REVISIÓN Y DISCUSIÓN METODOLÓGICA PARA LA DETERMINACIÓN DE LA BIOMASA FORESTAL

M.A. Balboa Murias, J.G. Álvarez González, A. Merino García y M. Barrio Anta

Universidad de Santiago de Compostela. Escuela Politécnica Superior de Lugo. Campus Universitario. 27002-LUGO (España). Correo electrónico: mibalboa@lugo.usc.es

RESUMEN

Se describen las metodologías de inventario de la biomasa arbórea basadas fundamentalmente en relaciones de proporcionalidad, así como otras más específicas y de utilización más puntual: "*Randomized Branch Sampling*" e "*Importance Sampling*". Se analizan las metodologías de ajuste de los modelos de biomasa arbórea y los procedimientos para garantizar en ellos la aditividad.

Palabras clave: *Aditividad, Biomasa, Ratio-type estimators, SUR*

INVENTARIO DE LA BIOMASA ARBÓREA

El empleo de modelos matemáticos sigue siendo en la actualidad el método más habitual para estimar la biomasa forestal. Estas ecuaciones estiman el peso de las diferentes fracciones arbóreas a partir de variables de árbol individual y de masa. Para su ajuste se hace necesario por tanto, conocer estas variables explicativas y determinar el peso seco de las fracciones arbóreas.

Sea cual sea la técnica empleada para inventariar la biomasa forestal, ha de incluir, por lo menos en su fase de desarrollo, una determinación de la biomasa del árbol individual mediante las denominadas técnicas de "árbol tipo", y su agregación para determinar la biomasa total por unidad de superficie (SATOO Y MADGWICK, 1982). Las técnicas que se basan los factores de proporcionalidad entre una parte y el todo son las más utilizadas. Lo habitual es pesar en verde la totalidad de la copa, y tomar una muestra para secar en laboratorio y determinar a partir de ella la humedad y la proporción de las diferentes fracciones (ramas, ramillos y hojas). Con respecto al tronco, se trocea, se pesan y se cubican todas sus trozas. En cualquier caso, se han de extraer rodajas transversales, a diferentes alturas, a partir de las cuales se determina la proporción de madera y corteza y sus respectivas humedades. Para determinar el peso de las trozas que no han sido pesadas y sí cubicadas se emplea la densidad de las trozas donde se han realizado estas dos operaciones o la de los discos que son llevados a laboratorio. Este tipo de metodología tiene un fundamento similar a la técnica de "*Ratio-Type Estimators*" (relaciones de proporcionalidad), también basada en las relaciones de escala entre una parte y el todo (BRIGGS et *al.,* 1987). Es sencilla y fácil de aplicar, aunque tiende a obtener valores sesgados de los pesos de las diferentes fracciones (VALENTINE et *al.,* 1984; CUNIA, 1979). El peso seco de hojas y ramas se determina por medición directa, y el del tronco (madera y corteza) mediante muestreo estratificado y posterior aplicación de estos factores de proporcionalidad. El tronco se fracciona en trozas de igual longitud, generalmente tres. En cada una de ellas se extraen tres discos transversales al eje. Su localización, tomando como referencia la distancia

ISSN: 1575-2410

© 2004 Sociedad Española de Ciencias Forestales

desde la base de la troza, se determina en función de un número elegido al azar, del 1 al 100, que se multiplica en forma decimal por la longitud total de la troza. Estos discos son pesados en campo, al igual que el resto del tronco, y son llevados a laboratorio para separar estas dos fracciones y determinar sus respectivos pesos secos. Finalmente, los factores de proporcionalidad relacionan el peso verde y seco de los discos con el peso verde de las trozas, para de esta forma poder determinar el peso seco del tronco.

Además de las metodologías de muestreo antes comentadas, existen otras técnicas más específicas, con las que se consiguen estimaciones no sesgadas de la biomasa arbórea. Entre ellas destacan *Randomized Branch Sampling* (RBS o muestreo aleatorio de sendas o caminos) e *Importance Sampling* (Importancia de muestreo) (VALENTINE et *al.,* 1984; GREGOIRE et *al.,* 1995b). Ambas se basan en seleccionar un "camino de muestreo" desde la base del árbol hasta un ramillo terminal de la copa. El camino empieza en el tronco, que constituye el primer segmento, y asciende hasta el primer verticilo o nudo; aquí se escoge aleatoriamente una de las ramas y se le asigna una probabilidad de selección (p_r), igual al producto de su longitud (l_r) por el cuadrado de su diámetro en la inserción (d_{insr}), dividido todo ello por la suma de dichos productos de las *n* ramas que constituyen el verticilo:

$$p_r = \frac{l_r \cdot d_{ins\,r}^{\,2}}{\sum_{i=1}^{n} l_i \cdot d_{ins\,i}^{\,2}} \qquad [1]$$

El camino de muestreo continúa por la rama escogida en el primer verticilo, que constituye el segundo segmento, e iría hasta la siguiente bifurcación, donde se escoge una nueva rama al azar y se le asigna su probabilidad de elección de forma análoga al segmento anterior. El camino de muestreo finaliza cuando se alcanza un ramillo terminal de la copa. La probabilidad de elección de ese camino de muestreo, P_k, es igual al producto de las probabilidades de cada uno de sus segmentos, teniendo en cuenta que el segmento inicial tiene una probabilidad de elección de uno, puesto que forma parte de todos los caminos de muestreo posibles:

$$P_k = \prod_{r=1}^{k} p_r \qquad [2]$$

Con esta técnica, especialmente válida para árboles de elevadas dimensiones y copas muy desarrolladas, el esfuerzo de muestreo se reduce considerablemente ya que la biomasa arbórea puede ser estimada a partir de un único camino de muestreo, aunque si se quiere obtener el error estándar de la estimación resulta necesario emplear dos o tres caminos de muestreo diferentes. En el camino o caminos de muestreo seleccionados se pesa la totalidad de cada una de las fracciones del árbol consideradas. El peso verde total de las mismas en el árbol será igual al peso obtenido en el muestreo dividido por la probabilidad del camino de muestreo hasta el segmento en el que la fracción considerada deja de formar parte de la muestra.

La otra metodología, *Importance Sampling*, desarrollada por VALENTINE et *al.* (1984), se centra especialmente a la biomasa de tronco. Como en RBS se elige un camino de muestreo y se multiplica cada segmento por el factor $1/P_k$ correspondiente. El fundamento de esta técnica consiste en considerar este camino como un conjunto de rodajas de grosor fijo y volumen conocido. Una de ellas se selecciona al azar, con una probabilidad proporcional a su volumen. A continuación se determina su peso seco y se divide por su probabilidad de selección. El resultado de esta operación constituye una estimación no sesgada del peso seco del árbol.

En la práctica, se mide el diámetro de tronco y de ramas en diferentes puntos del camino de muestreo y se define en ellos $A(L_s)$, que responde a la siguiente fórmula:

$$A(L_s) = \frac{D(L_s)^2}{P_k} \qquad [3]$$

donde $D(L_s)^2$ es es diámetro del segmento a una distancia Ls del inicio del camino de muestreo. Con estos valores, se ajusta la función S(L), que luego se integra en toda la longitud del camino (l), obteniendo una aproximación del volumen de madera del mismo (V):

$$V(l) = \int_0^l S(L)dL \qquad [4]$$

El siguiente paso consiste en localizar un punto, a una distancia α del origen del camino de muestreo, cuya probabilidad de selección es proporcional a $S(L)$, donde se corta un disco de

madera. En este punto se cumple que $V(\alpha)=n\cdot V(l)$, siendo n un número seleccionado al azar en una distribución (0,1). A continuación se determina el peso seco mayorado para ese disco de ancho determinado con la siguiente expresión:

$$W^*(\alpha) = \frac{W(\alpha)}{P_k} \qquad [5]$$

donde $W(\alpha)$ es el peso seco de la rodaja localizada a una distancia $L=\alpha$ de la base del árbol, con el P_k correspondiente al segmento del camino donde se localiza. La expresión final del peso seco del tronco responde a la siguiente expresión:

$$\hat{W}_T = \frac{W^*(\alpha)\cdot V(l)}{S(\alpha)} \qquad [6]$$

AJUSTE DE LOS MODELOS DE BIOMASA

Una vez determinados los pesos de las diferentes fracciones de los árboles medidos es necesario ajustar modelos matemáticos que relacionen esos pesos con una o más variables representativas de esos árboles o de la masa. Aunque existe una gran variedad de ecuaciones válidas para modelizar la biomasa, todas ellas derivan de alguna de las tres formas matemáticas siguientes (Pardé, 1980; SNOWDON, 1985; PARRESOL, 1999):

Lineal: $P = \beta_0 + \beta_1 \cdot x_1 + ... + \beta_j \cdot x_j + \varepsilon$ [7]

No lineal: $P = \beta_0 \cdot x_1^{\beta_1} \cdot x_2^{\beta_2} \cdot ... \cdot x_j^{\beta_j} + \varepsilon$ [8]

No lineal: $P = \beta_0 \cdot x_1^{\beta_1} \cdot x_2^{\beta_2} \cdot ... \cdot x_j^{\beta_j} \cdot \varepsilon$ [9]

donde P = biomasa total o de alguna de las fracciones del árbol; x = variables explicativas de árbol o de masa; ß = parámetros del modelo; e = error.

El primero de los casos [7] se puede ajustar mediante simples procedimientos de regresión lineal múltiple. En el caso del modelo [8], se precisan técnicas de regresión no lineal, con empleo de procedimientos iterativos para la estimación de parámetros. Las ecuaciones del tipo [9], no lineales y con error de tipo multiplicativo, se pueden ajustar mediante una transformación de variables aplicando logaritmos a ambos lados de la igualdad, y empleando mínimos cuadrados a continuación. De esta forma, además, se consigue generalmente corregir el problema de la heterocedasticidad (CARROL Y RUPPERT, 1988).

$$\log P = \beta_0 + \beta_1 \cdot \log x_1 + \beta_2 \cdot \log x_2 + ... + \beta_j \cdot \log x_j + \ln \varepsilon \qquad [10]$$

Los valores de los estadísticos de comparación que genera este modelo no son directamente comparables con los correspondientes a los modelos [7] y [8]. Lo más habitual es deshacer esta transformación y expresar la variable dependiente en unidades aritméticas, incorporando un factor corrector del sesgo, c, que aparece como multiplicativo en la expresión alométrica y que se calcula como:

$$c = e^{\left(\frac{EMC}{2}\right)} \qquad [11]$$

donde *EMC* es el error medio cuadrático del modelo logarítmico ajustado. Esta corrección, propuesta por MEYER (1944), se puede obviar si el ajuste de la función es bueno, aunque otros autores recomiendan su inclusión, sobre todo si el grado de precisión que se pretende obtener es elevado (BASKERVILLE, 1972; STERBA, 2001).

ADITIVIDAD EN LAS ECUACIONES DE BIOMASA

Una de las propiedades más importantes que deben cumplir estas ecuaciones de biomasa de las diferentes fracciones de un árbol es la llamada aditividad. Consiste en que la suma de las estimaciones de los pesos de todos estos componentes sea igual a la estimación del peso del árbol completo. Existen tres procedimientos para forzar esta propiedad en un sistema de ecuaciones de biomasa, dependiendo de cómo se agreguen los diferentes componentes (KOZAK, 1970; CHIYENDA Y KOZAK, 1984; PARRESOL, 1999).

El procedimiento más sencillo consiste en ajustar de forma individual los modelos de cada fracción y obtener la estimación del peso total como suma de los modelos de todas las fracciones (SRIVASTAVA Y GILES, 1987).

$$\begin{aligned} \hat{p}_1 &= f_1(x_1) \\ \hat{p}_2 &= f_2(x_2) \\ &\vdots \\ \hat{p}_k &= f_k(x_k) \\ \hat{p}_{total} &= \sum_{i=1}^{k} \hat{p}_i = \hat{p}_1 + \hat{p}_2 + \ldots + \hat{p}_k \end{aligned} \qquad [12]$$

Si se emplea esta metodología, cada modelo puede incluir diferentes variables explicativas. Sin embargo, el principal inconveniente de este procedimiento es que la varianza de la estimación de la biomasa total $s^2(\hat{p}_{total})$ es muy elevada, puesto que su valor depende de las varianzas de las estimaciones de cada fracción $s^2(\hat{p}_i)$, así como de las covarianzas entre estimaciones $Cov(\hat{p}_i,\hat{p}_j)$

$$s^2(\hat{p}_{total}) = \sum_{i=1}^{k} s^2(\hat{p}_i) + 2 \cdot \sum\sum Cov(\hat{p}_i, \hat{p}_j) \qquad [13]$$

$$Cov(\hat{p}_i, \hat{p}_j) = \lambda_{p_i p_j} \cdot \sqrt{s^2(\hat{p}_i) \cdot s^2(\hat{p}_j)}$$

$\lambda_{p_i p_j}$=correlación entre la estimación de la fracción i y la fracción j.

El segundo procedimiento se basa en emplear las mismas variables explicativas en las ecuaciones de regresión de los diferentes componentes del árbol y en la de biomasa total. De este modo, los valores de los parámetros del modelo de estimación del peso total son la suma de los parámetros obtenidos para cada variable explicativa al ajustar el modelo a cada fracción por separado. Pese a la obligación de emplear el mismo modelo en los k componentes y en la biomasa total, algunos autores han conseguido salvar este condicionante empleando mínimos cuadrados restringidos (CHIYENDA Y KOZAK, 1984).

El principal inconveniente de esta metodología radica en asumir que los *k* componentes de un mismo árbol son independientes, esto es, que no hay dependencia entre los errores, lo que no es cierto en la mayoría de los casos.

$$\begin{aligned}
\hat{p}_1 &= \beta_{01} + \beta_{11} \cdot x_1 + \beta_{21} \cdot x_2 + \ldots + \beta_{j1} \cdot x_j \\
\hat{p}_2 &= \beta_{02} + \beta_{12} \cdot x_1 + \beta_{22} \cdot x_2 + \ldots + \beta_{j2} \cdot x_j \\
&\vdots \\
\hat{p}_k &= \beta_{0k} + \beta_{1k} \cdot x_1 + \beta_{2k} \cdot x_2 + \ldots + \beta_{jk} \cdot x_j \\
\hat{p}_{total} &= (\beta_{01} + \beta_{02} + \ldots + \beta_{0k}) + (\beta_{11} + \beta_{12} + \ldots + \beta_{1k}) \cdot x_1 + \\
&+ (\beta_{21} + \beta_{22} + \ldots + \beta_{2k}) \cdot x_2 + \ldots + (\beta_{j1} + \beta_{j2} + \ldots + \beta_{jk}) x_j
\end{aligned} \qquad [14]$$

donde x_j = variables explicativas comunes a todos los modelos; $ß_{jk}$ = parámetros del modelo ajustado a la fracción i.

Asumiendo la hipótesis de independencia entre componentes, la varianza de la estimación de la biomasa total, $s^2(\hat{p}_{total})$, es la suma de las varianzas de las k fracciones arbóreas, $s^2(\hat{p}_i)$.

$$s^2(\hat{p}_{total}) = \sum_{i=1}^{k} s^2(\hat{p}_i) \qquad [15]$$

Por otro lado, la necesidad de emplear las mismas variables regresoras en todas las ecuaciones de biomasa implica en muchos casos la aparición de multicolinealidad.

La última metodología es la más flexible pero es también la de más difícil utilización. Se basa en el ajuste de un sistema de ecuaciones aparentemente no relacionadas formado por los modelos de los k componentes arbóreos junto con el de biomasa total. En este sistema [16], no es necesario que las ecuaciones de cada fracción arbórea presenten la misma estructura matemática ni las mismas variables independientes. Las variables explicativas del modelo de biomasa total son todas las variables regresoras que aparecen en los modelos de cada componente.

$$\begin{aligned}
\hat{p}_1 &= f_1(x_1) \\
\hat{p}_2 &= f_2(x_2) \\
&\vdots \\
\hat{p}_k &= f_k(x_k) \\
\hat{p}_{total} &= f_{total}(x_1, x_2, \ldots, x_k)
\end{aligned} \qquad [16]$$

Este sistema, sin relaciones analíticas entre ecuaciones, se suele resolver empleando regresión SUR (*Seemingly Unrelated Regression*), conocida también como mínimos cuadrados generalizados conjuntos. Se trata de una generalización del método de regresión por mínimos cuadrados ordinarios (*Ordinary Least Squares*, OLS) para un sistema de ecuaciones. Como OLS, el método SUR asume que todas las variables regresoras son variables independientes, pero SUR usa la correlación entre los errores de diferentes ecuaciones (es decir, $cov\ (e_i\ ,\ e_j) \neq 0$ para los pares de i y de j) para mejorar la eficiencia de las estimaciones. La aditividad en este caso se garantiza imponiendo restricciones a los parámetros del modelo de biomasa total. Otra ventaja de esta metodología radica en corregir el problema de dependencia entre los errores inherentes a las estimaciones de cada componente.

Lo habitual es ajustar inicialmente todas las ecuaciones usando mínimos cuadrados ordinarios (OLS), y determinar los residuos que son utilizados para estimar la matriz de covarianza,

$\hat{\Sigma}$. A continuación, basándose en ella, se realiza la regresión con mínimos cuadrados generalizados conjuntos y se determina la matriz de covarianza del error entre ecuaciones, (Σ). Teóricamente, la estimación de parámetros con el método SUR resulta, al menos, tan eficiente como con OLS para muestras amplias. Sin embargo, si la muestra es pequeña y la dependencia de errores no es muy marcada, se incrementa la variabilidad de las estimaciones SUR, debido a la necesidad de estimar la matriz de covarianzas a partir de mínimos cuadrados ordinarios. El método SUR es igualmente consistente y asintóticamente eficiente con los sistemas de ecuaciones no lineales, aunque en ese se habla de NSUR.

De las tres metodologías planteadas, mínimos cuadrados generalizados conjuntos es la que parece más recomendable. En el caso de que exista dependencia entre errores, SUR es superior al procedimiento I, ya que considera esta circunstancia. SUR también muestra un mejor comportamiento que la metodología II, que asume independencia entre las fracciones de biomasa de un mismo árbol y presenta además problemas de multicolinealidad al emplear las mismas variables regresoras en todas las ecuaciones. Incluso el aplicar SUR al sistema de ecuaciones [14] no redunda en ningún beneficio ya que la covarianza aumenta cuando los modelos de todas las fracciones tienen las mismas variables explicativas (SRIVASTAVA Y GILES, 1987), y los resultados serían los mismos a los obtenidos al considerar las ecuaciones independientes empleando mínimos cuadrados.

BIBLIOGRAFÍA

BASKERVILLE, G.L.; 1972. Use of logarithmic regression in the estimation of plant biomass. *Can. J. For. Res.* 2: 49-53.

BRIGGS, R.D.; CUNIA, T.; WHITE, E.H. & YAWNEY, H.W.; 1987. Estimating sample tree biomass by subsampling: some empirical results. *In*: E.H. Wharton, & T. Cunia, (comps.). *Estimating tree biomass regressions and their error. Proc. of the Workshop on Tree biomass functions and their contribution to the error of forest inventory estimates*: 119-127. USDA For. Serv. Gen. Tech. Rep. NE-117.

CARROLL, S.S. & RUPPERT, D.; 1988. *Transformation and Weighting in Regression.* Chapman & Hall. New York.

CHIYENDA, S.S. & KOZAK, A.; 1984. Additivity of component biomass regression equations when the underlying model is linear. *Can. J. For. Res.* 14: 441-446.

CUNIA, T.; 1979. On sampling trees for biomass table construction: some statistical comments. *In*: W.E. Frayer (ed.), *Forest resource inventories* 2: 643-664. Colorado State Univ. Fort Collins.

CUNIA, T. & BRIGGS, R.D.; 1984. Forcing additivity of biomass tables: some empirical results. *Can. J. For. Res.* 14: 376-384.

GREGOIRE, T.G.; VALENTINE, H.T. & FURNIVAL, G.M.; 1995b. Sampling methods to estimate foliage and other characteristics of individual trees. *Ecology* 76(4): 1181-1194.

KOZAK, A.; 1970. Methods of ensuring additivity of biomass components by regression analysis. *For. Chron.* 46(6): 402-404.

MEYER, H.A.; 1944. A correction for systematic error occurring in the application of the logarithmic volume equation. *Pennsylvania State University, Forest Research Paper* 7.

PARDÉ, J.; 1980. Forest biomass. *For. Abstr.* 41: 343-362.

PARRESOL, B.R.; 1999. Assessing tree and stand biomass: a review with examples and critical comparisons. *For. Sci.* 45: 573-593.

SATOO, T. & MADGWICK, H.A.I.; 1982. *Forest Biomass.* Forestry Sciences. Kluwer Academic Publishers Group. Dordretch.

SNOWDON, P.; 1985. Alternative sampling strategies and regression models for estimating forest biomass. *Aust. For. Res.* 15: 353-366.

SRIVASTAVA, V.K. & GILES, D.E.A.; 1987. *Seemingly unrelated regression equations models: estimation and inference.* Marcell Dekker. New York.

STERBA, H.; 2001. *Forest growth modelling.* Curso impartido en la Escuela Técnica Superior de Ingenieros de Montes. Madrid.

VALENTINE, H.T.; TRITTON, L.M. & FURNIVAL, G.M.; 1984. Subsampling trees for biomass, volume, or mineral content. *For. Sci.* 30: 673-681.