CIDE

NÚMERO 150

Juan Manuel Torres Rojo

USO DE LAS ECUACIONES SIMULTÁNEAS EN LA PREDICCIÓN DE PARÁMETROS DE DISTRIBUCIONES DIAMÉTRICAS

RESUMEN

La predicción de parámetros es un procedimiento ampliamente usado para predecir distribuciones de variables aleatorias a través de un conjunto de variables exógenas relacionadas con la distribución. Tal procedimiento de predicción por lo general, usa modelos lineales, cuyos ajustes generalmente son pobres y la mayoría de ellos viola algunos supuestos básicos del modelo de regresión lineal. Este artículo presenta una comparación entre las forma tradicional de ajuste de los modelos de predicción de parámetros y un procedimiento donde se corrigen los supuestos de normalidad de la variable dependiente, heteroscedasticidad, autocorrelación e independencia de los errores entre ecuaciones. Ambos procedimientos se comparan con distribuciones no usadas en el ajuste y procedentes de la misma población. Los resultados muestran que ambos procedimientos proporcionan estimadores significativamente diferentes y que los modelos de predicción corregidos, a pesar de tener menor bondad de ajuste brindan mejores predicciones. De aquí que las correcciones a violaciones de los supuestos básicos del modelo lineal mejoran las relaciones que deben conservar los parámetros para recuperar las funciones de densidad reales, que presenta el rodal.

Palabras clave: Predicción de parámetros, Distribución Weibull, Sistemas de Ecuaciones.

* Centro de Investigación y Docencia Económicas. Carr México-Toluca No. 3655 col. Lomas de Santa Fé, México D.F. 01210. E-mail: torresrj@dis1.cide.mx.

ABSTRACT

Parameter prediction is a widely used procedure to recover density functions of random variables from a set of exogenous variables related to the distribution. Such a prediction procedure uses lineal models, whose fits are generally poor and most of them violate some basic lineal regression model assumptions. This paper presents a comparison between the traditional way to fit parameter prediction models and a procedure where violations to the lineal regression model assumptions such as normality of the dependent variable, heteroscedasticity, autocorrelation and error independence of equations are corrected. Both procedures are compared with distributions not used to fit the models and from de same population. Results show that both procedures yield significantly different estimates and that the set of parameters estimated from the corrected prediction models yield better predictions, although they have lower goodness of fit than the traditionally estimated parameters. Hence corrections to the lineal model assumptions improve the relationships that the parameters must keep in order to recover the actual density function.

Key words: Parameter prediction, Weibull Distribution, Systems of equations.

INTRODUCCION

La técnica de predicción del crecimiento y rendimiento a través de distribuciones diamétricas se ha popularizado para modelar rodales regulares y uniespecíficos, aunque existen algunos ejemplos para rodales irregulares (Murphy y Farrar, 1988). En la generación de este tipo de modelos se han utilizado diferentes funciones de distribución de probabilidades (*fdp*) para caracterizar las distribuciones diamétricas de poblaciones forestales. Sin embargo, se ha generalizado el uso de la distribución Weibull (Bailey y Dell, 1973; Smalley y Bailey, 1974; Schreuder y Swank, 1974; Clutter y Belcher, 1978; Schreuder *et al.*, 1979; Bailey *et al.*, 1982) dada la amplia variedad de formas que puede modelar. Las aplicaciones en México del modelo Weibull para predecir el rendimiento actual y futuro incluyen los trabajos de Torres (1987) para *Pinus hartwegii*, Castillo (1988) para *P. patula*, Ramírez y Fierros (1989) para *P. caribaea* var. *hondurensis*, Acosta (1991) para *P. montezumae*, Torres *et al.* (1991) para *P. rudis* y Fuente (1998) para *P.rudis*.

Los primeros trabajos de esta clase se orientaron a modelar los estimadores de los parámetros de la *fdp* en forma explícita. El modelaje se hace generalmente a través de modelos lineales, usando como variables predictoras algunos atributos del rodal, tales como densidad, altura total, índice de sitio o edad. Tal procedimiento se ha denominado en forma general "predicción de parámetros". Las suposiciones de las relaciones lineales asumidas en estos modelos pueden tener algún fundamento biológico; sin embargo, la precisión de muchos resultados reportados en la bibliografía ha sido baja. Por ejemplo, Clutter y Belcher (1978) citan valores del coeficiente de determinación (R^2) de 0.107, 0.357 y 0.20, asociados con los modelos para predecir los parámetros de localización, escala y forma de la función Weibull; "a", "b" y "c", respectivamente. Como estos ejemplos se pueden citar un gran número de trabajos donde se muestran los bajos ajustes obtenidos en la predicción de estimadores (Smalley y Bailey, 1974; Clutter y Belcher, 1978; Dell *et al.* 1979; Feduccia *et al.* 1979). Existen ejemplos donde los ajustes son excelentes, sin embargo, el poder predictivo de tales estimadores es muy bajo (Cao y Burkhardt, 1984; Fuente, 1998)

Para solucionar este problema se desarrolló otra técnica de predicción, en la cual, en lugar de predecir directamente los estimadores, se predice alguna otra variable de mayor relación con las variables de estado del rodal, misma que mantiene una relación más directa con los parámetros de la distribución. De esta forma, surge la técnica conocida como "recuperación de parámetros", que como su nombre lo indica consiste en recobrar los estimadores de parámetros de la distribución, a partir de sus momentos no-centrales, momentos centrales, o bien, un conjunto de percentiles. Dada la mayor relación entre momentos (ó percentiles) y atributos del rodal, la técnica de recuperación de parámetros ha sido más socorrida para justificar mejores ajustes en los modelos usados (Borders y Patterson, 1990).

El procedimiento goza de una ventaja adicional, ya que al usar momentos en lugar de percentiles, las predicciones del área basal explícitamente calculada y el área basal recuperada de la distribución predicha resultan idénticas. Sin embargo, este procedimiento tiene dos desventajas importantes:

a) Al usar una función ligeramente truncada o con parámetros de localización adicionales, el conjunto de parámetros se debe recuperar por procedimientos iterativos, lo que en algunos casos puede resultar en una distribución totalmente diferente a la real y con muy malas predicciones (Borders y Patterson, 1990).

b) Cuando se aplica a masas naturales donde existen categorías sin frecuencias, o truncamientos, el procedimiento puede brindar estimaciones poco confiables (Borders y Patterson, 1990).

Algunos investigadores han enfatizado tales desventajas y comprueban que el procedimiento es muy ineficiente comparado con otros procedimientos de predicción de estructuras diamétricas (Borders y Patterson, 1990; Nepal y Somers, 1992, Vanclay, 1995).

Estas dos desventajas no son aplicables al procedimiento de predicción de parámetros mientras se cumplan las siguientes condiciones:

a) Que el conjunto de estimadores usado sea un conjunto de estimadores robustos.

b) Que se logren mantener las relaciones básicas entre los estimadores predichos a través de las relaciones lineales.

En este artículo se da seguimiento al segundo problema planteado anteriormente. Se propone un procedimiento para mejorar los ajustes entre variables de estado y estimadores de parámetros, mismo que también mantiene las relaciones básicas entre estos últimos. El artículo centra la atención en las relaciones que deben existir entre los parámetros de una distribución Weibull, así como los requisitos que debe cumplir el conjunto de ecuaciones predictoras. El artículo se ha organizado de la siguiente manera. En la segunda sección se presenta la metodología donde se detallan las propiedades que debe tener el conjunto de modelos de predicción de parámetros, así como la procedencia de los datos y el análisis. En la tercera sección muestran los resultados y finalmente, en la última sección se proporcionan las conclusiones.

METODOLOGIA.

Características del modelo

El procedimiento de estimación de parámetros no impone alguna restricción en el modelo que debe usarse para la predicción de cada estimador (Bailey y Dell 1973). Sin embargo, se sabe que existen relaciones entre los parámetros y entre éstos y algunas variables de la distribución que evidentemente se deben guardar. Así por ejemplo, es importante considerar que el parámetro "*a*" tiene una relación directa con el diámetro mínimo, por lo que este último es la mejor variable predictora de "*a*". Adicionalmente, el parámetro de localización tiene una relación muy conocida con el parámetro de escala, misma que tiene la forma:

$$b = X_{0.63} - a \qquad (1)$$

donde:

$X_{0.63}$: Valor del 63-avo percentil de la distribución acumulada.

Tal relación ha popularizado el uso de la variable combinada "$a + b$" para mejorar las predicciones de los parámetros de escala y localización (Smalley y Bailey, 1974; Feduccia *et al.*, 1979; Dell *et al.*, 1979) y evidentemente indica que ambas variables se determinan simultáneamente, esto es, son endógenas.

El parámetro de escala "*b*", muestra el desplazamiento de la distribución en el eje de las abscisas y como se señaló anteriormente, tiene relación estrecha con el parámetro de localización. Desde el punto de vista biológico, el valor del parámetro de escala depende en gran medida de la densidad relativa y la productividad del sitio, de aquí que variables que midan estas características de una población, tales como índice de sitio, diámetro cuadrático y medidas de densidad absoluta o relativa deben incluirse en los modelos de predicción para este parámetro.

Por su parte el parámetro de forma "*c*", también tiene una relación analítica con los parámetros de localización y escala. Sin embargo, tal relación no es lineal y es de la forma:

$$a = \mu_i' \, b\Gamma_i \qquad (2)$$

donde:

Γ_i :función gama de ($1 + i/c$)

μ_i' : i-*ésimo* momento no central

Como su nombre lo indica, el parámetro "*c*" refleja la forma de la distribución, misma que depende en gran medida de la edad y el nivel de competencia de la población. Varios autores en ecología de poblaciones han mostrado que las distribuciones con sesgo positivo (valores pequeños de *c*) son muy frecuentes en masas con alta competencia, mientras que en poblaciones sobre maduras son frecuentes las distribuciones con sesgo negativo (White y Harper, 1970). De aquí que la predicción de este parámetro debe considerar tanto la endogeneidad de los demás parámetros como variables que evalúen competencia.

El procedimiento original de predicción de parámetros únicamente considera que una relación lineal de buen ajuste es capaz de estimar eficientemente los tres parámetros y que una vez predichos, éstos pueden recobrar la distribución diamétrica en una forma aproximada. Sin embargo todo modelo lineal debe cumplir ciertos requisitos para que las predicciones o la inferencia sobre los parámetros del modelo sean apropiadas. En el caso particular del conjunto de ecuaciones utilizadas para predecir los parámetros, éstas deben cumplir con los siguientes requisitos:

Relación entre ecuaciones. Tal como se ha señalado, el procedimiento de predicción de parámetros consiste en estimar conjuntos de estimadores **u, v** y **w,** tal que permitan a través de un conjunto de variables de estado del rodal ($\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$) predecir los estimadores de los parámetros de la función Weibull. Esto es, un sistema de ecuaciones predictoras de la forma:

$$\begin{aligned} \mathbf{a} &= \mathbf{x}_1'\mathbf{u} + \mathbf{e}_1 \\ \mathbf{b} &= \mathbf{x}_2'\mathbf{v} + \mathbf{e}_2 \\ \mathbf{c} &= \mathbf{x}_3'\mathbf{w} + \mathbf{e}_3 \end{aligned} \qquad (3)$$

donde **a**, **b** y **c** representan los vectores de los estimadores de parámetros *a, b* y *c* respectivamente y $\mathbf{e}_i$ representa el vector de errores de la *i-ésima* ecuación. Evidentemente, si se considera que *a, b* y *c* están relacionados a través de las ecuaciones (1) y (2), entonces las variables dependientes son endógenas en cada ecuación y los conjuntos de errores $\mathbf{e}_i$ están correlacionados. De aquí que el conjunto de ecuaciones (3) debe estimarse a través de un procedimiento que considere la relación entre los errores de las tres ecuaciones y la endogeneidad de las variables dependientes, tal como los procedimientos de estimación de ecuaciones simultáneas.

Supuestos básicos: En cualquier estimación se busca que los estimadores cumplan con ciertas propiedades importantes tales como eficiencia, suficiencia y que sean insesgados. En el caso de los estimadores de los parámetros **u, v** y **w** (3), estas propiedades se evalúan a través de las características de los errores producidos. Si estos errores tienen media cero, tienen varianza uniforme (homoscedásticos) y no están correlacionados, entonces los estimadores tienen las propiedades deseables. Aún más, si los errores tienen una distribución en particular entonces es posible hacer inferencia sobre las características de los parámetros. Todo ello en conjunto permite que la predicción de valores sea más confiable.

Normalidad de la variable dependiente: Los estimadores calculados para cada distribución se pueden derivar de una amplia variedad de procedimientos; cada uno de los cuales reporta un conjunto de estimadores que aproximan la distribución diamétrica (Torres *et al.,* 1992). En poblaciones naturales con problemas de multimodalidad, truncamiento y la presencia de categorías vacías, frecuentemente el mejor conjunto de estimadores no es el analíticamente más eficiente. De aquí que la selección de un conjunto eficiente de estimadores requiere seleccionar dentro de una amplia variedad de éstos. Resulta evidente que los estimadores así obtenidos no tienen una distribución normal, ya sea porque se obtienen de diferentes procedimientos o porque se guardan relaciones no lineales (ecuaciones 1 y 2). Por esta razón resulta relevante que antes de iniciar cualquier ajuste se corrijan las variables dependientes por normalidad. La forma usual de corregir por ausencia de normalidad es a través de una transformación Box Cox de la forma:

$$\frac{\mathbf{y}^{\lambda}-1}{\lambda}=\mathbf{x'b}+\mathbf{e}$$

donde **b** es el vector de estimadores y **e** representa el vector de errores. En esta transformación se asume que existe un valor para λ tal que la variable dependiente transformada cumple con las siguientes características: se distribuye normalmente, es homoscedástica y tiene una relación lineal con los estimadores (Judge *et al.*, 1985).

Datos y análisis

La base de datos utilizada consistió de 147 observaciones provenientes de tres mediciones de 49 sitios de muestreo permanente (cuadrados de 50 x 50) ubicados en el Campo Experimental Forestal, San Juan Tetla, Puebla. Las características de los sitios y las variables medidas se detallan en Acosta (1991). Los datos se dividieron en dos conjuntos; el primer conjunto se integró por la primera y tercera medición para el cálculo de estimadores, mientras que el segundo conjunto se integró por la segunda medición y fue usado para la validación y comparación de las predicciones

El cálculo de estimadores de parámetros de la función Weibull se realizó con los diámetros de cada parcela. Los estimadores se calcularon con el sistema WEST (Magaña *et al.*, 1992) que permite calcular estimadores a través de 10 procedimientos diferentes y brinda los estadísticos Kolmogorov y Smirnov (*K-S*) y χ^2 para evaluar la bondad de ajuste de cada conjunto de estimadores. Adicionalmente, para cada parcela se calcularon las principales variables de estado tales como diámetros medio y cuadrático promedio, número de individuos, índice de sitio, edad e índice de densidad (Acosta, 1991) entre otras.

A fin de probar la calidad de las predicciones obtenidas de los modelos lineales de predicción de parámetros producidos, ya sea en forma tradicional o bien estimados con las correcciones arriba señaladas, se realizó lo siguiente. Para el ajuste tradicional se tomó el conjunto de estimadores más eficientes para cada parcela, esto es, aquel conjunto para el cual la prueba *K-S* arrojaba el valor más pequeño de desviación o bien el valor más alto de χ^2. Una vez seleccionado el mejor conjunto de estimadores de parámetros se corrieron varios modelos considerando diversas variables independientes que no solo tuvieran alguna relación con el parámetro, sino que mejoraran los ajustes. Los estimadores se calcularon por Cuadrados Mínimos Ordinarios usando el sistema SAS (Statistical Analysis System). Los modelos se evaluaron de acuerdo a su valor de R^2, el cuadrado medio del error (*CME*), el valor de *F* y la significancia y congruencia de los signos de los estimadores incluidos en el modelo.

Para el ajuste de los modelos con correcciones se realizó lo siguiente. Al igual que con los ajustes tradicionales, se seleccionó el conjunto de estimadores de parámetros de mejor ajuste para cada parcela siguiendo los criterios ya mencionados. Posteriormente, se definió un sistema de ecuaciones simultáneas que incluyera los tres parámetros y que además fuera una sistema perfectamente identificado: Esta propiedad se incluyó a fin de identificar un conjunto de restricciones en el sistema que obligase a una sola solución para el conjunto de estimadores. Adicionalmente, se requirió que el sistema incluyera solo variables que reflejaran las características básicas de la distribución diamétrica como diámetro medio y diámetro cuadrático. La especificación usada para el sistema de ecuaciones es:

$$\begin{aligned} a &= \alpha_1 b + \alpha_2 D + \alpha_3 Dq + e_1 \\ b &= \beta_0 + \beta_1 a + \beta_2 D + e_2 \\ c &= \gamma_1 a + \gamma_2 b + \gamma_3 D + e_3 \end{aligned} \qquad (4)$$

El sistema (4) es un sistema perfectamente identificado con tres variables endógenas (*a, b* y *c*), tres variables exógenas (*constante, D* y *Dq*) y dos restricciones en cada ecuación. Además se cuidó que el sistema solo incluyera variables exógenas que identifican a la distribución, en este caso los dos primeros momentos no centrales (*D* y *Dq*).

Una vez definido el sistema de ecuaciones se procedió a corregir por normalidad de la variable dependiente en cada ecuación. Para ello se utilizó el procedimiento de transformación Box-Cox (Box y Cox, 1964). Una vez que cada variable dependiente fue corregida, se procedió a realizar el ajuste de todo el sistema en forma simultánea, para lo cual se utilizó el procedimiento de Cuadrados Mínimos de Tres Etapas (CM3E) del sistema SAS.

La comparación se realizó usando el conjunto de datos reservado para validación. Este conjunto corresponde a la segunda medición de una misma población, por lo cual se esperarían diferencias mínimas. Para ello se calcularon los estadísticos básicos de cada parcela, posteriormente se calcularon los valores de $\hat{a}, \hat{b}$ y $\hat{c}$ con los modelos lineales tradicionales y los modelos corregidos. En seguida se procedió a evaluar tales estimadores con respecto a la distribución original usando los estadísticos *K-S*, χ^2 y el índice de error de Reynolds *et al.*, (1988). Adicionalmente se compararon las predicciones de área basal obteniendo el valor absoluto de la diferencia entre área basal predicha y observada $AB_O - AB_P$, así como el valor absoluto de la diferencia entre el diámetro promedio predicho menos el observado $D_O - D_P$.

El índice de error de Reynolds *et al.*, (1988) se define como una suma ponderada de la diferencias absolutas entre el número de árboles predicho y el observado en cada categoría diamétrica. La ponderación puede hacerse con volumen o valor, aunque los autores establecen que puede usarse cualquier otro ponderador. Dado que el objetivo de este trabajo es evaluar el poder predictivo de los dos diferentes sistemas de ecuaciones se uso como ponderador el área basal de cada categoría, de tal forma que el indice de Reynolds *et al.*, (1988) modificado tiene la forma:

$$IER = \sum_{i-mC}^{MC} \frac{n_i^P - n_i^O}{ab_i^O}$$

donde:

IER= Indice de Error de Reynolds *et al.*, (1988) modificado.

n_i^P = No. De árboles predicho en la *i-ésima* categoría diamétrica

n_i^O = No. De árboles observado en la *i-ésima* categoría diamétrica.

ab_i^O = Area basal observada en la *i-ésima* categoría diamétrica.

MC= Categoría diamétrica más alta.

mC= Categoría diamétrica más pequeña

RESULTADOS

Los modelos lineales ajustados en forma tradicional brindaron los siguientes estimadores:

$$\hat{a} = -0.888026\, b + 0.793927\, D + 0.186722\, Dq$$
$$(0.0072) \quad (0.0265) \quad (0.0244)$$
$$(^{**}) \quad (^{**}) \quad (^{**})$$
$$R^2=0.993,\ CME=0.2013,\ F=5757.67,\ DW=1.972$$

$$\hat{b} = 0.000345 + 1.073696\, D$$
$$(0.0001) \quad (0.0107)$$
$$(^{**}) \quad (^{**})$$
$$R^2=0.987,\ CME=31.1931,\ F=10059,\ DW=1.883$$

$$ln(\hat{c}) = 0.111187\, ln(D) + 0.018238\, AB$$
$$(0.0372) \quad (0.0041)$$
$$(^{**}) \quad (^{**})$$
$$R^2=0.936,\ CME=0.075,\ F=908.2,\ DW=2.074$$

donde:

AB = Area basal por hectárera (m^2).

ln(.)= Logaritmo natural de (.).

y las demás variables siguen la misma nomenclatura. Por su parte los valores en paréntesis representan los errores estándar de los estimadores, R^2 representa el coeficiente de determinación, *F* el valor de F, *DW* el estadístico Durbin –Watson y *CME* la varianza del modelo.

La transformación Box-Cox sobre cada uno de los modelos en el sistema (4) indicó que la única variable que requiere transformación es la ecuación del parámetro *c*. El intervalo de confianza al 95% para λ asociada a la transformación del parámetro *c* es: $-0.22 \le \lambda \le 0.05$, de aquí que se eligió la transformación *ln(c)* como la transformación de máxima verosimilitud. De esta forma el sistema de ecuaciones corregido por normalidad de la variable dependiente es:

$$a = \alpha_1 b + \alpha_2 D + \alpha_3 Dq + e_1$$
$$b = \beta_0 + \beta_1 a + \beta_2 D + e_2$$
$$ln(c) = \gamma_1 a + \gamma_2 b + \gamma_3 D + e_3$$

La estimación por Cuadrados Mínimos de Tres Etapas rindió los siguientes estimadores:

$$\hat{a} = -0.822744\, b + 0.745055\, D + 0.166071\, Dq$$
$$(0.06210) \quad (0.0566) \quad (0.0370)$$
$$(^{**}) \quad (^{**}) \quad (^{**})$$

$$R^2=0.922,\ CME=0.3352,\ F=485.25$$

$$\hat{b} = 2.781662 - 1.542264\ a + 1.075373\ D$$
$$(1.6626)\quad (0.306451)\quad (0.032407)$$
$$(NS)\quad (^{**})\quad (^{**})$$
$$R^2=0.905,\ CME=4.719,\ F=585.75$$

$$ln(\hat{c}) = -0.515771\,a - 0.493622\,b + 0.573069\,D$$
$$(0.155125)\quad (0.110901)\quad (0.124886)$$
$$(^{**})\quad (^{**})\quad (^{**})$$
$$R^2=0.879,\ CME=0.162,\ F=295.71$$

Con una matriz de varianza-covarianza Σ^{-1} entre ecuaciones de la forma:

$$\Sigma^{-1} = \begin{bmatrix} 0.3352 & 1.0685 & -0.0210 \\ 1.0685 & 4.7195 & -0.5217 \\ -0.0210 & -0.5217 & 0.1624 \end{bmatrix}$$

Observe que el modelo para predecir el parámetro a es similar tanto para la forma tradicional como para la forma corregida. Visto individualmente, el ajuste tradicional tiene menor varianza y estimadores más eficientes, sin embargo al corregir por la correlación entre los errores de ecuaciones se reduce la eficiencia entre estos modelos (comparados individualmente). Lo mismo sucede con las ecuaciones para estimar los parámetros b y c. A pesar de la reducción de eficiencia de cada ecuación individual, los estimadores lineales de los parámetros de la distribución Weibull mostraron mejor poder predictivo en la prueba de validación tal y como lo muestra el Cuadro 1 de estadísticos comparativos.

Presumiblemente el resultado se debe a que la estimación a través de un sistema de ecuaciones lineales (corregidas por normalidad y heteroscedasticidad) obliga a que los estimadores de parámetros conserven las relaciones que deben guardar (ecuaciones 1 y 2) y se restrinjan predicciones inconsistentes con las relaciones entre estos parámetros. Lo anterior resulta en distribuciones recuperadas más aproximadas a las distribuciones reales, mejorando con ello la calidad del modelo.

Resulta de gran interés confirmar que en muchos casos el conjunto de parámetros estimados fue totalmente diferente, sin embargo la distribución predicha por ambos conjuntos no fue significativamente diferente. El Cuadro 2 muestra las diferencias absolutas promedio entre los estimadores de parámetros de la distribución Weibull predichos con los dos procedimientos (tradicional y corregido). Como se puede apreciar, no existe evidencia empírica para rechazar la hipótesis de igualdad de valores entre ambos procedimientos y para todos los parámetros. Es de particular interés observar las diferencias en el parámetro a, dado que los modelos usados con ambos procedimientos son iguales y a pesar de ello los parámetros estimados son significativamente diferentes.

Cuadro 1. Promedio de estadísticos y criterios de comparación de las distribuciones predichas para la segunda medición

Estadístico / criterio de comparación	Cálculo Tradicional	Cálculo corregido
K-S promedio	0.4532	0.2893
χ^2 promedio	108.2152	196.3157
IER promedio	20.35	18.79
$D_O - D_P$ promedio	3.2804	1.8305
$\lvert \overline{AB}_O - \overline{AB}_P \rvert$ promedio	0.3876	0.1035
No. De distribuciones no significativas al 0.1 †	49	10
No. De distribuciones no significativas al 0.5 †	43	2
n ††	49	49

† Distribuciones que no fueron significativas con la prueba K-S al .10 y 0.05 de nivel de significancia.
†† Tamaño de la muestra.

Este resultado señala que es de suma importancia considerar en el ajuste de modelos de predicción dos aspectos básicos: i) El parámetro c no guarda una relación lineal con los momentos o con los parámetros a y b. ii) Los modelos deben considerar las relaciones entre parámetros, esto es, existe un problema de endogeneidad en el sistema de ecuaciones de predicción de parámetros. iii) El ajuste de los modelos debe realizarse considerando la endogeneidad de las variables y la alta relación entre los errores de las ecuaciones predictoras. Evidentemente en esta última observación debe considerarse un modelo que permita recuperar los parámetros dada una proyección. Así por ejemplo, el sistema de ecuaciones en (4) tiene que resolverse simultáneamente para identificar ecuaciones que permitan calcular un determinado parámetro a partir exclusivamente de D y Dq.

Cuadro 2. Diferencias entre estimadores de parámetros calculados con el método tradicional y las correcciones.

Estimador del parámetro	Diferencia absoluta promedio	Valor de t [†]	Prob.> t
a	1.3058	0.37	0.7155
b	1.9394	1.01	0.3134
c	0.0850	0.89	0.3751

[†] Estadístico de la prueba de la Hipótesis H_0: $\delta_c = \delta_t$ donde δ_t es el parámetro estimado con el método corregido y δ_t es el parámetro estimado con el método tradicional.

CONCLUSIONES

El procedimiento más común de predicción de estructuras diamétricas en México es el de recuperación de funciones de densidad con el procedimiento de predicción de parámetros. Este trabajo ha mostrado que la forma tradicional de ajustar las ecuaciones de predicción podría predecir distribuciones sesgadas si no se consideran tanto las relaciones básicas entre los parámetros, como la linealidad de las relaciones entre los parámetros y las variables exógenas del modelo. Adicionalmente se ha mostrado que el procedimiento de ajuste no debe ser el tradicional, sino uno que incluya la relación entre los errores de las diferentes ecuaciones. Al considerarse tales recomendaciones en la formulación de modelos como en los ajustes, se obtienen distribuciones proyectadas significativamente más aproximadas a las reales.

BIBLIOGRAFIA

Acosta M., M. 1991. Modelo de crecimiento para Pinus montezumae Lamb. en el CEF San Juan Tetla, Puebla. Tesis de Maestría en Ciencias Forestales. División de Ciencias Forestales, Universidad Autónoma Chapingo. Chapingo, México. 88 p.

Bailey, R. L. y Dell, T. R. 1973. Quantifying diameter distributions with the Weibull function. For. Sci. 19:97-104.

Bailey, R. L., N. C. Abernethy y E. P. Jones. 1982. Diameter distribution models for repeatedly thinned slash pine plantations. USDA For. Serv. Gen. Tech. Rep. SO-34. 53 p.

Borders, B.E. y W.D. Patterson. 1990. Projecting stand tables: A comparison of the Weibull diameter distribution method, a percentile-based projection method , and a basal area growth projection method. For. Sci. 36:413.424.

Box, G.E.P. y D.R. Cox. 1964. An analysis of transformations. Journal of the Royal Statistical Society. 26:211-243

Cao, Q.V. y H.E. Burkhart. 1984. A segmented distribution approach for modeling diameter frequency data. For. Sci. 30:129-137.

Castillo S., M.A. 1988. Modelo para estimación de incremento y producción maderable neta en Pinus caribaea var. hondurensis Barr. y Golf., de la Sabana, Oax. Tesis. . División de Ciencias Forestales, Universidad Autónoma Chapingo. Chapingo, México. 81 p.

Clutter, J.L. y D.M. Belcher. 1978. Yield of site prepared slash pine plantations in lower coastal plain of Georgia and Florida. *In* Growth for long term forecasting for timber yields. Div. of For. and Wildlife Res. V.P.I. Rep. FWS-1-78.

Dell, T.R., D.P. Feduccia, T.E. Campbell, W.F. Mann y B.H. Polmer. 1979. Yields of unthinned slash pine plantations on cutover sites in the west gulf region. USDA. For. Serv. Res. Pap. SO-147.

Feduccia, D.P. T.R. Dell, W.F. Mann, T.E. Campbell y B.H. Polmer. 1979. Yields of unthinned loblolly pine plantations on cutover sites in the west gulf region. USDA For. Serv. Res. Pap. SO-148.

Fuente E., A de la. 1998. Crecimiento y predicciones del rendimiento de rodales coetáneos de Pinus rudis Endl. de Pueblos mancomunados, Ixtlán, Oax. Tesis de Doctorado. Colegio de Postgraduados. Montecillo, México 175p.

Judge, G.G., W.E. Griffiths, R. C. Hill, H. Lütkepohl y T-Ch Lee. 1985. The theory and practice of Econometrics. 2nd Ed. John Wiley and Sons. 1019 p.

Magaña T., O.S., J.M. Torres R. y M. Acosta M. 1992. WEST: Programa para estimar los parámetros de la función de distribución de probabilidades Weibull. Documento de trabajo, CEVAMEX, INIFAP.

Murphy, P.A. y R.M. Farrar. 1988. A framework for stand structure projection of uneven-aged lololly-shortleaf pine stands. For. Sci. 34:321-332.

Nepal , S.K. y G.L. Somers. 1992. A generalized approach to stand table projection. For. Sci. 38:120-133.

Ramírez M., H. y A.M. Fierros G. 1989. Estimación del rendimiento y crecimiento de *Pinus caribaea var. hondurensis* a través de su distribución diamétrica. In Salazar, R. (Dc.) Cuarta Reunión del Grupo de Trabajo de IUFRO. Silvicultura de los Neotrópicos. Memoria. Antigua, Guatemala. 3 - 7 de abril de 1989. CATIE, Turrialba, Costa Rica. pp. 459 - 474.

Reynolds, M.R. Jr., T.E. Burke y W. Huang. 1988. Goodness-of-tests and model selection procedures for diameter distribution models. For. Sci.34:373-379.

Shreuder, H.T. y W.T. Swank. 1974. Coniferous stands characterized with the Weibull distribution. Can. J. For. Res. 4(3):518-523.

Shreuder, H.T., W.L. Hafley y F.A. Bennett. 1979. Yield prediction for unthinned natural slash pine stands. For. Sci. 25(1):25-30.

Smalley, G. W. y Bailey, R. L. 1974. Yield Tables and stand structure for loblolly pine plantations in Tennessee, Alabama and Georgia highlands. USDA. For. Serv. Res. Pap. SO-96. 81 p.

Torres R., J. M. 1987. Economic analysis of several alternatives of forest management for Pinus hartwegii. Master Thesis. Oregon State University. 126 p.

Torres R., J.M., C. Rodríguez F., O.S. Magaña T., H. Aguirre D. y A.M. Fierros G. 1991. Predicción de la producción de *Pinus rudis* en Aloapan, Oaxaca. In. Memorias del Taller Internacional sobre Modelos Forestales. Instituto Forestal. Santiago de Chile 5-7 de Marzo de 1991.

Torres R., J.M., M. Acosta M. y O.S. Magaña T. 1992. Métodos para estimar los parámetros de la función Weibull y su potencial para ser predichos a través de atributos del rodal. Agrociencia, Serie Rec. Nat. Renovables 2:57-76.

Vanclay, Jerome. 1995. Growth models for tropical forests: A synthesis of models and methods. For. Sci. 41(1): 7-42.

White, J. y J.L. Harper. 1970. Correlated changes in plant size and number in plant populations. J. Ecol. 58(2):467-85.

AGRADECIMIENTOS

El autor agradece la colaboración del Dr. Octavio Magaña Torres en la estimación y validación de modelos, así como al M.C. Miguel Acosta Mireles por haber proporcionado los datos de las parcelas experimentales usadas en el análisis.