-2, \alpha/2}$ = t distribution value for N_{i-2} degrees of freedom at a significance level α

del cono.

Para detectar valores atípicos en los datos, el procedimiento *loess* (Bi, 2000) del paquete estadístico SAS/STAT™ (2004) se empleó con un parámetro de suavizado de 0.3 (Corral-Rivas, Diéguez-Aranda, Corral-Rivas, & Castedo, 2007). Los residuos de cada ajuste *loess* se dividieron en 10 intervalos de altura relativa en los que se examinó su distribución y se calcularon los valores límite para la discriminación de los datos atípicos; es decir, el cuartil inferior menos dos veces el rango intercuartílico y el cuartil superior más dos veces el rango intercuartílico (Bi, 2000).

Muestra preliminar

La metodología desarrollada por Demaerschalk y Kozak (1974) se utilizó en la estimación del tamaño mínimo de muestra para el desarrollo de funciones de perfil. El método consiste en encontrar dos variables correlacionadas linealmente; en este caso, la altura relativa (hi/H) y el diámetro relativo (di/D), siendo di/D la variable dependiente (y) y hi/H la variable independiente (x). De acuerdo con esta metodología, el primer paso consiste en seleccionar un número preliminar de árboles, a partir del cual se calcula el tamaño de muestra final. Kitikidou y Chatzilazarou (2008) utilizaron 20 árboles como muestra inicial; sin embargo, debido a que el número mínimo de árboles reportado en otros estudios oscila entre 40 (Saarinen et al., 2019) y 60 (Cormier, Reich, Czaplowski, & Bechtold, 1992), en el presente, la muestra preliminar fue 50. La muestra preliminar se seleccionó con la función “aleatorio” del programa Excel, mediante la cual, los árboles de la muestra total se combinaron 1 000 veces para después extraer 50 árboles al azar de cada especie en cada región forestal. Posteriormente, se realizó un análisis de regresión lineal entre hi/H y di/D de la muestra preliminar de cada especie.

Tamaño de muestra

El siguiente paso consistió en definir el criterio de confiabilidad que indica el error aceptable de la estimación final del tamaño de muestra. Dicho error está definido por el intervalo de confianza de las medias de los valores predichos para cada valor de la variable independiente, por lo que Kitikidou y Chatzilazarou (2008) predefinen dicho error aceptable como $w_i = 0.05 \cdot X$ o $w_i = 0.10 \cdot X$ en función del nivel de precisión deseado (5 o 10 %), donde X es el error estándar de la predicción. Dado que el valor obtenido es el error aceptable, es necesario el cálculo del ancho del intervalo de confianza (W_i) requerido que satisface la relación $w_i = W_i$, el cual se obtiene a partir de los valores predichos obtenidos de la regresión lineal entre hi/H y di/D . El ancho de dicho intervalo de confianza está dado por la ecuación siguiente:

$\hat{\sigma}$ = standard estimate error, calculated from the regression analysis of the preliminary sample
 N_i = sample size for the i -th value of the independent variable X_i
 \bar{X} = mean of the independent variable
 $VarX$ = variance of the independent variable.

The calculation of the t value requires knowledge of the value of N (sample size), for which an iterative procedure is used (Kitikidou & Chatzilazarou, 2008), which assumes that the sample size N_i , for the i -th value of the independent variable, is implicitly present in the critical value of t when the maximum acceptable width of the confidence interval of the estimated values of W_i is defined. Therefore, the sample size N_i , for each specific value of X_i (hi/H), is calculated by solving N as follows:

$$N = \frac{4t^2_{N-2, \alpha/2} \sigma^2 [1 + (x_j - (X_j - X_i))^2 / VarX]}{W_i^2}$$

where

N = sample size that satisfies $w_i = W_i$; that is, the value of N_i that equals W_i with the acceptable error (w_i) is the number of trees selected as the final sample size required for the fit of the taper equation. The solution of this equation requires the use of an iterative procedure in which different sample sizes are tested. Demaerschalk and Kozak (1974) define the most demanding value of X_p , with respect to the required value of W_p , as one that relates directly to the number closest to the minimum sample size required. The most demanding value X_i is the one for which the ratio $(w_i - W_i)/w_i$ is maximum. In practice, the most demanding value X_i is closely related to the extreme values of the data range of the independent variable hi/H .

Regression hypothesis

The Kolmogorov-Smirnov Z statistical test was applied to verify that the residues of the linear regression between hi/H and di/D presented a normal distribution; homogeneity was verified with the Levene test (1960). In addition, the values of the residues were plotted against those predicted with linear regression, to detect problems of heteroscedasticity in the errors.

Adjusted taper-volume function

The compatible taper-volume function by Fang et al. (2000) was selected, which has proven to be a highly flexible model for both coniferous and broadleaved species of temperate and some tropical climates (García-Espinoza et al., 2019; López-Martínez et al., 2019; Vargas-Larreta et al., 2017). The function has the following expression:

$$d_i = c_1 \sqrt{h^{(k-b_1)/b_1} (1-q)^{(k-\beta)/\beta} \alpha_1^{1+1/2} \alpha_1^{1/2}}$$

donde,

W_i = ancho del intervalo de confianza más amplio de los valores estimados para cada valor de la variable dependiente

$t_{N_i-2, \alpha/2}$ = valor de la distribución de t para N_{i-2} grados de libertad a un nivel de significancia α

$\hat{\sigma}$ = error estándar de la estimación, calculado del análisis de regresión de la muestra preliminar

N_i = tamaño de muestra para el i -ésimo valor de la variable independiente X_i

\bar{X} = media de la variable independiente

$VarX$ = varianza de la variable independiente.

El cálculo del valor de t requiere conocer el valor de N (tamaño de muestra), para lo cual se utiliza un procedimiento iterativo (Kitikidou & Chatzilazarou, 2008), el cual asume que el tamaño de muestra N_i , para el i -ésimo valor de la variable independiente, está implícitamente presente en el valor crítico de t cuando se define el ancho máximo aceptable del intervalo de confianza de los valores estimados de W_i . Por tanto, el tamaño de muestra N_i , para cada valor específico de X_i (hi/H), se calcula resolviendo N de la manera siguiente:

$$N = \frac{4t^2_{N-2, \alpha/2} \sigma^2 [1 + (x_j - (X_j - X_i))^2 / VarX]}{W_i^2}$$

donde N = tamaño de muestra que satisface $w_i = W_i$; es decir, el valor de N_i que iguala W_i con el error aceptable (w_i) es el número de árboles seleccionado como el tamaño de muestra final requerido para el ajuste de la ecuación de perfil. La solución de esta ecuación requiere el uso de un procedimiento iterativo en el cual se van probando tamaños distintos de muestra. Demaerschalk y Kozak (1974) definen el valor más exigente de X_p , con respecto al valor requerido de W_p , como aquel que se relaciona de manera directa con el número más próximo al tamaño de muestra mínimo requerido. El valor más exigente X_i es aquel para el cual la relación $(w_i - W_i)/w_i$ es máximo. En la práctica, el valor más exigente X_i está estrechamente relacionado con los valores extremos del rango de datos de la variable independiente hi/H .

Hipótesis de regresión

La prueba estadística Z de Kolmogorov-Smirnov se aplicó para verificar que los residuos de la regresión lineal entre hi/H y di/D presentaran una distribución normal; la homogeneidad se verificó con la prueba de Levene (1960). Adicionalmente, los valores de los residuos se graficaron contra los predichos con la regresión lineal, para detectar problemas de heterocedasticidad en los errores.

Función de ahusamiento-volumen ajustada

Se seleccionó la función compatible de ahusamiento-volumen de Fang et al. (2000), la cual ha demostrado

$$c_1 = \sqrt{\frac{a_0 d^{a_1} h^{a_2 - k/b_1}}{b_1(r_0 - r_1) + b_2(r_1 - \alpha_1 r_2) + b_3 \alpha_1 r_2}}$$

$$\beta = b_1^{1/(1+l_2)b_2} b_2^{l_1} b_3^{l_2}; \alpha_1 = (1-p_1)^{\frac{(b_2-b_1)k}{b_1 b_2}}; \alpha_2 = (1-p_2)^{\frac{(b_3-b_2)k}{b_2 b_3}}$$

$$r_0 = (1-h_{st}/h)^{k/b_1}; r_1 = (1-p_1)^{k/b_1}; r_2 = (1-p_2)^{k/b_2}; k = \pi/40000;$$

$$l_1 = 1 \text{ si } P_1 \leq q \leq P_2; 0 \text{ otherwise}$$

$$l_2 = 1 \text{ si } P_2 < q \leq 1; 0 \text{ otherwise}$$

where,

p_1 and $p_2 = h_1/H$ y h_2/H , respectively (relative heights from the ground at which the two inflection points that the model assumes occur)

$$q = h_i/H$$

di = stem diameter at height i

h_{st} = stump height

a_0, a_1, a_2, b_1, b_2 and b_3 = parameters to be estimated.

The Fang et al. function (2000) was fitted to the data pairs $di/D - hi/H$ of the complete sample and the minimum sample calculated for each species and region. The fit was performed using the seemingly unrelated regression procedure (SUR) with the PROC MODEL procedure of the SAS program (SAS Institute Inc., 2004).

Evaluation of developed profile functions

The statistics used for the evaluation and comparison of the functions were: mean bias (\bar{E}), root mean square error (REMC) and coefficient of determination for nonlinear regression (R^2). The expressions of these statistics are as follows:

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

$$R^2 = r^2_{y_{ij} \hat{y}_{ij}}$$

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

where,

y_i, \hat{y}_i and \bar{y}_i = observed, predicted and mean value of the dependent variable (di/D), respectively

r = correlation coefficient between observed and predicted values of the dependent variable

n = number of observations

p = number of equation parameters.

Results

The mean, maximum and minimum values, variance, and standard error of the variables of the sample used

ser un modelo altamente flexible, tanto para especies de coníferas como para latifoliadas de clima templado y algunas tropicales (García-Espinoza et al., 2019; López-Martínez et al., 2019; Vargas-Larreta et al., 2017). La función tiene la expresión siguiente:

$$d_i = c_1 \sqrt{h^{(k-b_1)/b_1} (1-q)^{(k-\beta)/\beta} \alpha_1^{l_1+1} \alpha_2^{l_2}}$$

$$c_1 = \sqrt{\frac{a_0 d^{a_1} h^{a_2 - k/b_1}}{b_1(r_0 - r_1) + b_2(r_1 - \alpha_1 r_2) + b_3 \alpha_1 r_2}}$$

$$\beta = b_1^{1/(1+l_2)b_2} b_2^{l_1} b_3^{l_2}; \alpha_1 = (1-p_1)^{\frac{(b_2-b_1)k}{b_1 b_2}}; \alpha_2 = (1-p_2)^{\frac{(b_3-b_2)k}{b_2 b_3}}$$

$$r_0 = (1-h_{st}/h)^{k/b_1}; r_1 = (1-p_1)^{k/b_1}; r_2 = (1-p_2)^{k/b_2}; k = \pi/40000;$$

$$l_1 = 1 \text{ si } P_1 \leq q \leq P_2; 0 \text{ en caso contrario}$$

$$l_2 = 1 \text{ si } P_2 < q \leq 1; 0 \text{ en caso contrario}$$

donde,

p_1 y $p_2 = h_1/H$ y h_2/H , respectivamente (alturas relativas desde el suelo a las que se producen los dos puntos de inflexión que el modelo supone)

$$q = h_i/H$$

di = diámetro del fuste a la altura i

h_{st} = altura del tocón

a_0, a_1, a_2, b_1, b_2 y b_3 = parámetros a ser estimados durante el ajuste.

La función de Fang et al. (2000) se ajustó a los pares de datos $di/D - hi/H$ de la muestra completa y de la muestra mínima calculada para cada especie y región. El ajuste se realizó mediante el procedimiento de Regresión Aparentemente no Relacionada (SUR, por sus siglas en inglés) con el procedimiento PROC MODEL del programa SAS (SAS Institute Inc., 2004).

Evaluación de las funciones de perfil desarrolladas

Los estadísticos utilizados para la evaluación y comparación de las funciones fueron: sesgo promedio (\bar{E}), raíz del error medio cuadrático (REMC) y coeficiente de determinación para regresión no lineal (R^2). Las expresiones de estos estadísticos son las siguientes:

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

$$R^2 = r^2_{y_{ij} \hat{y}_{ij}}$$

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

Table 1. Statistical summary of the variables measured in three *Pinus* species.
Cuadro 1. Resumen estadístico de las variables medidas en tres especies de *Pinus*.

Region/ Región	Species/Especie	Statistical/Estadísticos	D (cm)	H (m)	di/D	hi/H	
Casas Grandes	<i>P. durangensis</i>	N	152	152	1 294	1 294	
		Mean/Media	31.19	14.04	0.85	0.33	
		Variance/Varianza	173.75	23.19	0.11	0.10	
		Standard Error/Error estándar	13.18	4.81	0.33	0.32	
		Minimum/Mínimo	8.00	4.77	0.06	0.00	
			Maximum/Máximo	57.35	24.16	1.63	0.99
	<i>P. engelmannii</i>	N	162	162	1 368	1 368	
		Mean/Media	30.79	13.77	0.85	0.33	
		Variance/Varianza	195.93	27.14	0.10	0.10	
		Standard Error/Error estándar	13.99	5.21	0.32	0.32	
Minimum/Mínimo		9.00	4.34	0.03	0.00		
		Maximum/Máximo	61.25	27.32	1.57	0.99	
Baja Tarahumara	<i>P. durangensis</i>	N	150	150	1 323	1 323	
		Mean/Media	27.25	15.18	0.85	0.33	
		Variance/Varianza	175.26	23.49	0.097	0.11	
		Standard Error/Error estándar	13.23	4.84	0.312	0.32	
		Minimum/Mínimo	7.50	6.35	0.047	0.00	
			Maximum/Máximo	56.5	27.25	1.642	0.99
	<i>P. engelmannii</i>	N	149	149	1 357	1 357	
		Mean/Media	30.74	16.02	0.85	0.34	
		Variance/Varianza	188.28	30.56	0.09	0.10	
		Standard Error/Error estándar	13.72	5.52	0.30	0.31	
Minimum/Mínimo		8.75	6.38	0.07	0.00		
		Maximum/Máximo	57.70	28.41	1.81	0.98	
Guadalupe y Calvo	<i>P. arizonica</i>	N	134	134	1 088	1 088	
		Mean/Media	28.06	15.83	0.84	0.35	
		Variance/Varianza	157.28	30.72	0.11	0.12	
		Standard Error/Error estándar	12.54	5.54	0.33	0.34	
		Minimum/Mínimo	7.50	5.06	0.05	0.00	
			Maximum/Máximo	57.00	28.3	1.56	0.99
	<i>P. engelmannii</i>	N	152	152	1 254	1 254	
		Mean/Media	27.45	15.65	0.79	0.37	
		Variance/Varianza	95.83	21.55	0.12	0.12	
		Standard Error/Error estándar	9.78	46.64	0.35	0.35	
Minimum/Mínimo		9.40	4.80	0.01	0.01		
		Maximum/Máximo	53.00	25.30	1.86	0.99	

N = number of trees, D = breast height diameter, H = total height, di/D = relative diameter, hi/H = relative height.

N = número de árboles, D = diámetro normal, H = altura total, di/D = diámetro relativo, hi/H = altura relativa.

are presented in Table 1. The behavior of the relative diameter (di/D) with respect to the relative height (hi/H) of the data included in the adjustments is shown in Figure 2.

Determination of minimum sample size

According to Table 2, the fitted and reliable values (X_i), named by Kitikidou and Chatzilazarou (2008) as more demanding values of the independent variable, were in a range of 0.975 (*P. engelmannii* in Baja Tarahumara) to 0.995 (*P. durangensis* in Casas Grandes); these coincided with the higher value of hi/H in all cases. The calculation of the required acceptable error with 5 and 10 % accuracy resulted in the w_i values shown in Table 2. With the values of X_i and w_i the final value of N was

donde,

y_i, \hat{y}_i y \bar{y}_i = valor observado, predicho y promedio de la variable dependiente (di/D), respectivamente

r = coeficiente de correlación entre los valores observados y predichos de la variable dependiente

n = número de observaciones

p = número de parámetros de la ecuación.

Resultados

Los valores medio, máximo y mínimo, la varianza y el error estándar de las variables de la muestra utilizada se presentan en el Cuadro 1. El comportamiento del diámetro relativo (di/D) respecto a la altura relativa (hi/H) de los datos incluidos en los ajustes se muestra en la Figura 2.

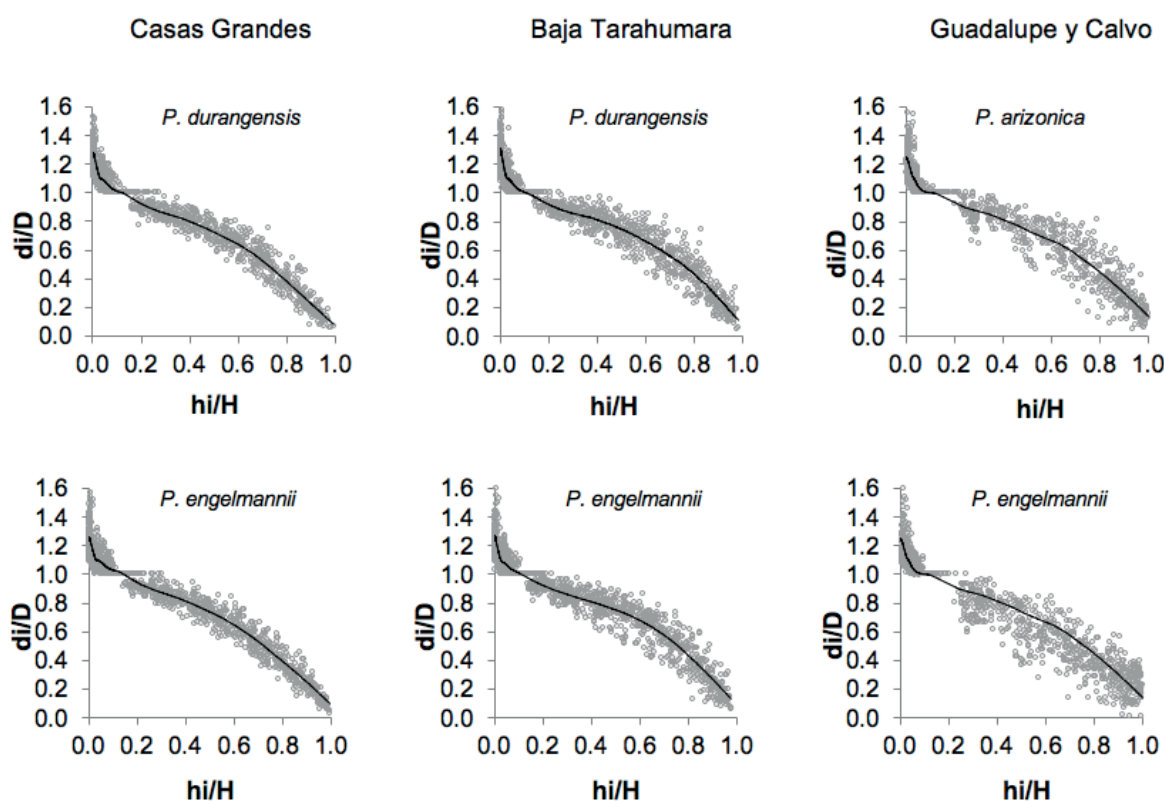


Figure 2. Relative diameter-relative height ratio (di/D - hi/H) in three *Pinus* species from three forest regions, after data cleansing according to the methodology of Bi (2000). The solid line is the average prediction through nonparametric local fit with a 0.3 parameter spline smoothing curve.

Figura 2. Relación diámetro relativo-altura relativa (di/D - hi/H) en tres especies de *Pinus* de tres regiones forestales, después de la depuración de datos de acuerdo con la metodología de Bi (2000). La línea sólida es la predicción media a través del ajuste local no paramétrico con una curva suavizada *spline* de parámetro 0.3.

Table 2. Minimum sample size calculated to fit the compatible taper-volume function by Fang et al. (2000) in three species of *Pinus*.

Cuadro 2. Tamaño mínimo de muestra calculado para ajustar la función compatible de ahusamiento-volumen de Fang et al. (2000) en tres especies de *Pinus*.

Region/ Región	Species/ Especie	X_i	EE	w_i		Sections/ Secciones		Trees/Árboles	
				5 %	10 %	5 %	10 %	5 %	10 %
Casas Grandes	<i>P. durangensis</i>	0.9948	0.3271	0.0163	0.0327	2 388	420	299	53
	<i>P. engelmannii</i>	0.9898	0.3235	0.0161	0.0323	2 409	424	302	54
Baja Tarahumara	<i>P. durangensis</i>	0.9798	0.3380	0.0169	0.0338	2 850	502	357	63
	<i>P. engelmannii</i>	0.9750	0.3316	0.0165	0.0331	3 067	539	384	68
Guadalupe y Calvo	<i>P. arizonica</i>	0.9826	0.3262	0.0163	0.0326	3 322	584	416	73
	<i>P. engelmannii</i>	0.9939	0.1722	0.0172	0.0345	3 983	700	498	88

X_i = adjusted and reliable values of the independent variable; EE = standard error of the linear regression between relative diameter-relative height ($d_i/D-h_i/H$); w_i = acceptable error.

X_i = valores ajustados y confiables de la variable independiente; EE = error estándar de la regresión lineal entre diámetro relativo-altura relativa ($d_i/D-h_i/H$); w_i = error aceptable.

obtained, considered as the minimum size required to fit taper functions with a reliability level of 5 and 10 %.

The number of sections required for the calculated final sample ranged from 420 (53 trees) for *P. durangensis* in Casas Grandes, and 700 (88 trees) for *P. engelmannii* in Guadalupe y Calvo, with a reliability level of 10 %. In case the desired accuracy level is 5 %, the number of sections increases considerably, since a minimum sample of 299 would be required (*P. durangensis* in Casas Grandes) to 498 trees (*P. engelmannii* in Guadalupe y Calvo) (Table 2).

In Table 3, the Kolmogorov-Smirnov residue normality test indicated that the residues of the regressions approximate the normal distribution (significance of $Z > 0.05$) for a probability of 5 %, which means that the hypothesis tests are valid. On the other hand, the significance of Levene's statistic L was greater than the probability $\alpha = 0.05$ in the variance homogeneity test, so the null hypothesis that residues have homogeneous variance is accepted.

Figure 3 shows that although the residues of the linear regression between d_i/D and h_i/H seem to present a defined pattern, this corresponds to the typical distribution of residues in this type of relationships (Kitikidou & Chatzilazarou, 2008), therefore, based on the test of homogeneity of variances, the absence of heterocedasticity is assumed.

Compatible taper-volume function

Table 4 presents the values of the estimated parameters of the Fang et al. (2000) function, from the complete

Determinación de tamaño mínimo de muestra

De acuerdo con el Cuadro 2, los valores ajustados y confiables (X_i), denominados por Kitikidou y Chatzilazarou (2008) como valores más exigentes de la variable independiente, se ubicaron en un intervalo de 0.975 (*P. engelmannii* en Baja Tarahumara) a 0.995 (*P. durangensis* en Casas Grandes); estos coincidieron con el valor superior de h_i/H en todos los casos. El cálculo del error aceptable requerido con 5 y 10 % de precisión dio como resultado los valores de w_i mostrados en el Cuadro 2. Con los valores de X_i y W_i se obtuvo el valor final de N , considerado como el tamaño mínimo requerido para ajustar funciones de perfil con un nivel de confiabilidad de 5 y 10 %.

El número de secciones requerido para la muestra final calculada osciló entre 420 (53 árboles) para *P. durangensis* en Casas Grandes, y 700 (88 árboles) para *P. engelmannii* en Guadalupe y Calvo, con un nivel de confiabilidad de 10 %. En caso de que el nivel de precisión deseado sea 5 %, el número de secciones aumenta considerablemente, ya que se requeriría una muestra mínima de 299 (*P. durangensis* en Casas Grandes) a 498 árboles (*P. engelmannii* en Guadalupe y Calvo) (Cuadro 2).

En el Cuadro 3, la prueba de normalidad de residuos de Kolmogorov-Smirnov indicó que los residuos de las regresiones se aproximan a la distribución normal (significancia de $Z > 0.05$) para una probabilidad del 5 %, lo que significa que las pruebas de hipótesis son válidas. Por su parte, la significancia del estadístico L de Levene fue mayor que la probabilidad $\alpha = 0.05$ en la prueba de homogeneidad de varianzas, por lo que

Table 3. Kolmogorov-Smirnov normality test and Levene variance homogeneity test values for regression residues between relative diameter and height (d_i/D and h_i/H) of three *Pinus* species.

Cuadro 3. Valores de las pruebas de normalidad de Kolmogorov-Smirnov y de homogeneidad de varianzas de Levene para los residuos de la regresión entre el diámetro y la altura relativos (d_i/D y h_i/H) de tres especies de *Pinus*.

Region / Región	Species / Especie	Standard deviation / Desviación estándar	Kolmogorov-Smirnov		Levene	
			Z	Significance / Significancia	L	Significance / Significancia
Casas Grandes	<i>P. durangensis</i>	0.0884	0.2811	0.0646	0.193	0.660
	<i>P. engelmannii</i>	0.0851	0.3586	0.0613	0.465	0.496
Baja Tarahumara	<i>P. durangensis</i>	0.1015	0.0944	0.0539	3.037	0.082
	<i>P. engelmannii</i>	0.1018	0.3395	0.0523	1.029	0.311
Guadalupe y Calvo	<i>P. arizonica</i>	0.1058	0.5281	0.0578	0.673	0.412
	<i>P. engelmannii</i>	0.1278	0.3350	0.0515	0.042	0.838

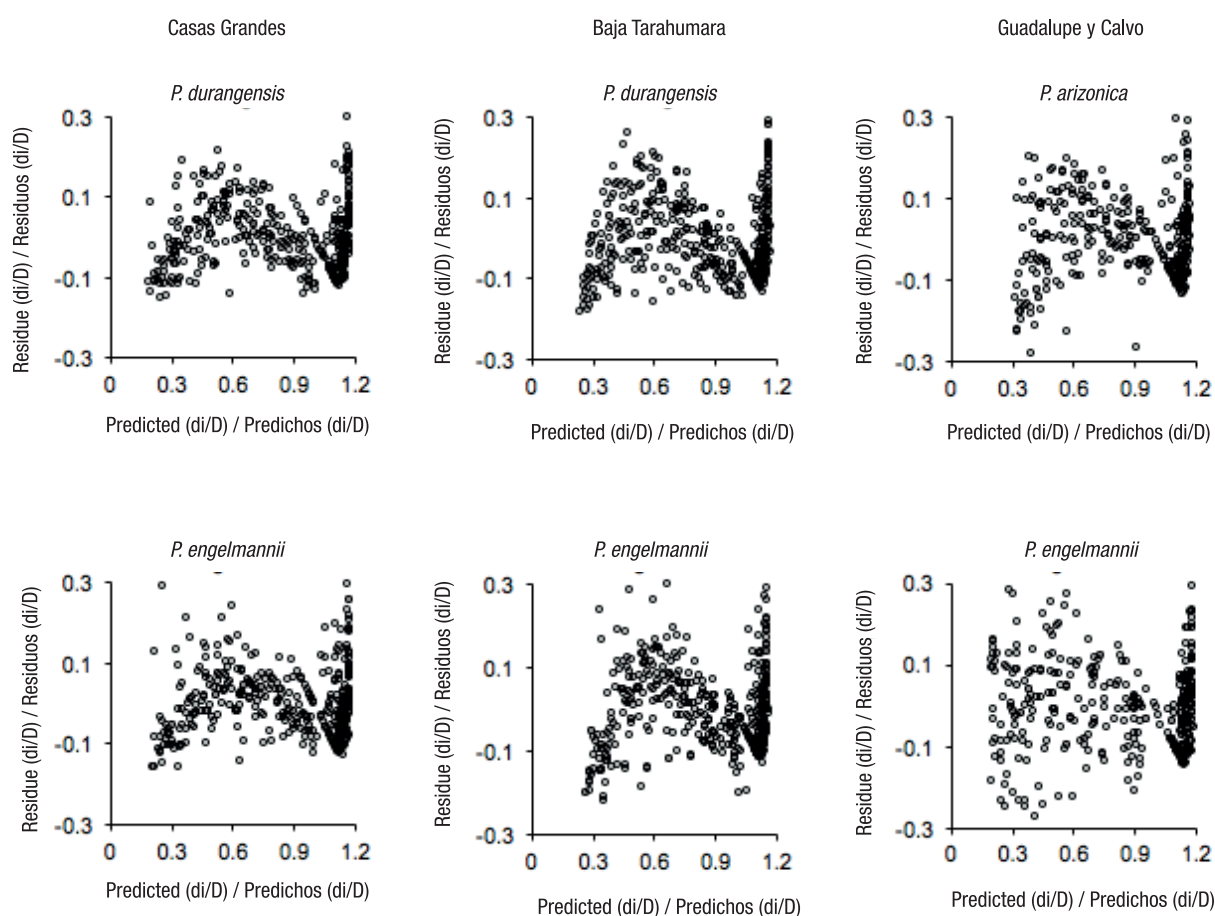


Figure 3. Residues observed against relative diameter estimates (d_i/D) of linear regression models for three *Pinus* species in three forest regions of Chihuahua.

Figura 3. Residuos observados contra los estimados del diámetro relativo (d_i/D) de los modelos de regresión lineal para tres especies de *Pinus* en tres regiones forestales de Chihuahua.

Table 4. Estimated and statistical goodness of fit parameters of the compatible taper-volume function by Fang et al. (2000) for three species of *Pinus* in Chihuahua, Mexico. Cuadro 4. Parámetros estimados y estadísticos de bondad de ajuste de la función compatible de alusamiento-volumen de Fang et al. (2000) para tres especies de *Pinus* en Chihuahua, México.

Region/ Región	Species/ Especie	Fit with full sample / Ajuste con la muestra completa										Fit with calculated minimum sample						
		N	Parameter/ Parámetro	Estimator/ Estimador	EE	t-value/ Valor-t	P > t	\hat{E}	REMC	R ²	N	Estimator/ Estimador	EE	t-value/ Valor-t	P > t	\hat{E}	REMC	R ²
Casas Grandes	<i>Pinus durangensis</i>	152	a ₀	0.000074	0.0000	22.25	<.0001	0.009	2.036	0.985	53	0.000070	0.0000	12.49	<.0001	0.029	1.940	0.981
		a ₁	1.935462	0.0175	110.37	<.0001					1.921104	0.0298	64.47	<.0001				
		a ₂	0.872526	0.0207	42.08	<.0001					0.910056	0.0364	25.03	<.0001				
		P ₁	0.048887	0.0019	25.22	<.0001					0.049701	0.0032	15.51	<.0001				
		P ₂	0.615636	0.0216	28.53	<.0001					0.607429	0.0326	18.63	<.0001				
		b ₁	0.000008	0.0000	25.27	<.0001					0.000008	0.0000	5.77	<.0001				
	<i>Pinus engelmannii</i>	b ₂	0.000038	0.0000	97.78	<.0001					0.000038	0.0000	55.08	<.0001				
		b ₃	0.000031	0.0000	60.74	<.0001					0.000030	0.0000	37.66	<.0001				
		a ₀	0.000072	0.0000	27.13	<.0001	0.029	1.859	0.988	54	0.000064	0.0000	15.43	<.0001	0.032	1.853	0.983	
		a ₁	1.939417	0.0152	127.46	<.0001					1.961464	0.0269	73.02	<.0001				
		a ₂	0.883090	0.0179	49.38	<.0001					0.892157	0.0317	28.18	<.0001				
		P ₁	0.049435	0.0021	23.18	<.0001					0.050916	0.0042	12.01	<.0001				
Baja Tarahumara	<i>Pinus durangensis</i>	150	P ₂	0.589797	0.0180	32.80	<.0001				0.601891	0.0451	13.34	<.0001				
		b ₁	0.000010	0.0000	24.43	<.0001					0.000010	0.0000	12.78	<.0001				
		b ₂	0.000039	0.0000	99.02	<.0001					0.000038	0.0000	54.73	<.0001				
		b ₃	0.000031	0.0000	79.53	<.0001					0.000032	0.0000	39.17	<.0001				
		a ₀	0.000085	0.0000	22.66	<.0001	0.012	2.044	0.982	63	0.000098	0.0000	14.58	<.0001	0.015	1.994	0.981	
		a ₁	1.949807	0.0188	103.69	<.0001					1.838358	0.0304	60.47	<.0001				
	<i>Pinus durangensis</i>	a ₂	0.808635	0.0257	31.52	<.0001					0.891305	0.0440	20.25	<.0001				
		P ₁	0.044911	0.0018	24.36	<.0001					0.043581	0.0026	17.06	<.0001				
		P ₂	0.588468	0.0332	17.75	<.0001					0.736681	0.0388	18.99	<.0001				
		b ₁	0.000007	0.0000	23.10	<.0001					0.000007	0.0000	15.87	<.0001				
		b ₂	0.000039	0.0000	80.53	<.0001					0.000037	0.0000	82.86	<.0001				
		b ₃	0.000033	0.0000	65.27	<.0001					0.000031	0.0000	25.83	<.0001				

N = number of sample trees; EE = standard estimator error; \hat{E} = bias; REMC = root of mean quadratic error.

N = número de árboles de la muestra; EE = error estándar del estimador; \hat{E} = Sesgo; REMC = raíz del error medio cuadrático.

Table 4. Estimated and statistical goodness of fit parameters of the compatible taper-volume function by Fang et al. (2000) for three species of *Pinus* in Chihuahua, Mexico. (cont.)
Cuadro 4. Parámetros estimados y estadísticos de bondad de ajuste de la función compatible de abusamiento-volumen de Fang et al. (2000) para tres especies de *Pinus* en Chihuahua, México. (cont.)

Region/ Región	Species/ Especie	Fit with full sample / Ajuste con la muestra completa										Fit with calculated minimum sample						
		N	Parameter/ Parámetro	Estimator/ Estimador	EE	t-value/ Valor-t	P > t	\hat{E}	REMC	R ²	N	Estimator/ Estimador	EE	t-value/ Valor-t	P > t	\hat{E}	REMC	R ²
Baja Tarahumara	<i>Pinus engelmannii</i>	149	a ₀	0.000094	0.0000	19.86	<0.001	-0.029	2.360	0.978	68	0.000093	0.0000	13.30	<0.001	0.014	2.388	0.973
			a ₁	1.892595	0.0211	89.56	<0.001					1.926250	0.0322	59.75	<0.001			
			a ₂	0.855311	0.0262	32.70	<0.001					0.821162	0.0370	22.17	<0.001			
			P ₁	0.047248	0.0022	21.62	<0.001					0.052128	0.0035	14.91	<0.001			
			P ₂	0.714252	0.0172	41.59	<0.001					0.705480	0.0210	33.52	<0.001			
			b ₁	0.000008	0.0000	20.49	<0.001					0.000009	0.0000	14.72	<0.001			
			b ₂	0.000040	0.0000	104.18	<0.001					0.000041	0.0000	65.55	<0.001			
			b ₃	0.000031	0.0000	43.08	<0.001					0.000029	0.0000	29.34	<0.001			
			a ₀	0.000074	0.0000	15.31	<0.001	0.017	2.668	0.969	73	0.000059	0.0000	13.61	<0.001	-0.053	2.077	0.967
			a ₁	2.015916	0.0356	56.62	<0.001					2.013093	0.0385	52.77	<0.001			
Guadalupe y Calvo	<i>Pinus arizonica</i>		a ₂	0.778588	0.0438	17.77	<0.001				0.867904	0.0485	17.91	<0.001				
			P ₁	0.050732	0.0028	18.31	<0.001					0.053447	0.0043	12.44	<0.001			
			P ₂	0.861684	0.0380	22.69	<0.001					0.396530	1.4410	0.28	<0.001			
			b ₁	0.000008	0.0000	16.25	<0.001					0.000009	0.0000	12.13	<0.001			
			b ₂	0.000038	0.0000	98.52	<0.001					0.000043	0.0000	18.89	<0.001			
			b ₃	0.000044	0.0000	24.82	<0.001					0.000042	0.0000	54.66	<0.001			
			a ₀	0.000066	0.0000	13.31	<0.001	0.090	2.728	0.961	88	0.000057	0.0000	10.72	<0.001	0.004	3.004	0.952
			a ₁	2.042404	0.0281	72.65	<0.001					1.900400	0.0395	48.09	<0.001			
			a ₂	0.735248	0.0327	22.48	<0.001					0.965490	0.0531	18.57	<0.001			
			P ₁	0.078316	0.0049	15.84	<0.001					0.061481	0.0049	12.45	<0.001			
<i>Pinus engelmannii</i>		P ₂	0.851854	0.0185	46.08	<0.001				0.933220	0.0147	63.28	<0.001					
		b ₁	0.000011	0.0000	19.04	<0.001				0.000009	0.0000	11.80	<0.001					
		b ₂	0.000034	0.0000	76.87	<0.001				0.000034	0.0000	70.84	<0.001					
		b ₃	0.000049	0.0000	27.63	<0.001				0.000073	0.0001	3.74	<0.001					

N = number of sample trees; EE = standard estimator error; \hat{E} = bias; REMC = root of mean quadratic error.
 N = número de árboles de la muestra; EE = error estándar del estimador; \hat{E} = Sesgo; REMC = raíz del error medio cuadrático.

sample and the calculated minimum sample, as well as the goodness of fit statistics for each species. The fit of the function by Fang et al. (2000) with the minimum sample size calculated generates values of R^2 slightly lower than those obtained with the full sample size for all species. Differences in R^2 in both fittings ranged from 0.2 % (*P. arizonica* in Guadalupe y Calvo) and 0.9 % (*P. engelmannii* in Guadalupe y Calvo), with an average difference of 0.42 %.

On the other hand, the REMC values for the fitted function with the full sample and with the calculated minimum sample were also similar (Table 4). The lowest REMC values were 1.859 and 1.853 cm for the complete and minimum sample, respectively, while the highest were 2.728 and 3.004 cm. The REMC value was lower in four (*P. durangensis* and *P. engelmannii* in Casas Grandes, *P. durangensis* in Baja Tarahumara and *P. arizonica* in Guadalupe y Calvo) of the six cases with the minimum calculated sample; that is, the accuracy of the model was higher. The mean bias was slightly lower with the complete sample in four of the six cases.

se acepta la hipótesis nula de que los residuos tienen varianza homogénea.

La Figura 3 muestra que a pesar de que los residuos de la regresión lineal entre di/D y hi/H parecieran presentar un patrón definido, este corresponde a la distribución típica de los residuos en este tipo de relaciones (Kitikidou & Chatzilazarou, 2008), por tanto, con base en la prueba de homogeneidad de varianzas, se asume la ausencia de heterocedasticidad.

Función compatible de ahusamiento-volumen

El Cuadro 4 presenta los valores de los parámetros estimados de la función de Fang et al. (2000), a partir de la muestra completa y la muestra mínima calculada, así como los estadísticos de bondad de ajuste para cada especie. El ajuste de la función de Fang et al. (2000) con el tamaño de muestra mínimo calculado genera valores de R^2 ligeramente menores que los obtenidos con el tamaño de muestra completo para todas las especies. Las diferencias en R^2 en ambos ajustes oscilaron

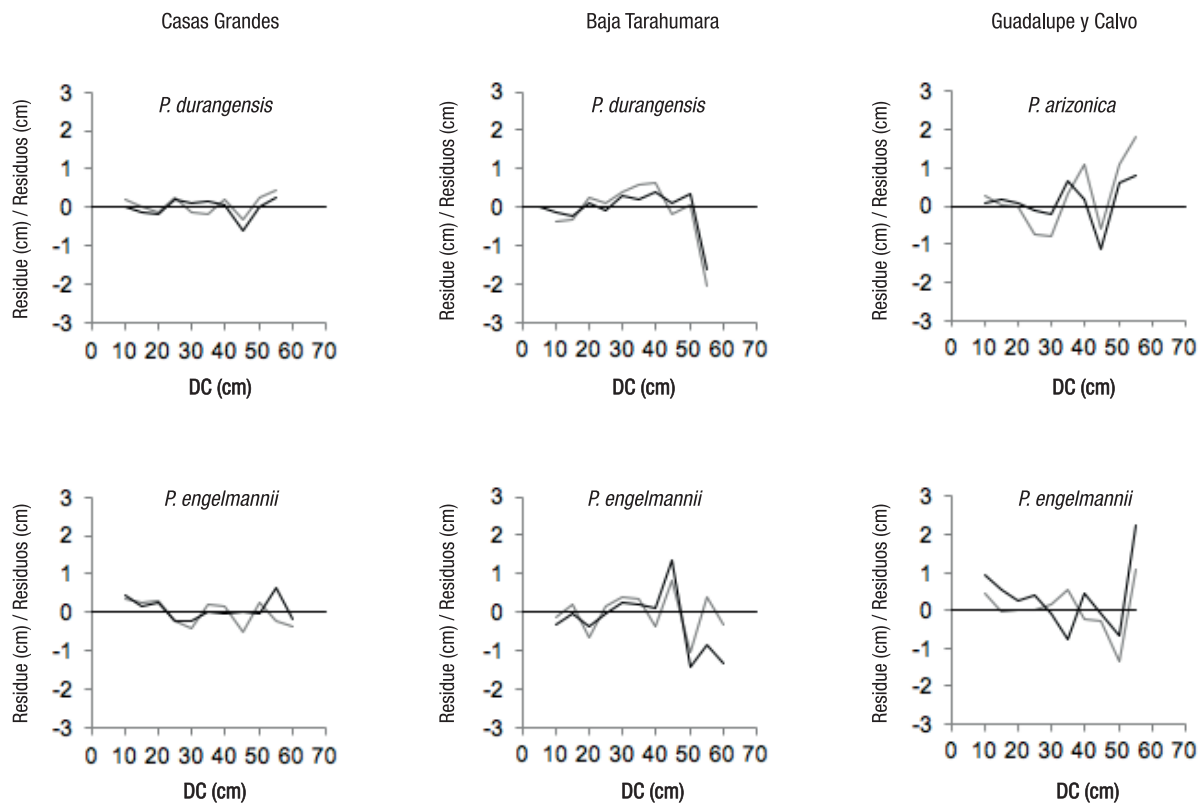


Figure 4. Evolution of bias by diametric class (DC) in the estimation of relative diameter (di/D) with data from the complete sample (black line) and the calculated minimum sample (gray line) for each *Pinus* species in three forest regions.

Figura 4. Evolución del sesgo por clase diamétrica (CD) en la estimación del diámetro relativo (di/D) con datos de la muestra completa (línea negra) y la muestra mínima calculada (línea gris) para cada especie de *Pinus*

Figure 4 compares the evolution of bias by diametric class (CD) in the estimate of di/D with the complete sample and the calculated minimum sample. In both cases, the trend was similar in all three species, although the bias of the entire sample was slightly closer to the zero line. The error trend was similar in the two fittings up to the CD of 30 cm; for the rest, the difference in bias was less than 0.15 % in the three species.

Figure 5 shows that the shape of the stem described by both equations is similar, except for *P. arizonica* in Guadalupe y Calvo, since the equation generated with the minimum sample seems to be more accurate to describe the stem. The observed differences were minimal and are related to the inflection points (p_1 and p_2) of the function by Fang et al. (2000). The values of the first inflection point (p_1) were similar in all cases (Table 4), generating forms similar to the lower part of the stem with both equations; on the contrary, the values of the second inflection point (p_2) showed greater differences, which is observed in the junctions of the tapers of some species.

Given the small differences in the tapers described with both equations, the effect of sample size on stem volume was evaluated. For this purpose, the observed volumes of the complete sample and those estimated with the equations developed with the complete sample and the calculated minimum sample were used. The results are shown in Table 5. Overall, estimates of the total stem volume with the complete sample were closer to the observed volume, except for *P. engelmannii* in Guadalupe y Calvo, where the estimated volume with the equation of the calculated minimum sample was closer to that observed, with a difference of less than 1 %.

Discussion

Most studies on stem taper modeling use variable sample sizes, without specifying the procedure used for selecting the number of trees used in the fit. For example, Novo, Rojo, and Álvarez (2003) adjusted 21 taper functions with data from 236 *P. sylvestris* L. trees, and Red et al. (2005) compared 31 taper functions with data from 203 *P. pinaster* Ait. trees in Galicia, Spain; Diéguez-Aranda et al. (2006) and Barrio-Anta, Diéguez-Aranda, Castedo-Dorado, Álvarez-González, and von Gadow (2007) fitted, respectively, taper systems to 228 *P. sylvestris* trees and 251 *Quercus robur* L. trees in northwestern Spain; Gezan, Moreno, and Ortega (2009) used 251, 180, and 127 trees for the development of stem models of *Nothofagus alpina* (Peopp et Endl.) Oerst., *N. obliqua* (Mirb.) Oerst. and *N. dombeyi* (Mirb.) Oerst., respectively, in Chile. Heidarrson and Pukkala (2011) also used samples of 87 and 89 trees to develop

entre 0.2 % (*P. arizonica* en Guadalupe y Calvo) y 0.9 % (*P. engelmannii* en Guadalupe y Calvo), con una diferencia promedio de 0.42 %.

Por otra parte, los valores de REMC para la función ajustada con la muestra completa y con la muestra mínima calculada también fueron similares (Cuadro 4). Los valores más bajos de REMC fueron 1.859 y 1.853 cm para la muestra completa y mínima, respectivamente, mientras que los más altos fueron 2.728 y 3.004 cm. El valor de REMC fue menor en cuatro (*P. durangensis* y *P. engelmannii* en Casas Grandes, *P. durangensis* en Baja Tarahumara y *P. arizonica* en Guadalupe y Calvo) de los seis casos con la muestra mínima calculada; es decir, la precisión del modelo fue mayor. El sesgo promedio fue ligeramente menor con la muestra completa en cuatro de los seis casos.

La Figura 4 compara la evolución del sesgo por categoría diamétrica (CD) en la estimación de di/D con la muestra completa y la muestra mínima calculada. En ambos casos, la tendencia fue similar en las tres especies, aunque el sesgo de la muestra completa estuvo ligeramente más cercano a la línea del cero. La tendencia de los errores fue similar en los dos ajustes hasta la CD de 30 cm; para el resto, la diferencia en el sesgo fue menor de 0.15 % en las tres especies.

En la Figura 5 se observa que la forma del fuste descrita por ambas ecuaciones es similar, excepto para *P. arizonica* en Guadalupe y Calvo, ya que la ecuación generada con la muestra mínima parece ser más precisa para describir el fuste. Las diferencias observadas fueron mínimas y están relacionadas con los puntos de inflexión (p_1 y p_2) de la función de Fang et al. (2000). Los valores del primer punto de inflexión (p_1) fueron similares en todos los casos (Cuadro 4), generando formas parecidas de la parte baja del fuste con ambas ecuaciones; por el contrario, los valores del segundo punto de inflexión (p_2) presentaron mayores diferencias, lo cual se observa en los cruces de los perfiles de algunas especies.

Dadas las pequeñas diferencias en los perfiles descritos con ambas ecuaciones, se evaluó el efecto del tamaño de la muestra en el volumen del fuste. Para ello se utilizaron los volúmenes observados de la muestra completa y los estimados con las ecuaciones desarrolladas con la muestra completa y la muestra mínima calculada. Los resultados se muestran en el Cuadro 5. En general, las estimaciones del volumen total del fuste con la muestra completa fueron más cercanas al volumen observado, excepto para *P. engelmannii* en Guadalupe y Calvo, donde el volumen estimado con la ecuación de la muestra mínima calculada fue más cercano al observado, con una diferencia menor de 1 %.

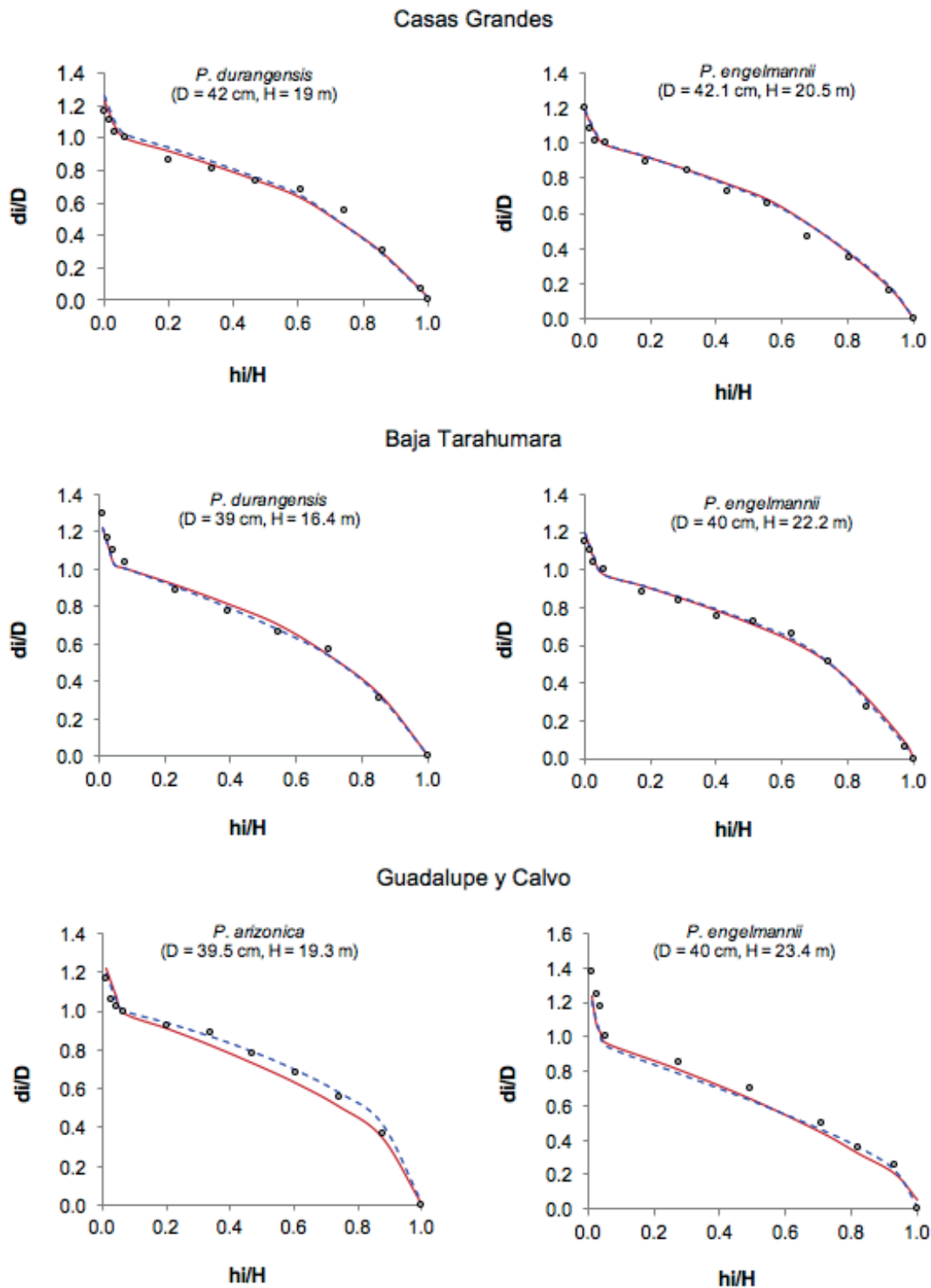


Figure 5. Stem taper of a tree by *Pinus* species and forest region, using the equation developed with the complete sample (red line) and the minimum sample calculated (blue line). The circles indicate the observed relative diameter (d_i/D) and height (h_i/H) values.

Figura 5. Perfil fustal de un árbol por especie de *Pinus* y región forestal, utilizando la ecuación desarrollada con la muestra completa (línea roja) y la muestra mínima calculada (línea azul). Los círculos indican los valores observados de diámetro (d_i/D) y altura (h_i/H) relativos.

Table 5. Comparison of the observed volume against the estimated volumes of three *Pinus* species with the equations generated with the complete sample and the calculated minimum sample.

Cuadro 5. Comparación del volumen observado contra los volúmenes estimados de tres especies de *Pinus* con las ecuaciones generadas con la muestra completa y la muestra mínima calculada.

Region/ Región	Species/ Especie	Observed volume (m ³)/Volumen observado (m ³)	Estimated volume (m ³)/Volumen estimado (m ³)			
			Complete sample/ Muestra completa	Difference (%)/ Diferencia (%)	Minimum sample/ Muestra mínima	Difference (%)/ Diferencia (%)
Casas Grandes	<i>P. durangensis</i>	118.98	118.89	0.08	118.67	0.27
	<i>P. engelmannii</i>	131.15	130.58	0.43	130.07	0.82
Baja Tarahumara	<i>P. durangensis</i>	104.73	103.76	1.60	102.93	1.70
	<i>P. engelmannii</i>	132.74	132.07	0.50	133.69	0.72
Guadalupe y Calvo	<i>P. arizonica</i>	101.49	100.39	1.08	102.59	1.09
	<i>P. engelmannii</i>	84.39	81.79	3.08	83.55	0.99

taper equations for *P. contorta* Dougl. and *Larix sibirica* Ledeb., respectively, in Iceland; and López, Barrios, and Trincado (2015) used 130 trees in the development of taper models for *Eucalyptus tereticornis* Smith plantations in Colombia.

The number of trees calculated in this study (53 to 88) is within the range reported in studies to estimate the optimal sample size for the development of taper functions. In this sense, Kitikidou and Chatzilazarou (2008) recommend, in a general way, the use of a minimum sample size of 77 trees for the development of such equations; however, similar research reports sizes even smaller than those calculated in the present study. For example, Cormier et al. (1992) indicate that a sample of 50 trees was sufficient for the satisfactory fit of taper functions for *P. taeda* L. in the southeast of the United States. Subedi, Sharma, and Parton (2011) state that 30 trees is a sample size suitable for parameterizing a taper equation; in this case that of Sharma and Zhang (2004) for *P. banksiana* Lamb. and *Picea mariana* Mill. Britton, Sterns & Poggenb. in northern Canada. Moreover, Saarinen et al. (2019) recommend at least 46 trees for the fit of taper equations of *P. sylvestris*, *Picea abies* (L.) Karst and *Betula pendula* (L.) and point out that a smaller number would affect the fit.

The goodness of fit statistics, obtained with the minimum estimated sample, are within the intervals reported by Silva-González, Nava-Moreno, Hernández, and Colín (2018) in fits of the function by Fang et al. (2000) for 111 trees of *P. arizonica* and 91 of *P. engelmannii* in the south of Chihuahua; REMC values varied between 1.82 and 2.61 cm, explaining 97 % of

Discusión

La mayoría de los estudios sobre modelización del perfil fustal utiliza tamaños variables de muestra, sin especificar el procedimiento utilizado para la selección del número de árboles empleado en el ajuste. Por ejemplo, Novo, Rojo, y Álvarez (2003) ajustaron 21 funciones de perfil con datos de 236 árboles de *P. sylvestris* L. y Rojo et al. (2005) compararon 31 funciones de perfil con datos de 203 árboles de *P. pinaster* Ait. en Galicia, España; Diéguez-Aranda et al. (2006) y Barrio-Anta, Diéguez-Aranda, Castedo-Dorado, Álvarez-González, y von Gadow (2007) ajustaron, respectivamente, sistemas de ahumamiento a 228 árboles de *P. sylvestris* y 251 árboles de *Quercus robur* L. en el noroeste de España; Gezan, Moreno, y Ortega (2009) usaron 251, 180 y 127 árboles para el desarrollo de modelos fustales de *Nothofagus alpina* (Peopp et Endl.) Oerst., *N. obliqua* (Mirb.) Oerst. y *N. dombeyi* (Mirb.) Oerst., respectivamente, en Chile. Asimismo, Heidarsson y Pukkala (2011) utilizaron muestras de 87 y 89 árboles para el desarrollo de ecuaciones de perfil para *P. contorta* Dougl. y *Larix sibirica* Ledeb., respectivamente, en Islandia; y López, Barrios, y Trincado (2015) usaron 130 árboles en el desarrollo de modelos de ahumamiento para plantaciones de *Eucalyptus tereticornis* Smith en Colombia.

El número de árboles calculado en este estudio (53 a 88) se ubica en el intervalo reportado en estudios de estimación del tamaño de muestra óptimo para el desarrollo de funciones de perfil. En este sentido, Kitikidou y Chatzilazarou (2008) recomiendan, de manera general, el uso de un tamaño mínimo de

the variance of diameter along the stem. Similarly, the goodness of fit statistics coincides with those reported by Corral et al. (1999) for sample sizes between 322 and 507 trees of five species in Durango, and with those obtained by López, Cruz, Nájera, and Hernández (2015) for sample sizes of 163 and 223 trees of *Pinus oocarpa* Schiede ex Schltdl. and *Pinus douglasiana* Martínez, respectively, in southwestern Durango; as well as with the results of García-Espinoza et al. (2019), who adjusted the function of Fang et al. (2000) to data from 165 trees of *P. pseudostrobus* in Michoacán.

On the other hand, the goodness of fit is superior to that reported in similar studies in the same region of study, such as the Pompa-García, Hernández, Prieto-Ruiz, and Davalos (2009) and Pompa-García, Corral-Rivas, Díaz-Vazquez, and Martínez-Salvador (2009), who fitted taper-volume compatible functions data of 663 trees of *P. durangensis* and 319 of *P. arizonica* in the southwest of Chihuahua. Finally, Vargas-Larreta et al. (2017) fitted the function by Fang et al. (2000) for 97 tree species of temperate and tropical forests in Mexico, with a fixed sample size of 150 trees, regardless of species or region. The R^2 values obtained by these authors were slightly higher than 0.97, but in some cases were lower than those obtained in this study with the minimum sample size. The differences observed in the goodness of fit are logical and could be explained by the influence of the growth conditions in the study areas, because the species is distributed in a mosaic of shapes, altitudes, temperature, and precipitation, which generate anatomical differences in the species (Bloomberg, Mason, Jarvis, & Sedcole, 2008).

The selection of sample size for the development of volume models, whether total or merchantable, should cover the greatest diversity of shapes and sizes (Torres & Magaña, 2001). In this regard, Cormier et al. (1992) note that, due to the high variability in tree shape, small sample sizes in the adjustment of profile equations increase the risk of obtaining large errors; however, in this study it was observed that the reduction in the number of trees, by means of the calculated minimum sample, does not generate gain in terms of accuracy of diameter estimates along the stem. Although R^2 values for the taper function with the complete sample were slightly higher than those obtained in the fit with the minimum estimated sample (Table 4), the difference was not statistically significant (between 0.001 and 0.009; 0.004 on average). This indicates that the equation developed with the calculated minimum sample explains the variability of the diameter along the stem equally accurately as the equation generated from the complete sample.

As was expected, the total volume is closer to that observed when using the equation from the complete

muestra de 77 árboles para el desarrollo de este tipo de ecuaciones; sin embargo, investigaciones similares reportan tamaños incluso menores que los calculados en el presente estudio. Por ejemplo, Cormier et al. (1992) señalan que una muestra de 50 árboles fue suficiente para el ajuste satisfactorio de funciones de perfil de *P. taeda* L. en el sureste de Estados Unidos. Subedi, Sharma, y Parton (2011) afirman que 30 árboles es un tamaño de muestra adecuado para la parametrización de una ecuación de ahusamiento; en este caso, la de Sharma y Zhang (2004) para *P. banksiana* Lamb. y *Picea mariana* Mill. Britton, Sterns & Poggenb. en el norte de Canadá. Por otra parte, Saarinen et al. (2019) recomiendan 46 árboles como mínimo para el ajuste de ecuaciones de ahusamiento de *P. sylvestris*, *Picea abies* (L.) Karst y *Betula pendula* (L.), y señalan que un número menor afectaría el ajuste.

Los estadísticos de bondad de ajuste, obtenidos con la muestra mínima estimada, se encuentran dentro de los intervalos reportados por Silva-González, Nava-Moreno, Hernández, y Colín (2018) en ajustes de la función de Fang et al. (2000) para 111 árboles de *P. arizonica* y 91 de *P. engelmannii* en el sur de Chihuahua; los valores de REMC variaron entre 1.82 y 2.61 cm, explicando 97 % de la varianza del diámetro a lo largo del fuste. De igual manera, los estadísticos de bondad de ajuste coinciden con los reportados por Corral et al. (1999) para tamaños de muestra entre 322 y 507 árboles de cinco especies en Durango, y con los obtenidos por López, Cruz, Nájera, y Hernández (2015) para tamaños de muestra de 163 y 223 árboles de *Pinus oocarpa* Schiede ex Schltdl. y *Pinus douglasiana* Martínez, respectivamente, en el suroeste de Durango; así como con los resultados de García-Espinoza et al. (2019), quienes ajustaron la función de Fang et al. (2000) a los datos de 165 árboles de *P. pseudostrobus* en Michoacán.

Por otra parte, la bondad de ajuste es superior a la reportada en estudios similares en la misma región de estudio, como los de Pompa-García, Hernández, Prieto-Ruiz, y Dávalos (2009) y Pompa-García, Corral-Rivas, Díaz-Vázquez, y Martínez-Salvador (2009), quienes ajustaron funciones compatibles de ahusamiento-volumen a datos de 663 árboles de *P. durangensis* y 319 de *P. arizonica* en el suroeste de Chihuahua. Finalmente, Vargas-Larreta et al. (2017) ajustaron la función de Fang et al. (2000) para 97 especies arbóreas de bosques templados y tropicales de México, con un tamaño de muestra fijo de 150 árboles, independientemente de la especie o región. Los valores de R^2 obtenidos por estos autores fueron ligeramente superiores a 0.97, pero en algunos casos fueron menores que los obtenidos en este estudio con el tamaño mínimo de muestra. Las diferencias observadas en la bondad de ajuste son lógicas y podrían ser explicadas por la influencia de las condiciones de crecimiento en las áreas de estudio,

sample, as the number of observations increases, the variance of the volumes tends to stabilize, therefore, the errors of the estimates are minimized; however, from the practical point of view, the differences in the volume calculated with both equations were small (1.1 %) and can be considered, for the purposes of this study, non-significant. These differences can mainly be due to the characteristics of the site and the forestry practices implemented, which define, to a large extent, the shape of the tree stem.

The equations developed with the complete sample and the calculated minimum sample yielded adequate estimates of the total volume of the tree lot (Table 5). An explanation for the similar behavior of the two equations is that they both describe the diametric structure of the population appropriately, indicating that the minimum tree number estimated for taper-volume function fit guarantees accurate estimates, as long as it represents the range of diameters and heights of the population. Otherwise, a sample with a concentration of small or large diameters could cause inherent bias in the developed model, starting from the minimum sample size required. In this study, the original sample was twice the minimum sample required in most cases, confirming that a smaller sample than those traditionally used can be used with the certainty that the volume estimates will be statistically the same.

Conclusions

The Fang function fitted to the estimated minimum sample size produced estimates of the stem volume as accurate as those obtained with the complete sample in all cases; that is, no significant differences were found in the accuracy of the function, so it is recommended to use the 10 % estimate, which reduces the sampling effort. The use of small samples in the adjustment of taper-volume functions allows accurate estimates to be obtained if the appropriate representation of the study population is ensured.

End of English version

References / Referencias

- Barrio-Anta, M., Diéguez-Aranda, U., Castedo-Dorado, F., Álvarez-González, J. G., & von Gadow, K. (2007). Merchantable volume system for pedunculate oak in northwestern Spain. *Annals of Forest Science*, 64(5), 511–520. doi: 10.1051/forest:2007028
- Bi, H. (2000). Trigonometric variable-form taper equations for Australian eucalyptus. *Forest Science*, 46(3), 397–409. doi: 10.1093/forestscience/46.3.397
- Bloomberg, M., Mason, E. G., Jarvis, P., & Sedcole, R. (2008). Predicting seedling biomass of radiate pine from allometric variables. *New Forests*, 36, 103–114. doi: 10.1007/s11056-008-9086-7
- debido a que las especies de distribuyen en un mosaico de formas, altitudes, temperatura y precipitación, que generan diferencias anatómicas en las especies (Bloomberg, Mason, Jarvis, & Sedcole, 2008).
- La selección del tamaño de muestra para la elaboración de modelos de volumen, ya sean totales o comerciales, debe cubrir la mayor diversidad de formas y tamaños (Torres & Magaña, 2001). En este sentido, Cormier et al. (1992) señalan que, debido a la alta variabilidad en la forma de los árboles, los tamaños de muestra pequeños en el ajuste de ecuaciones de perfil incrementan el riesgo de obtener errores grandes; sin embargo, en este estudio se observó que la reducción del número de árboles, mediante la muestra mínima calculada, no genera ganancia en términos de precisión de las estimaciones del diámetro a lo largo del fuste. A pesar de que los valores de R^2 para la función de ahusamiento con la muestra completa fueron ligeramente mayores que los obtenidos en el ajuste con la muestra mínima estimada (Cuadro 4), la diferencia no fue significativa desde un punto de vista estadístico (entre 0.001 y 0.009; 0.004 en promedio). Esto indica que la ecuación desarrollada con la muestra mínima calculada explica la variabilidad del diámetro a lo largo del fuste de manera igualmente precisa que la ecuación generada a partir de la muestra completa.
- Como era de esperarse, el volumen total es más cercano al observado cuando se utiliza la ecuación de la muestra completa, ya que conforme el número de observaciones aumenta, la varianza de los volúmenes tiende a estabilizarse y, por lo tanto, los errores de las estimaciones se minimizan; sin embargo, desde el punto de vista práctico, las diferencias en el volumen calculado con ambas ecuaciones fueron pequeñas (1.1 %) y pueden considerarse, para los fines de este estudio, no significativas. Estas diferencias pueden deberse fundamentalmente a las características del sitio y a las prácticas silvícolas implementadas, las cuales definen, en gran medida, la forma del fuste de los árboles.
- Las ecuaciones desarrolladas con la muestra completa y la muestra mínima calculada arrojaron estimaciones adecuadas de volumen total del lote de árboles (Cuadro 5). Una explicación del comportamiento similar de las dos ecuaciones es que ambas describen la estructura diamétrica de la población de manera adecuada, lo que indica que el número de árboles mínimo estimado para el ajuste de la función de ahusamiento-volumen garantiza estimaciones precisas, siempre y cuando represente el rango de diámetros y alturas de la población. En caso contrario, una muestra con una concentración de diámetros pequeños o grandes podría causar sesgo inherente en el modelo desarrollado, a partir del tamaño de muestra mínimo requerido. En este estudio, la muestra original fue el

- Burkhart, H. E., & Tomé, M. (2012). *Modeling forest trees and stands*. Dordrecht, The Netherlands: Springer.
- Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). (2014). *La biodiversidad en Chihuahua: Estudio de estado*. México: Author. Retrieved from <https://bpo.sep.gob.mx/#/recurso/498/document/1>
- Cormier, K. L., Reich, R. M., Czaplowski, R. L., & Bechtold, W. A. (1992). Evaluation of weighted regression and sample size in developing a taper model for loblolly pine. *Forest Ecology and Management*, 53, 65–76. doi: 10.1016/0378-1127(92)90034-7
- Corral, R. S., Návar, Ch. J., & Fernández, S. F. (1999). Ajustes de funciones de ahusamiento a los perfiles fustales de cinco pináceas de la región de El Salto, Durango. *Madera y Bosques*, 5(2), 53–65. doi: 10.21829/myb.1999.521347
- Corral-Rivas, J. J., Diéguez-Aranda, U., Corral-Rivas, S., & Castedo, D. F. (2007). A merchantable volume system for major pine species in El Salto, Durango (Mexico). *Forest Ecology and Management*, 238(1-3), 118–129. doi: 10.1016/j.foreco.2006.09.074
- Crecente-Campo, F., Rojo-Alboreca, A., & Diéguez-Aranda, U. A. (2009). A merchantable volume system for *Pinus sylvestris* L. in the major mountain ranges of Spain. *Annals of Forest Science*, 66(8), 808. doi: 10.1051/forest/2009078
- Demaerschalk, J. P., & Kozak, A. (1974). Suggestions and criteria for more effective regression sampling. *Canadian Journal of Forest Research*, 4(3), 341–348. doi: 10.1139/x74-051
- Diéguez-Aranda, U., Castedo-Dorado, F., Álvarez-González, J. G., & Rojo, A. (2006). Compatible taper function for Scots pine plantations in northwestern Spain. *Canadian Journal of Forest Research*, 36(5), 1190–1205. doi: 10.1139/x06-008
- Fang, Z., Borders, B. E., & Bailey, R. L. (2000). Compatible volume-taper models for Loblolly and Slash pine based on a system with segmented-stem form factors. *Forest Science*, 46(1), 1–12. doi: 10.1093/forests/46.1.1
- García-Espinoza, G. G., Aguirre-Calderón, O. A., Vargas-Larreta, B., Martínez-Ángel, L., García-Magaña, J. J., & Hernández-Ramos, J. (2019). Compatible taper and volume system for *Pinus pseudostrabus* Lindl. in Nuevo San Juan Parangaricutiro, Michoacan, Mexico. *Agrociencia*, 53(1), 115–131. Retrieved from https://www.researchgate.net/publication/331478269_Compatible_taper_and_volume_system_for_pinus_pseudostrabus_Lindl_In_Nuevo_San_Juan_Parangaricutiro_Michoacan_Mexico
- Gezan, S. A., Moreno, M. P., & Ortega, A. (2009). Modelos fustales para renovales de roble, raulí y coigüe en Chile. *Bosque*, 30(2), 61–69. doi: 10.4067/S0717-92002009000200001
- Heidarsson, L., & Pukkala, T. (2011). Taper functions for lodgepole pine (*Pinus contorta*) and siberian larch (*Larix sibirica*) in Iceland. *Icelandic Agricultural Sciences*, doble de la muestra mínima requerida en la mayoría de los casos, lo que confirma que una muestra menor que las usadas tradicionalmente puede utilizarse con la certeza de que las estimaciones de volumen serán estadísticamente iguales.

Conclusiones

La función de Fang ajustada con el tamaño de muestra mínimo estimado produjo estimaciones del volumen fustal tan precisas como las obtenidas con la muestra completa en todos los casos; es decir, no se encontraron diferencias significativas en la precisión de la función, por lo que se recomienda el uso de la estimación al 10 %, la cual aminora el esfuerzo de muestreo. El uso de muestras pequeñas en el ajuste de funciones de ahusamiento-volumen permite la obtención de estimaciones precisas si se garantiza la representación adecuada de la población objeto de estudio.

Fin de la versión en español

24, 3–11. Retrieved from https://ias.is/wp-content/uploads/Icelandic_Agricultural_Sciences_24_2011/Taper-functions-for-lodgepole-pine-Pinus-contorta.pdf

Jiménez, P. J., Aguirre, C. O., Treviño, G. E., & Domínguez, C. A. (1998). Desarrollo de un sistema matemático para la elaboración de tarifas volumétricas en especies arbóreas. *Madera y Bosques*, 4(2), 67–77. doi: 10.21829/myb.1998.421360

Kitikidou, K. (2010). An alternative sample size estimation for optimizing taper equations fitting. *Operational Research*, 10(2), 199–208. doi: 10.1007/s12351-009-0057-3

Kitikidou, K., & Chatzilazarou, G. (2008). Estimating the sample size for fitting taper equations. *Journal of Forest Science*, 54(4), 176–182. doi: 10.17221/789-JFS

Kozak, A. (1988). A variable-exponent taper equation. *Canadian Journal of Forest Research*, 18(11), 1363–1368. doi: 10.1139/x88-213

Lee, W. K., Seo, J. H., Son, Y. M., Lee, K. H., & von Gadow, K. (2003). Modelling stem profiles for *Pinus densiflora* in Korea. *Forest Ecology and Management*, 172, 69–77. doi: 10.1016/S0378-1127(02)00139-1

Levene, H. (1960). Robust tests for equality of variances. In I. Olkin (Ed.), *Contributions to probability and statistics* (pp. 278–292). Palo Alto, California, USA: Stanford University Press.

Li, R., & Weiskittel, A. R. (2010). Comparison of model forms for estimating stem taper and volume in the primary conifer species of the North American Acadian Region. *Annals of Forest Science*, 67, 302. doi: 10.1051/forest/2009109

- López, M. A., Barrios, A., & Trincado, G. (2015). Taper models with an autoregressive error structure for *Eucalyptus tereticornis* plantations in Colombia. *Madera y Bosques*, 21(2), 73–88. Retrieved from http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-04712015000200005
- López, M. J., Cruz, C. F., Nájera, L. J., & Hernández, F. J. (2015). Modelos de ahusamiento y volumen comercial para *Pinus oocarpa* y *P. douglasiana* en la región de Pueblo Nuevo, Durango. *Investigación y Ciencia de la Universidad Autónoma de Aguascalientes*, 64, 47–53. Retrieved from <https://investigacion.uaa.mx/RevistaIyC/archivo/revista64/Articulo%207.pdf>
- López-Martínez, J. O., Vargas-Larreta, B., Aguirre-Calderón, O., Aguirre-Calderón, C., Macario-Mendoza, P., Martínez-Salvador, M., & Álvarez-González, J. G. (2019). Compatible taper-volume systems for major tropical species in Mexico. *Forestry: An International Journal of Forest Research*, 93(1), 56–74. doi: 10.1093/forestry/cpz033
- Návar, J. J., & Domínguez, C. A. (1997). Ajuste de modelos de volumen y funciones que describen el perfil diamétrico de cuatro especies de pino plantadas en el nordeste de México. *Investigación Agraria: Sistemas y Recursos Forestales*, 6(1), 147–162. Retrieved from <https://recyt.fecyt.es/index.php/IA/article/view/4961>
- Navarro, J., Borja, A., Musálem, M. A., Ramírez, H., & Granados, D. (2002). Ecuaciones y tablas de volúmenes comerciales con y sin corteza para *Peltogyne mexicana* Martínez, una especie amenazada del estado de Guerrero, México. *Revista Chapingo Serie Ciencias Forestales y del Ambiente*, 8(2), 133–139. Retrieved from <https://www.redalyc.org/articulo.oa?id=62980207>
- Novo, N., Rojo, A., & Álvarez, G. J. (2003). Funciones de perfil del tronco para *Pinus sylvestris* L. en Galicia. *Investigación Agraria: Sistemas y Recursos Forestales*, 12(1), 123–136. Retrieved from <http://revistas.inia.es/index.php/fs/article/view/792/789>
- Pompa-García, M., Corral-Rivas, J. J., Díaz-Vázquez, M. A., & Martínez-Salvador, M. (2009). Función de ahusamiento y volumen compatible para *Pinus arizonica* Engelm. en el suroeste de Chihuahua. *Revista Ciencia Forestal en México*, 34(105), 119–136. Retrieved from <https://cienciasforestales.inifap.gob.mx/editorial/index.php/forestales/article/view/695/1778>
- Pompa-García, M., Hernández, C., Prieto-Ruiz, J. A., & Dávalos, S. R. (2009). Modelación del volumen fustal de *Pinus durangensis* en Guachochi, Chihuahua, México. *Madera y Bosques*, 15(1), 61–73. doi: 10.21829/myb.2009.1511197
- Rojo, A., Perales, X., Sánchez-Rodríguez, F., Álvarez-González, J. G., & von Gadow, K. (2005). Stem taper functions for maritime pine (*Pinus pinaster* Ait.) in Galicia (Northwestern Spain). *European Journal of Forest Research*, 124(3), 177–186. doi: 10.1007/s10342-005-0066-6
- Saarinen, N., Kankare, V., Pyörälä, J., Yrttimaa, T., Liang, X., Wulder, M. A., ...Vastaranta, M. (2019). Assessing the effects of sample size on parametrizing a taper curve equation and the resultant stem-volume estimates. *Forests*, 10(10), 848. doi: 10.3390/f10100848
- SAS Institute Inc. (2004). SAS/ETS 9.1 User's Guide. Cary, NC: SAS Institute Inc.
- Sharma, M., & Zhang, S. Y. (2004). Variable-exponent taper equations for jack pine, black spruce, and balsam fir in eastern Canada. *Forest Ecology and Management*, 198, 39–53. doi: 10.1016/j.foreco.2004.03.035
- Silva-González, E., Nava-Moreno, M. A., Hernández, F. J., & Colín, J. G. (2018). Funciones compatibles de ahusamiento-volumen para tres especies de *Pinus* en la Unidad de Manejo Forestal 0808 del estado de Chihuahua. *Investigación y Ciencia de la Universidad Autónoma de Aguascalientes*, 26(73), 58–67. Retrieved from <https://investigacion.uaa.mx/RevistaIyC/archivo/revista73/Articulo%207.pdf>
- Subedi, N., Sharma, M., & Parton, J. (2011). Effects of sample size and tree selection criteria on the performance of taper equations. *Scandinavian Journal of Forest Research*, 26, 555–567. doi: 10.1080/02827581.2011.583677
- Tamarit-Urías, J. C., De los Santos-Posadas, H. M., Aldrete, A., Valdez-Lazalde, J. R., Ramírez-Maldonado, H., & Guerra-De la Cruz, V. (2014). Sistema de cubicación para árboles individuales de *Tectona grandis* L. f. mediante funciones compatibles de ahusamiento-volumen. *Revista Mexicana de Ciencias Forestales*, 5(21), 58–75. doi: 10.29298/rmcf.v5i21.358
- Tapia, J., & Návar, J. (2011). Ajuste de modelos de volumen y funciones de ahusamiento para *Pinus pseudostrobus* Lindl. en bosques de la Sierra Madre Oriental de Nuevo León, México. *Foresta Veracruzana*, 13(2), 19–28. Retrieved from <https://www.redalyc.org/pdf/497/49721457004.pdf>
- Tlaxcala-Méndez, R. M., De los Santos-Posadas, H. M., Hernández-de la Rosa, P., & López-Ayala, J. L. (2015). Variación del factor de forma y el ahusamiento en procedencias de cedro rojo (*Cedrela odorata* L.). *Agrociencia*, 50(1), 89–105. Retrieved from <https://www.colpos.mx/agrocien/Bimestral/2016/ene-feb/art-7.pdf>
- Torres, R. J., & Magaña, T. O. (2001). *Evaluación de plantaciones forestales*. México: Noriega Editores.
- Uranga-Valencia, L. P., De los Santos-Posadas, H. M., Valdez-Lazalde, J. R., López-Upton, J., & Navarro-Garza, H. (2015). Volumen total y ahusamiento para *Pinus patula* Schiede ex Schltdl. et Cham. en tres condiciones de bosque. *Agrociencia*, 49(7), 787–801. Retrieved from <https://www.colpos.mx/agrocien/Bimestral/2015/oct-nov/art-7.pdf>
- Vargas-Larreta, B., Corral-Rivas, J. J., Aguirre-Calderón, O. A., López-Martínez, J. O., Santos-Posadas, H. M., Zamudio-Sánchez, F. J., & Aguirre-Calderón, C. G. (2017). SiBiFor: Sistema Biométrico Forestal para el manejo de los bosques de México. *Revista Chapingo Serie Ciencias Forestales y del Ambiente*, 23(3), 437–455. doi: 10.5154/r.rchscfa.2017.06.040