

# METODOLOGÍA PARA LA CONSTRUCCIÓN DE CURVAS DE CALIDAD DE ESTACIÓN

J. G. Álvarez González\*, M. Barrio Anta, U. Diéguez Aranda y A. Rojo Alboreca

Departamento de Ingeniería Agroforestal. Escuela Politécnica Superior. Campus Universitario s/n 27002-LUGO (España). \* Correo electrónico: [algonjg@lugo.usc.es](mailto:algonjg@lugo.usc.es)

## Resumen

Se presenta una metodología de construcción de curvas de calidad de estación basada en datos procedentes del análisis de tronco de árboles dominantes. Este procedimiento consiste en el empleo de una estructura de datos transformada que incluye todos los posibles intervalos de crecimiento y el uso de ecuaciones de crecimiento expresadas en diferencias algebraicas. El objetivo es obtener familias de curvas polimórficas, invariantes en la edad de referencia y que permitan despejar de la ecuación ajustada el índice de sitio. Debido a la estructura longitudinal de los datos empleados, se corrige la dependencia de los mismos considerando la estructura del error como un proceso autorregresivo. Por último, se presenta un método para seleccionar la edad de referencia más adecuada y se incide sobre la importancia de las comparaciones gráficas en la elección del modelo definitivo.

Palabras clave: *Altura dominante, Índice de sitio, Ecuaciones en diferencias algebraicas, Modelos autorregresivos*

## INTRODUCCIÓN

La calidad de estación se refiere a la capacidad de un lugar determinado para el crecimiento de los árboles y/o cualquier otra vegetación, denominándose también “productividad forestal”. Las estimaciones precisas de esta productividad son un elemento clave en la gestión forestal, ya que permiten determinar la posibilidad anual y el periodo de rotación de las cortas, además de ayudar a la toma de decisiones acerca de elección de las especies más adecuadas para una localización concreta. Idealmente, dicha productividad debería cuantificarse directamente al final del ciclo productivo en términos de volumen de madera; sin embargo, esta aproximación no es factible debido a las largas rotaciones de las masas forestales, y a que sería necesario conocer tanto los volúmenes extraídos en cortas intermedias como los perdidos por causas naturales. Por tanto, es necesario utilizar indicadores de la pro-

ductividad potencial. Para ello, se pueden considerar tanto factores intrínsecos (evolución con la edad de alguna variable de masa) como factores extrínsecos a la masa (clima, suelo, vegetación, etc.). Los primeros han proporcionado generalmente mejores resultados, y dentro de ellos, la altura dominante se ha mostrado como el más adecuado por su baja dependencia de la densidad de la masa y de los tratamientos selvícolas habituales, así como por estar estrechamente relacionada con la producción total en volumen. Por tal motivo, el potencial de crecimiento o calidad de estación se suele expresar mediante el índice de sitio, que es el valor de la altura dominante de la masa a una determinada edad de referencia.

## MÉTODOS

Existe un gran número de funciones de crecimiento, aunque no todas ellas cumplen los

requisitos deseables en una función de este tipo (GOELTZ & BURK, 1992): polimorfismo, ser creciente y con punto de inflexión, existencia de asíntota horizontal a edades avanzadas, comportamiento lógico, base teórica, ser invariante en edad de referencia, y ser sencilla, es decir, con el menor número de parámetros posible. Actualmente, los modelos altura-edad expresados en forma de diferencias algebraicas (EDA) son la forma más utilizada para ajustar las ecuaciones de crecimiento en altura (CLUTTER *et al.*, 1983; GOELTZ & BURK, 1992; PARRÉSOL & VISSAGE, 1998), ya que garantizan el cumplimiento de gran parte de los requisitos exigidos a estas funciones. El procedimiento de ecuaciones EDA consiste en despejar un parámetro en el modelo de crecimiento y expresarlo en función de la altura  $H_1$  a la edad inicial  $t_1$  y el resto de parámetros; realizar la misma operación para una altura  $H_2$  a la edad  $t_2$ ; igualar ambas expresiones; y despejar  $H_2$ , de modo que quede expresada en función de  $t_1$ ,  $H_1$ ,  $t_2$  y del resto de parámetros. La elección del parámetro a eliminar determina el comportamiento del modelo; así, la sustitución del parámetro de la asíntota produce curvas anamórficas, y la sustitución de cualquier otro parámetro produce curvas polimórficas con igual o diferentes asíntotas.

Para la construcción de ecuaciones de índice de sitio se emplean habitualmente datos longitudinales procedentes del análisis de tronco de árboles dominantes, que permiten reconstruir las curvas de evolución de la altura con la edad. Su empleo, no obstante, plantea tres problemas: (1) la determinación de la altura verdadera a la edad calculada por el conteo de anillos en cada sección, (2) la elección de la estructura más adecuada de los datos originales, y (3) la dependencia de los errores inherente a este tipo de datos.

Existen numerosos métodos para la corrección del sesgo introducido al suponer que la sección de corte coincide con la finalización del crecimiento en ese período vegetativo. No obstante, DYER & BAILEY (1987), en un estudio comparativo de seis de esos métodos, concluyeron que el empleo del algoritmo de CARMEAN (1972) con la posterior corrección de NEWBERRY (1991) para la troza final era el método más preciso para eliminar el sesgo.

A partir de los datos originales procedentes del análisis de tronco se pueden generar diferentes estructuras de datos para su uso en el ajuste de modelos expresados como ecuaciones en diferencias algebraicas. De entre ellas, la que considera todos los posibles intervalos de crecimiento, incluyendo los descendentes, es la que proporciona resultados más estables y consistentes, optimizando la capacidad predictiva del modelo (GOELTZ & BURK, 1992; HUANG & TITUS, 1993; HUANG, 1999). No obstante, el empleo de esta estructura lleva a rechazar las hipótesis del error, debido a que introduce automáticamente una falta de independencia entre las observaciones (PARRÉSOL & VISSAGE, 1998), aunque las estimaciones de los parámetros son insensibles a la estructura del error y sólo su error estándar, es decir, el intervalo de confianza de las estimaciones de los parámetros, puede verse afectado (GOELTZ & BURK, 1992).

El potencial problema de autocorrelación de los datos se puede solucionar utilizando mínimos cuadrados generalizados no lineales (MONSERUD, 1984; GOELTZ & BURK, 1992; HUANG, 1997), y expandiendo el término del error mediante un modelo autorregresivo de la siguiente manera (GOELTZ & BURK, 1992; PARRÉSOL & VISSAGE, 1998):

$$H_{ij} = f(H_j, t_i, t_j, \beta) + e_{ij} \quad \text{con} \quad [1]$$

$$e_{ij} = \rho e_{i-1,j} + \gamma e_{i,j-1} + \varepsilon_{ij}$$

donde  $H_{ij}$  representa la predicción de la altura  $i$  utilizando  $H_j$  (altura  $j$ ),  $t_i$  (edad  $i$ ), y  $t_j$  (edad  $j \neq i$ ) como variables predictoras;  $\beta$  es el vector de parámetros a estimar;  $e_{ij}$  es el correspondiente término del error; el parámetro  $\rho$  tiene en cuenta la autocorrelación entre el residuo actual y el residuo obtenido al estimar  $H_{i-1}$  utilizando  $H_j$  como predictora; el parámetro  $\gamma$  tiene en cuenta la autocorrelación entre el residuo actual y el residuo obtenido al estimar  $H_i$  utilizando  $H_{j-1}$  como predictora; y  $\varepsilon_{ij}$  son los errores independientes y homogéneamente distribuidos.

El ajuste simultáneo de la estructura de la media (dada por la ecuación de crecimiento) y de la estructura del error (dada por el modelo autorregresivo) se puede realizar empleando el procedimiento MODEL del paquete estadístico SAS/ETS® (SAS

INSTITUTE INC., 2000), que permite una actualización dinámica de los residuos.

Debe tenerse en cuenta que en la expresión definitiva de los modelos generalmente se ignoran los valores de los parámetros  $\rho$  y  $\gamma$  de la estructura de autocorrelación del error, debido a que el principal propósito de su uso es obtener unas estimaciones de los parámetros  $\beta$  del modelo de la media insesgados y más eficientes; además, en la práctica no se conocen los errores  $e_{i-1}$  o  $e_{i-2}$  sin un análisis de tronco (HUANG, 1999; PARRÉSOL & VISSAGE, 1998).

El análisis de la capacidad de ajuste de los modelos debe basarse en comparaciones numéricas y gráficas. Así, a partir de los residuos obtenidos en la fase de ajuste, pueden calcularse, entre otros, los siguientes estadísticos: sesgo, coeficiente de determinación ajustado y criterio de información de Akaike. Con respecto a las comparaciones gráficas, su empleo es necesario para comprobar que las curvas se ajustan a los datos en todo su rango; además, diferentes modelos pueden presentar los mismos estadísticos de comparación pero un comportamiento distinto. Entre los diferentes métodos gráficos disponibles se recomiendan los siguientes: (1) superponer las curvas ajustadas sobre las trayectorias de las alturas observadas a lo largo del tiempo, (2) representar los residuos frente a los valores predichos por el modelo, y (3) analizar la evolución del sesgo y del error medio cuadrático para las distintas clases de edad. Adicionalmente, se pueden incluir gráficos para verificar la constancia del índice de sitio con la edad para cada árbol.

Para la validación de los modelos y dado que no existe una metodología estándar, se recomienda utilizar la técnica denominada validación cruzada por árbol, que consiste en obtener los residuos eliminados. Éstos se obtienen para cada árbol reajustando de nuevo el modelo pero sin las observaciones de ese árbol. A partir de ellos, se pueden calcular estadísticos similares a los empleados en la fase de ajuste. En sentido estricto, sin embargo, la validación de un modelo sólo se puede realizar con su empleo a lo largo del tiempo, por lo que esta metodología es en realidad un criterio más de decisión para la selección del modelo más adecuado.

### Determinación de la edad de referencia y del error de las predicciones

La utilización del índice de sitio como indicador de la calidad de estación requiere la elección de una determinada edad de referencia. Ésta debe seleccionarse siguiendo las siguientes consideraciones (GOELZ & BURK, 1992): (1) debería ser inferior o igual a la rotación más joven generalmente utilizada en la gestión de la especie, (2) debería ser cercana a la edad de rotación, y (3) debería ser un estimador fiable de la altura a otras edades. Para optimizar esta última consideración, pueden utilizarse diferentes edades de referencia y sus correspondientes alturas observadas para estimar la altura a otras edades. Los resultados así obtenidos deben compararse con los observados en la realización del análisis de tronco, mediante el cálculo del error relativo (RE%), que se determina mediante la ecuación [2] (HUANG *et al.*, 2003). La edad de referencia se puede seleccionar como aquella en la que los errores relativos son menores, siempre que exista un número suficiente de datos a esa edad.

En España ha sido habitual hacer coincidir la edad de referencia con la culminación del crecimiento medio o con los dos tercios del turno de máxima renta en especie para una calidad media (PTA, 1991), sin embargo éste método es menos objetivo que el anterior.

$$RE\% = \frac{\sqrt{\sum_{i=1}^n (H_i - \hat{H}_i)^2 / (n - p)}}{\bar{H}} 100 \quad [2]$$

donde  $H_i$ ,  $\hat{H}_i$  y  $\bar{H}$  son respectivamente los valores observado, predicho y medio de la altura dominante;  $n$  es el número de observaciones y  $p$  el número de parámetros del modelo.

Para conocer hasta cuándo son fiables las estimaciones realizadas a partir de un par de valores altura-edad dados, se puede determinar la altura  $H_2$  a la edad  $t_2$ , considerando la altura  $H_1$  a la edad  $t_1$  como variables predictoras, usando diferentes intervalos de 2, 4, 6, etc. años. Una vez estimadas las alturas de esta manera, se puede calcular la raíz del error medio cuadrático REMC de la altura para los distintos intervalos de tiempo considerados, y el error crítico *Ecrit*, expresado como un porcentaje de la media

observada, obtenido a partir del estadístico  $\chi_n^2$  de FREESE (1960):

$$E_{crit.} = \frac{\sqrt{\tau^2 \sum_{i=1}^n (y_i - \hat{y}_i)^2 / \chi_{crit.}^2}}{\bar{y}} \quad [3]$$

donde  $n$  es el número de observaciones del conjunto de datos,  $y_i$  el valor observado, y el valor predicho por el modelo,  $\bar{y}$  el promedio de los valores observados,  $\tau$  el valor de la distribución normal a un determinado nivel de probabilidad ( $\tau=1.960$  para  $\alpha=0.05$ ), y  $\chi_{crit.}^2$  se obtiene para  $\alpha=0.05$  y  $n$  grados de libertad. Si el error máximo admisible expresado como un porcentaje de la media observada es inferior al error crítico, el test  $\chi_n^2$  indicará que el modelo no proporciona predicciones satisfactorias; en caso contrario, indicará que las predicciones son aceptables.

### EJEMPLO PRÁCTICO

Para ilustrar la metodología expuesta, a continuación se describe un ejemplo práctico desarrollado a partir de los datos de la tabla 1.

El primer análisis exploratorio, recomendado para detectar posibles anomalías o falta de datos, consiste en superponer los gráficos de perfil de los árboles utilizados, una vez efectuada las correcciones de CARMEAN (1972) y NEWBERRY (1991), con la estructura real de la media. En la figura 1a se observa que los datos del ejemplo tienen una distribución ideal, abarcando homogéneamente el rango de edades y de calidades encontradas; en la Figura 1b se muestra un ejem-

plo mucho más real, donde la ausencia de datos en las mejores calidades a edades avanzadas (zona punteada) provoca una caída artificial de la media que puede influir en el ajuste de las ecuaciones de crecimiento en esa zona.

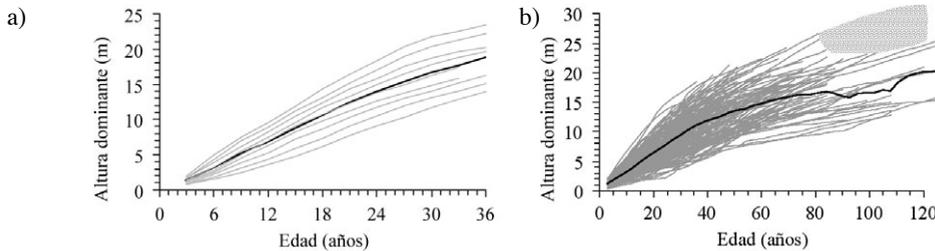
Para el ajuste de las curvas se utilizó un modelo EDA polimórfico que resulta de resolver el parámetro  $b_1$  en el modelo de Chapman-Richards  $H=b_0(1-e^{-b_1A})^{b_2}$ . Para evidenciar la presencia de autocorrelación entre los residuos de un mismo árbol se realizó el ajuste del modelo sin tener en cuenta los parámetros  $\rho$  y  $\gamma$ , y se calculó el test  $t$  de DURBIN (1970), que indicó una alta correlación entre residuos ( $\rho=0,48$   $t=15,95$   $P<0,0001$ ;  $\gamma=0,11$   $t=3,80$   $P<0,0001$ ). El mismo resultado se puso de manifiesto con el valor de 0,80 obtenido para el estadístico de DURBIN-WATSON. Posteriormente, se realizó un ajuste simultáneo de la estructura de la media y del error mediante el procedimiento MODEL del programa SAS/ETS® (SAS INSTITUTE INC., 2000), obteniéndose un coeficiente de determinación ajustado de 0,9929 y un valor del estadístico de DURBIN-WATSON de 1,83 (muy próximo 2), con lo que se evidencia la corrección de la dependencia de residuos. La ecuación [4] muestra el modelo ajustado incluyendo la estructura del error considerada:

$$H_2 = 29,34598 \left( 1 - \left( 1 - \left( \frac{H_1}{29,34598} \right)^{1/1,393015} \right)^{t_2/t_1} \right)^{1,393015} + 0,35642e_{i-1,j} + 0,01159e_{i,j-1} \quad [4]$$

En la figura 2 se muestran los gráficos de error relativo en la predicción de las alturas para

Árbol/Edad	3	6	9	12	15	18	21	24	27	30	33	36
1	0,80	1,60	2,50	3,65	4,75	6,25	7,75	9,00	10,25	11,75	12,88	13,97
2	0,80	1,76	3,14	4,36	5,93	7,30	9,07	10,50	12,00	13,09	14,00	15,10
3	0,92	2,20	3,82	5,29	7,19	8,68	10,33	11,70	12,96	14,07	15,15	16,18
4	1,71	3,92	6,23	8,28	10,44	12,33	14,25	15,75	17,50	18,63	19,61	20,25
5	1,71	4,26	6,91	8,87	11,23	13,14	15,15	17,06	19,00	20,34	21,23	22,11
6	1,88	4,85	7,45	9,50	11,84	14,36	16,33	18,33	20,34	21,81	22,59	23,40
7	1,23	2,70	5,00	7,00	9,00	10,75	12,25	13,75	15,00	16,20	17,40	
8	1,20	2,40	4,46	6,33	8,25	9,75	11,25	12,65	13,77	14,80	15,88	
9	1,43	3,25	5,50	7,50	9,51	11,67	13,24	14,90	16,52	17,75	18,82	19,61

Tabla 1. Datos de altura dominante-edad utilizados para el ajuste de las curvas de calidad una vez efectuadas las correcciones de CARMEAN (1972) y NEWBERRY (1991)



**Figura 1.** Superposición de los gráficos de perfil y la estructura real de la media para los datos usados en el ejemplo (a) y para otro conjunto de datos no tan homogéneo (b)

diferentes edades de referencia, tanto para el modelo seleccionado en este ejemplo (2a) como para otra situación real (2b). La zona sombreada contiene el rango de valores más apropiados para la elección de la edad de referencia, ya que en ella se comete un menor error relativo y existe una muestra suficientemente representativa de datos. Así, en el ejemplo desarrollado (Figura 2a), prácticamente todos los árboles alcanzan los 35 años por lo que se elegirá la edad de referencia de 25 años por presentar el mínimo valor de error relativo; sin embargo, en la figura 2b la edad de referencia debe seleccionarse en torno a 50 años, pues a esa edad se alcanza un compromiso entre un error relativo bajo y un número suficiente de árboles que han alcanzado esa edad.

Para usar el modelo EDA ajustado en la estimación de la altura dominante  $H$  de la una masa a una edad  $t$  determinada, dados un índice de sitio  $IS$  y su correspondiente edad de referencia  $t_{ref}$  asociada, basta con sustituir  $H_2$  por  $H$ ,  $t_2$  por  $t$ ,  $H_1$  por  $IS$  y  $t_1$  por  $t_{ref}$  en la ecuación [4].

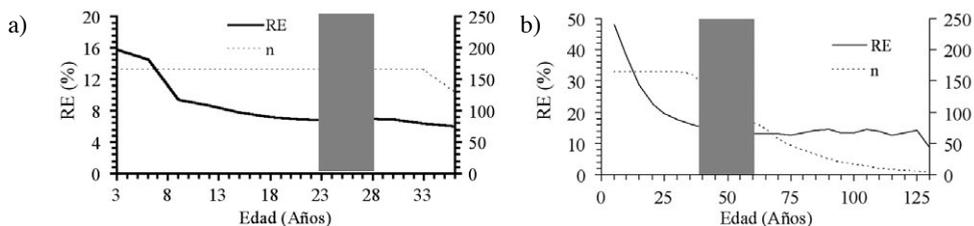
$$H = 29,34598 \left( 1 - \left[ 1 - \left( \frac{IS}{29,34598} \right)^{1/1,393015} \right]^{t_{ref}} \right)^{1,393015} \quad [5]$$

De manera similar, para estimar el índice de sitio a partir de un par altura dominante-edad dado, basta con sustituir  $H_2$  por  $IS$ ,  $t_2$  por  $t_{ref}$ ,  $H_1$  por  $H$  y  $t_1$  por  $t$  también en la ecuación [4]:

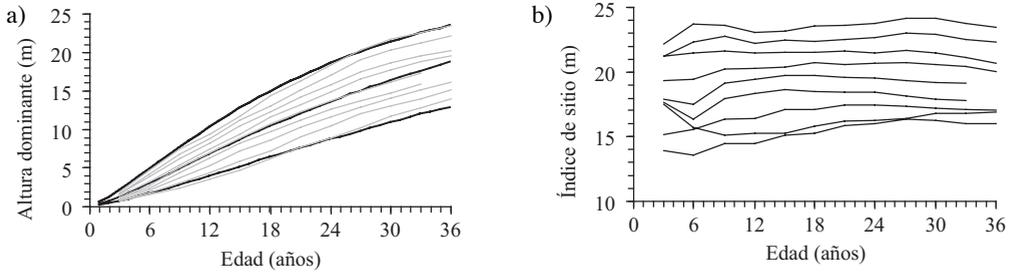
$$IS = 29,34598 \left( 1 - \left[ 1 - \left( \frac{H}{29,34598} \right)^{1/1,393015} \right]^{t_{ref}} \right)^{1,393015} \quad [6]$$

En la figura 3a se muestran las curvas obtenidas para índices de sitio de 9, 14 y 19 m a la edad de 25 años, superpuestas sobre los gráficos de perfil de los árboles empleados en el ajuste. Este gráfico es muy importante para verificar que las curvas ajustadas siguen la tendencia de los datos en todo el rango de edades, puesto que en ocasiones modelos con los mismos estadísticos de comparación presentan formas de las curvas bastante distintas (unas subestiman en las primeras edades y sobreestiman a edades avanzadas y viceversa).

Puesto que el índice de sitio es un atributo fijo de la masa que, por tanto, debe permanecer estable a lo largo del tiempo (HUANG, 1999), para verificar su constancia se elaboró un gráfico que enfrenta las predicciones del índice de sitio con el modelo utilizado para cada uno de



**Figura 2.** Gráfico de error relativo en la predicción de la altura en función de la edad de referencia para el modelo utilizado en el ejemplo (a) y para otro conjunto de datos no tan homogéneo (b)



**Figura 3.** a) Curvas de calidad construidas a partir de la ecuación [5] para índices de sitio de 9, 14 y 19 m a la edad de referencia de 25 años. b) Predicciones de índice de sitio frente a edad para los datos del ejemplo obtenidos tras los análisis de tronco

los árboles a los que se realizó el análisis de tronco; el gráfico revela esa constancia del índice de sitio predicho, salvo para edades jóvenes en las el índice de sitio no es un buen indicador de la calidad de estación (Figura 3b).

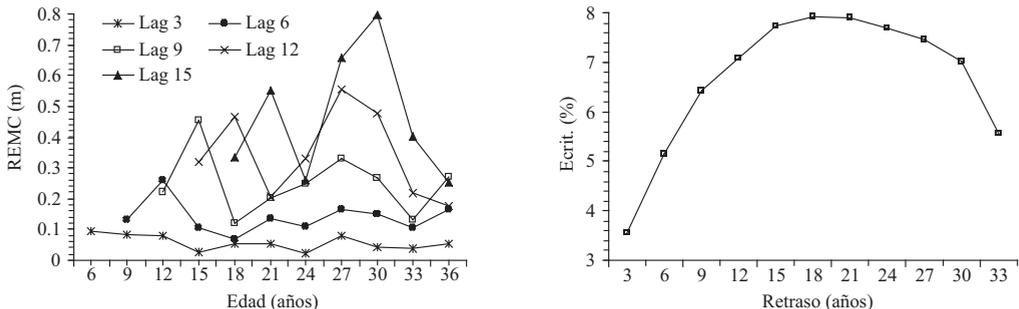
Con respecto hasta cuándo son fiables las curvas para estimar la altura dominante a partir de un par de valores altura-edad dados, el gráfico de REMC frente a diferentes retrasos de edad (Figura 4a) muestra que a medida que aumenta el retraso, REMC hace la mismo para todas las edades. Por su parte, el gráfico de error crítico (Figura 4b) indica, para diferentes retrasos, el mínimo valor del error, expresado como un porcentaje de la media observada, que debería admitirse para considerar que las predicciones del modelo son satisfactorias; así, por ejemplo, si se pretenden realizar predicciones de la altura dominante a 18 años vista, debería admitirse un error relativo al menos del 8% (Figura 4a). En este sentido, es necesario indicar que en modelización forestal un error relativo, a un 95% de probabilidad fiducial, entre un

10 y un 20% parece razonable como límite para elegir los niveles de aceptación y rechazo (HUANG et al., 2003).

El empleo de modelos invariantes en edad de referencia proporciona gran flexibilidad, ya que se pueden adaptar a cambios de gestión que requieran la modificación de la edad de referencia. Así, tanto la ecuación [5] como la [6] pueden resolverse para cualquier edad de referencia sin afectar a las predicciones de la altura dominante o del índice de sitio para una calidad dada. Además, con estos modelos, las estimaciones de altura y de índice de sitio se obtienen directamente, sin necesidad de utilizar procedimientos iterativos.

**BIBLIOGRAFÍA**

CARMEAN, W.H.; 1972. Site index curves for upland oaks in the central states. *For. Sci.* 18: 109-120.  
 CLUTTER, J.L.; FORTSON, J.C.; PIENAAR, L.V.; BRISTER, H.G. & BAILEY, R.L.; 1983. *Timber*



**Figura 4.** Gráficos de (a) REMC y (b) error crítico frente a diferentes retrasos de edad

- management: a quantitative approach*. John Wiley & Sons, Inc. New York.
- DURBIN, J.; 1970. Testing for serial correlation in least-square regression when some of the regressors are lagged dependent variables. *Econometrica* 50: 410-421.
- DYER, M.E. & BAILEY, R.L.; 1987. A test of six methods for estimating true heights from stem analysis data. *For. Sci.* 33(1): 3-13.
- FREESE, F.; 1960. Testing accuracy. *For. Sci.* 6:139-145.
- GOELZ, J.C.G. & BURK, T.E.; 1992. Development of a well-behaved site index equation: jack pine in north central Ontario. *Can. J. For. Res.* 22: 776-784.
- HUANG, S.; TITUS, S.J.; (1993). An index of site productivity for uneven-aged and mixed-species stands. *Can. J. For. Res.* 23: 558-562.
- HUANG S., YANG Y. & WANG Y. (2003). A critical look at procedures for validating growth and yield models. In: A. Amaro, A. D. Reed & P. Soares (eds.), *Modelling forest systems*: 271-293. CABI Publishing. Wallingford.
- HUANG, S.; 1997. Development of a subregion-based compatible height-site index-age model for black spruce in Alberta. *Alberta Land and Forest Service, For. Manag. Res. Note* 5, Pub. n° T/352. Edmonton.
- HUANG, S.; 1999. Development of compatible height and site index models for young and mature stands within an ecosystem-based management framework. In: A. Amaro & M. Tomé (eds.), *Empirical and process-based models for forest tree and stand growth simulation*: 61-98. Lisboa.
- MONSERUD, R.A (1984). Height growth and site index curves for inland Douglas-fir based on stem analysis data and forest habitat type. *For. Sci.* 30: 943-965.
- NEWBERRY, J.D.; 1991. A note on Carmean's estimate of height from stem analysis data. *For. Sci.* 37(1): 368-369.
- PARRESOL, B.R. & VISSAGE, J.S.; 1998. White pine site index for southern forest survey. *Southern Research Station, Res. Pap.* SRS-10. U.S. Department of Agriculture. Forest Service. Asheville, NC.
- PITA, P.A.; 1991. Potencialidad de las estaciones forestales. Curvas de calidad. En: *Seminario sobre Inventario y Ordenación de Montes. Unidad Temática 1*: 18-39. Segovia.
- SAS INSTITUTE INC.; 2000. *SAS/ETS User's Guide, Version 8.0*. Cary, NC.