APLICACIÓN DEL MÉTODO DE LOS COEFICIENTES VARIABLES PARA LA MODELIZACIÓN DEL CRECIMIENTO EN ALTURA DE LAS MASAS FORESTALES

JOSÉ L. BENGOA MTZ. DE MANDOJANA

AREA DE ECOLOGÍA (DPTO. DE CIENCIAS AGROFORESTALES). E.T.S. DE INGENIERÍAS AGRARIAS (PALENCIA)

RESUMEN

En este artículo se aplica el método de los coeficientes variables para la modelización del crecimiento en altura de los montes bajos riojanos de *Quercus pyrenaica*. A modo de ejemplo se hacen algunas estimaciones de la altura de una masa forestal bajo determinados supuestos. Por último se hacen algunas indicaciones acerca de la aplicación del método.

P. C.: Curvas de Calidad, Modelo de crecimiento, Coeficientes variables, Indice de sitio.

SUMMARY

In this paper, random coeficients method is applied for height growth modeling of *Quercus pyrenaica* coppice of La Rioja. As an example, some estimation about future height of an hipotetical stand are made. Finaly some items about aplication of the model are pointed out.

K.W. Height growth model, Random coeficients, Height quality class, Site index.

INTRODUCCION Y OBJETIVOS

Las curvas de calidad representan el crecimiento en altura de las masas forestales. Tanto las curvas de calidad, como las tablas de producción (de las que, con frecuencia forman parte) son modelos de crecimiento, es decir, representaciones simplificadas del crecimiento de un sistema. DAVIS & JOHNSON (1987) y VANCLAY (1994) indican varios aspectos que deben ser evaluados por los usuarios, antes de la utilización de cualquier modelo de crecimiento (inclusive las tablas de producción). Estos requisitos pueden resumirse en dos: los modelos de crecimiento, no sólo deben ser útiles, sino además, deben ser estadísticamente correctos. Con frecuencia, la corrección estadística y, en particular, la fiabilidad de las predicciones no está suficientemente analizada ni en las curvas de calidad ni en las tablas de producción. En este trabajo se muestra una técnica para modelizar el crecimiento en altura de las masas forestales (o cualquier otra variable equivalente) que permite estimar intervalos de confianza para las predicciones de la altura futura o pasada de la masa y para el índice de sitio, conocidos uno o varios valores edad-altura de la misma. A modo de ejemplo se aplica esta técnica utilizando los datos procedentes de parcelas de *Quercus pyrenaica* de La Rioja. Con el modelo construido en base a dichas parcelas se muestra, con tres ejemplos, la fiabilidad que se puede obtener en la predicción de la altura y se estima con cuántos años de antelación se puede predecir la altura dominante de una masa para que el error sea inferior al 20% de la altura predicha (α=0.05).

MATERIAL Y METODOS

Para mostrar la aplicación del método de los coeficientes variables a la elaboración de un modelo de crecimiento en altura (curvas de calidad) se han utilizado datos procedentes de 41 parcelas temporales distribuidas por todo el área ocupada por los montes bajos regulares (intervalo de edades de hasta 5 años) de *Quercus pyrenaica* de La Rioja. Quedan excluídas las masas que presentan una fracción de cabida cubierta inferior al 70% y las que presentan desviaciones anormales en la trayectoria del crecimiento en altura o diámetro, probablemente en respuesta a resalveos, claras fuertes, incendios, plagas, puntisecado o circustancias similares (no se pretende modelizar dichos eventos; en consecuencia, las estimaciones del modelo están condicionadas a que no se presenten tales circustancias). Se ha procurado muestrear una amplia gama de situaciones desde el punto de vista medio-ambiental con objeto de abarcar las distintas formas de crecimiento (calidades) de estas masas en la citada Comunidad Autónoma. Las parcelas son cuadradas, de superficie variable (desde 10x10 m hasta 30x30) dependiendo del tamaño de los árboles. En cada parcela se seleccionaron dos árboles (T1 y T2) del estrato superior que fueron apeados y cortados en trozas de 1 m. Las correspondientes secciones se llevaron al laboratorio para realizar el análisis de tronco. Los cálculos del análisis de tronco se realizaron con la aplicación informática TDIF 2.0 (BENGOA, 1995). Para la estimación de la altura se utilizó el método de CARMEAN (1972). Para cada edad, se considera que la altura dominante de la parcela es el valor más alto de las alturas de los dos árboles apeados, siguiendo, en cierta forma, el criterio de HAMILTON *et al.* (1981). La edad de las parcelas va desde los 13 hasta los 80 años (tan solo 4 parcelas pasan de los 60 años); la altura dominante se sitúa entre 4.5 y 14.4 m; el diámetro cuadrático medio, entre 2.9 y 22.1, (aunque sólo una parcela supera los 14 cm); el número de pies por hectárea va desde 624 hasta 16000 (la parcela de 624 pies/ha es la que presenta un diámetro cuadrático medio de 22.1 cm) y el área basimétrica se sitúa entre 8.4 y 36.8 m^2/ha. La mayor parte de estos robledales proceden del semiabandono del método de beneficio tradicional de estas masas, que era el de monte bajo con cortas a hecho a turnos cortos para la producción de leñas.

Para el análisis de las distintas curvas edad-altura se va a utilizar el método de los coeficientes variables (ELSTON & GRIZZLE, 1962; RAO, 1965, citados en NIGH & SIT, 1996) adaptado al caso según las indicaciones de LAPPI & BAILEY (1988). Las principales características del método se exponen a continuación de forma resumida. Para detalles sobre la aplicación del método, se puede acudir a LAPPI & BAILEY (1988). El modelo considera que la trayectoria de la curva edad-altura de una determinada parcela puede representarse mediante una función de crecimiento (como p. ej. la función de Richards) más una variable aleatoria que representa la variabilidad no explicada por dicha función. La curva correspondiente a cada una de las parcelas puede describirse dando distintos valores a los coeficientes de la función:

$$h_i(t) = A_i\,[1 - D_i\, e^{-B_i t}]^{1/C_i} + u_i(t)$$

Donde: t es la edad de la parcela i; $h_i(t)$ es su altura dominante; A_i, B_i, C_i y D_i son los coeficientes de la función de crecimiento y $u_i(t)$ es la componente aleatoria del crecimiento. Por tratarse de una función no lineal respecto de sus parámetros, el tratamiento estadístico del espacio probabilístico asociado a dichos coeficientes resulta, probablemente inabordable. Para afrontar dicho tratamiento estadístico es conveniente adoptar un modelo lineal: LAPPI & BAILEY (1988) optan por acudir al desarrollo en serie de Taylor de la función de crecimiento, (deteniendo dicho desarrollo en serie en las primeras derivadas de la función respecto de sus coeficientes) alrededor de los valores medios poblacionales ($A=A_m$ $B=B_m$, $C=C_m$ y $D=D_m$). Se considera que, al igual que ocurre con la función original, dicho desarrollo en serie es capaz de adoptar las diferentes formas del crecimiento. En adelante, con objeto de facilitar la exposición, se considerará que sólo los coeficientes A y C son variables, dejando B y D fijos. El D se iguala a uno, para que la función pase por el origen de coordenadas. El B también queda fijo porque presenta bastante correlación con los demás (A y C), lo que conlleva un incremento en el error de su estimación (aspecto de particular importancia en este modelo).

Aunque el modelo de LAPPI & BAILEY (1988) considera las diferencias entre los árboles dentro de cada parcela, además de las existentes entre las parcelas (versión jerárquica), en esta exposición se utiliza una versión no jerárquica, que considera únicamente la variación entre parcelas. De esta forma se abrevia la exposición y se evitan algunos problemas asociados al modelo jerárquico, que se indican más adelante. En definitiva, se parte de que el desarrollo en serie de la función expresa adecuadamente el crecimiento en altura de las distintas parcelas:

$$h_i(t) = h_m(t) + A_i\, f_A(t) + C_i\, f_C(t) + u_i(t)$$

Donde: $h_i(t)$ es la altura dominante de la parcela i.

$h_m(t)$ es la curva altura dominante - edad, media para toda la población considerada.

$f_A(t)$ y $f_C(t)$ son las derivadas parciales de h(t) respecto a A y B, para los valores para $A=A_m$, $B=B_m$, $C=C_m$ y $D=D_m$.

A_i y C_i son coeficientes variables, que se consideran variables aleatorias que toman distintos valores en las distintas parcelas. Su media es cero, hipótesis de partida.

$u_i(t)$ es el error residual, incorrelado con las demás variables aleatorias y, en este ejemplo, incorrelado para distintos valores de t. Se considera que este error residual presenta la misma varianza para todas las edades y todos los árboles.

Conocidos $h_m(t)$ y por lo tanto $f_A(t)$ y $f_C(t)$, este es un modelo lineal en el que a cada parcela le corresponden unos determinados valores de A y C (A_i y C_i). La estimación de los parámetros del modelo consiste en la estimación de las varianzas y covarianzas de A_i y C_i y de la varianza de $u_i(t)$ para el conjunto de las parcelas (una vez estimados A_m, B_m, C_m y D_m). De esta forma queda definido el espacio probabilístico de las citadas variables. En el supuesto de que se considerara que los residuos están correlados, también debería estimarse su correspondiente autocorrelación. La varianza residual $var(u_i)$, puede estimarse mediante la media ponderada de los estimadores de la varianzas residuales de las distintas parcelas. Si el factor de ponderación es igual a los grados de libertad correspondientes al ajuste en cada parcela, este estimador es insesgado (RAO, 1975 citado en LAPPI & BAILEY, 1988). Para estimar las varianzas de A_i y C_i y su correspondiente covarianza, es necesario tener en cuenta que, para cada parcela, se cuenta con estimadores de dichos parámetros $\hat{A}_i$ y $\hat{C}_i$ (obtenidos en la regresión lineal). Por lo tanto la varianza de dichos valores (entre las parcelas) es debida a la variación de los coeficientes (los verdaderos) entre las parcelas y al error de su estimación (ver p. ej. BOX *et al.*, 1989:586). Para el conjunto de las parcelas, la cuasivarianza muestral de $\hat{A}_i$ es un estimador de la suma de la varianza de A_i más la varianza del error de estimación (s_A^2), de donde: $var(A_i)= var(\hat{A}_i) - s_A^2$. Lo mismo puede decirse para la varianza de C_i y la covarianza de ambas variables. En consecuencia, la matriz de estimadores de las varianzas-covarianzas de A_i y C_i puede obtenerse sustrayendo la matriz de varianzas-covarianzas muestrales de $\hat{A}_i$ y $\hat{C}_i$ menos la media ponderada de las matrices de varianzas-covarianzas debidas al error de estimación, obtenidas en el modelo lineal para cada parcela (factor de ponderación igual a los grados de libertad correspondientes al ajuste de cada cada parcela). La estimación de las covarianzas puede realizarse a partir de: $cov(x,y)=(var(x+y)-var(x)-var(y))/2$. La varianza de $(\hat{A}_i+\hat{C}_i)$ pueden estimarse de la misma forma que las de sus componentes (aplicando la definición de la varianza).

El modelo resultante permite realizar predicciones de los parámetros del modelo y de la altura de una parcela a una determinada edad a partir de una muestra formada por uno o varios valores (edad - altura) de dicha parcela. Las variables cuyos valores son conocidos, junto con aquella cuyo valor se desea estimar, constituyen una variable n-dimensional en la que se puede establecer una regresión lineal múltiple (modelo en desviaciones) de la variable desconocida respecto a las variables conocidas (ver ARNAIZ 1978: distribución normal n-dimensional). Para detalles sobre la aplicación del método ver LAPPY & BAILEY (1988).

RESULTADOS

La estimación de los parámetros de la curva media poblacional (función de Richards, anteriormente expuesta) se ha realizado por diferentes métodos, obteniendo resultados parecidos: (1) mediante ajuste de todas las series de datos edad-altura de las distintas parcelas de forma conjunta a la función de Richards y (2) mediante ajuste de dicha función, de forma individualizada a los datos de cada parcela, para promediar, a continuación los coeficientes de la función (se ha ensayado la media aritmética y ponderada con distintos factores de ponderación). En todos los casos se trata de ajuste, mínimo-cuadrático, no lineal, por el método del desarrollo en serie (DRAPPER & SMITH, 1981). La curva media poblacional obtenida mediante la media aritmética de los coeficientes de cada parcela es la siguiente:

$$h= 18.93(1 - e^{-0.813\,t})^{1/0.0294}$$

El desarrollo en serie de esta función, con las indicaciones anteriormente apuntadas, se ajustó a los datos edad-altura de cada parcela (modelo lineal). El estimador de la desviación típica residual es 0.36 m (media ponderada para todas las parcelas). Los estimadores de las varianzas y covarianza de los coeficientes A y C se exponen en el cuadro nº1. En base a estos datos y siguiendo las indicaciones de LAPPI & BAILEY (1988) adaptadas al modelo no jerárquico, se ha estimado la altura y desviación típica (la del estimador de la altura) para una parcela hipotética, bajo tres supuestos (ver cuadro nº 3), dando por válida la hipótesis de distribución normal n-dimensional de los coeficientes, y varianza de la altura conocida.

DISCUSIÓN

Como puede observarse, el intervalo de confianza para la altura de la parcela a los 50 años es muy amplio, en especial, si únicamente se conoce la altura que tenía a los 10 años. De hecho, con este método se puede mostrar que si se desea conseguir estimaciones con un márgen de error inferior a ±20% de la altura estimada, con el 95% de probabilidad, no se pueden hacer previsiones a largo plazo (si se conoce la altura a los 10 años, la previsión no puede hacerse a más de 8 años vista; si se conoce a los 20 años, la previsión puede hacerse a 15 años vista; si se conoce la altura de la parcela a los 30 años, la previsión puede alcanzar hasta la edad de 50 años, en todos los casos para las parcelas aquí utilizadas). No se han hecho cálculos para edades superiores a los 50 años porque se cuenta con pocas parcelas que superen esta edad y su comportamiento no es suficientemente conocido. De hecho el valor de la asíntota de la curva media poblacional (19 m) no pretende indicar que se alcance, por término medio esta altura máxima, ya que dicha predicción corresponde a edades muy alejadas del intervalo de edades indicado (0-50 años). Para tener una cierta idea del valor de la asíntota, es necesario contar con parcelas de, al menos, 150 o 200 años (los análisis de tronco de parcelas de 100 años, muestran que esta especie puede presentar a esta edad un ritmo de crecimiento en altura aceptable, no estancado).

Quizás pudiera resultar sorprendente que la estimación de la altura correspondiente al tercer ejemplo sea menor que la de los otros dos. Sin embargo dicho resultado es coherente con los datos de partida ya que éstos reflejan una curvatura temprana y particularmente pronunciada, que tiene como consecuencia que la altura a los 50 años sea más baja de lo esperado si sólo se cuenta con un punto de la curva edad-altura. De hecho, la capacidad de este modelo para utilizar, en la inferencia, varias alturas a distintas edades, junto con la de admitir la variación de más de un coeficiente y, claro está, la de proporcionar intervalos de confianza le confieren unas posibilidades muy amplias, no sólo para estimar la altura de las parcelas sino también para analizar y caracterizar las distintas formas de crecimiento.

Debe tenerse en cuenta que, de acuerdo con la propuesta de Lappi & Bailey, en la inferencia se considera que las varianzas y covarianzas de A y C son datos conocidos, no estimados, lo que no deja de ser una incorrección. Para valorar hasta que punto esta hipótesis puede estar lejos de la realidad se ha estimado el error de estimación de dichas varianzas. Téngase en cuenta que las varianzas y covarianzas de A_i y C_i se obtienen sustrayendo las varianzas y covarianzas debidas al error de estimación de los coeficientes $\hat{A}_i$ y $\hat{C}_i$, de sus varianzas y covarianzas muestrales. Si ambos valores son parecidos y presentan un error de estimación elevado, la diferencia puede ser incluso negativa, reflejo evidente de que el modelo no es adecuado para dichos datos. En el ejemplo que se expone en este trabajo, como la varianza debida al error de estimación es mucho menor que la debida a las diferencias entre parcelas, los errores en la primera no tienen mucha incidencia en la diferencia de ambas (ver cuadro nº 2). En cambio, cuando se aplica el modelo jerárquico (LAPPI & BAILEY, 1988), debido a que las diferencias entre los árboles de las parcelas son bastante pequeñas, el error en la estimación de las varianzas es alto en comparación con el debido a las diferencias entre árboles (salvo que la muestra de árboles sea bastante grande), lo que puede suponer una objeción importante para su aplicación. La consideración del coeficiente B como no variable esta relacionada con este razonamiento: debido a su correlación con los otros dos coeficientes, su inclusión en el modelo incrementa notablemente el error de su estimación.

Un aspecto que no queda suficientemente resuelto con el modelo utilizado en este trabajo es el de la autocorrelación de los residuos del modelo lineal, que da lugar a una infravaloración del error de estimación de los coeficientes. Distanciando suficientemente los datos edad-altura (aprox. más de 5 años o 1 m), la autocorrelación desciende notablemente y su influencia en los resultados del modelo es relativamente pequeña. En cualquier caso, se obtienen resultados más precisos, sin mucho esfuerzo, modelizando los crecimientos (correspondientes a intervalos no solapados) en lugar de la altura acumulada. Otra alternativa, aunque probablemente no suficientemente satisfactoria es la de acudir al ajuste por mínimos cuadrados generalizados. GARCÍA (1983) propuso un modelo que considera dicha autocorrelación, aunque no proporciona estimadores de las varianzas de los coeficientes, por lo que no le es aplicable esta metodología.

AGRADECIMIENTOS

Este artículo ha sido elaborado en base a los trabajos realizados en el Dpto. de Sistemas Forestales del INIA (actualmente CIFOR), durante mi estancia como becario en este centro, colaborando en los proyectos nº 8147 del INIA y "MEDCOP", de la Comisión Europea.

REFERENCIAS BIBLIOGRÁFICAS

ARNAIZ, G. 1978. Introducción a la estadística. Ed. Lex nova. Valladolid. 818 pp.

BENGOA, J. L. 1996. TDIF 2.0. Aplicación informática para el cálculo de árboles tipo y análisis de tronco. Montes 43: 51-55.

BOX, G. E.; HUNTER, W. G. & HUNTER, J. S. 1989. Estadística para investigadores. Ed. Reverté.

CARMEAN, W. H. 1972. Site index curves for Upland Oaks in the Central States. For. Sci., 18(2): 109-120.

DAVIS, L. S. & JOHNSON, K. N. 1987. Forest management. Mc. Graw Hill. New York.

DRAPPER & SMITH, 1981. Applied Regression analysis. John Wiley & Sons. New York.

ELSTON, R. C. & GRIZZLE, J. E. 1962. Estimation on time response curves and their confidence bands. Biometrics, 18: 148-159.

GARCÍA, O. 1983b. A stochastic differential equation model for the height growth of forest stands. Biometrics, 39: 1059-1072.

HAMILTON, G. J., CHRISTIE, J. M. & EDWARDS, J. 1981. Yield models for forest management. Forestry Commision Booklet No. 48.

LAPPI, J. & BAILEY, R. L. 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. For. Sci. 34(4): 907-927.

RAO, C. R. 1965. The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves. Biometrika, 52 447-458.

RAO, C. R. 1975. Simultaneous estimation of parameters in different linear models and applications to biometric problems. Biometrics,, 31: 545-554.

VANCLAY, J. 1994. Modelling forest growth and yield. Applications to mixed tropical forests. CAB International, Wallingford.

Varianzas y covarianzas de los Coeficientes			
Causa de variación	A	C	(A,C)
Error de Estimación	0.7177	0.00159	-0.8653
Diferencias entre parcelas	10.920	0.04973	-0.3848

Cuadro nº 1.

Intervalo de confianza del 90% para las varianzas de los coeficientes.				
Causa de variación	Intevalo de confianza			
	var ($\hat{A}$)		var ($\hat{C}$)	
Error de Estimación	0.38	1.93	0.00084	0.0043
Diferencias entre parcelas más error de estimación *	8.34	17.57	0.0356	0.0742

* (Estimadores de var ($\hat{A}_i$) y var ($\hat{C}_i$)). Estos cálculos se han realizado suponiendo que se pudiera considerar que (n-k-1)vâr(u_i) / var(u_i) sigue una distribución χ^2_{p-k-1} y contando con que vâr ($\hat{A}_i$) y vâr ($\hat{C}_i$) siguen una distribución χ^2_{n-1} (n es el número de parcelas; p es el número medio de pares de valores edad-altura por parcela y k, el número de coef. variables = 2).

Cuadro nº 2.

Ejemplos de estimación de la altura con intervalos de confianza.						
	Datos conocidos			Datos estimados		
Nº Ejemplo	Edad (años)	Altura (m)	Edad (años)	Altura (m)	S_H (m)	Intervalo 95% (m)
1	10	3.0	50	13.3	1.99	9.4 - 17.2
2	20	5.0	50	12.0	1.63	8.8 - 15.2
3	10 20	3.0 5.0	50	10.5	1.35	7.80 - 131

Cuadro nº 3. S_H es el estimador de la desviación típica del estimador de la altura.