



7º CONGRESO FORESTAL ESPAÑOL

**Gestión del monte: servicios
ambientales y bioeconomía**

26 - 30 junio 2017 | Plasencia
Cáceres, Extremadura

7CFE01-094

Edita: Sociedad Española de Ciencias Forestales
Plasencia. Cáceres, Extremadura. 26-30 junio 2017
ISBN 978-84-941695-2-6

© Sociedad Española de Ciencias Forestales

Estimación de variables dasométricas en rodales y medidas de error asociadas: dos casos de estudio en Burgos y Madrid

TORDESILLAS TORRES, A.¹, SASTRE PARDIAL, L.¹, MAURO GUTIÉRREZ, F.², PASCUAL CASTAÑO, C.¹, MANZANERA DE LA VEGA, J.A.¹

¹ Grupo de investigación SILVANET. E.T.S.I. Montes. Universidad Politécnica de Madrid.

² Forest Engineering Resources and Management. College of Forestry. Oregon State University.

Resumen

El uso de datos LiDAR en inventarios forestales ha sido desarrollado de forma notable durante las dos últimas décadas. Sin embargo, la metodología tradicional de desarrollar modelos de regresión para unidades equiparables a un cuartel o monte y realizar estimaciones a nivel de rodal en base a dicho modelo (estimadores sintéticos) puede ser problemática y producir estimaciones sesgadas a nivel de rodal cuando la variabilidad entre rodales es significativa. Si la variabilidad entre rodales es significativa se recomienda el uso de estimadores EBLUP (Empirical Best Linear Unbiased Predictor) basados en modelos mixtos.

En este estudio se emplean datos LiDAR para producir estimaciones a nivel de rodal para las variables Área Basimétrica (G), Densidad (N) y Volumen (V) en la ladera de la peña del Águila (Valle de la Fuenfría, Madrid) y en el Valle de las Caderechas (Burgos). Los valores medios de la reducción de la incertidumbre al incluir información LiDAR en Burgos, fueron 82% para V, 58% para N y 38% para G. Para Madrid estos valores fueron 29.43%, 10.53% y 57.53%. La menor reducción de la incertidumbre en Madrid se explica por el elevado número de parcelas empleado en esta zona.

Palabras clave

LiDAR, inventario forestal, modelos mixtos, EBLUP, estimación en áreas pequeñas

1. Introducción

El inventario forestal debe proporcionar estimaciones sobre las existencias presentes en distintas subunidades de gestión del monte como cuarteles, cantones o rodales (áreas de interés de aquí en adelante), con el fin de hacer que las decisiones del proceso de planificación forestal se realicen en base a información actualizada. Las observaciones procedentes de una misma unidad de muestreo o gestión pueden presentar una alta correlación, originada por una serie de factores que influyen de igual forma para todas las mediciones procedentes de la misma unidad. Esto es esperable ya que las unidades de muestreo o de gestión pueden presentar diferencias en cuanto a calidad de estación, orientaciones, pendientes, tipo de masa (masa pura o mixta; monte bajo, medio o alto; clase natural de edad), etc.... Igualmente, cabe esperar que esta diferenciación sea especialmente notable en el caso de masas ordenadas o en el caso de masas con alta heterogeneidad espacial en la distribución de las distintas clases de edad. Por tanto, parece razonable considerar la potencial variabilidad entre sub unidades cuando se establecen modelos de regresión para predecir existencias en un monte usando datos LiDAR.

Si no se considera la variabilidad entre unidades de muestreo a la hora de construir los modelos de regresión, y se establece un modelo general conjunto para toda el área de estudio (monte), se corre el riesgo de que la variabilidad entre las unidades de muestreo sea significativa, y que las predicciones para cada una de ellas (basadas en un modelo general para todo el monte), estén afectadas por importantes problemas de sesgo, tal como se muestra en la Figura 1.

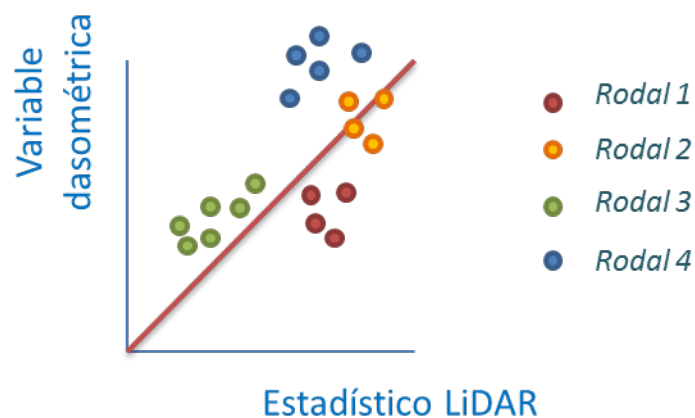


Figura 1. Representación simplificada e ilustrativa del posible sesgo derivado de construir un modelo general para todo el monte. Al realizar estimaciones a nivel de subunidad (rodal) con dicho modelo se producirán importantes errores de sesgo por exceso (rodales 1 y 2) y por defecto (rodales 3 y 4)

Una posible solución sería construir modelos diferenciados para cada una de las subunidades para las que se desea obtener estimaciones. Sin embargo, esta no parece ser una solución factible cuando se consideran unidades de gestión como rodales o cantones, debido al bajo número de parcelas que usualmente se inventarían en cada una de ellas. Aumentar el número de parcelas inventariadas en estas subunidades no es una solución factible en la mayoría de los casos debido al alto coste que tienen los trabajos de campo. Dado que cantones y rodales contienen normalmente un bajo número de parcelas, puede decirse que estas unidades respondan a la definición de área pequeña (RAO & MOLINA, 2015). Una de las principales alternativas a la hora de generar estimaciones para áreas pequeñas es el empleo de estimadores EBLUP (Empirical Best Linear Unbiased Predictor) basados en modelos mixtos, ya que permiten mejorar la estimación de variables forestales a nivel de cantón o rodal y que incorporan la variabilidad entre subunidades sin necesidad de tener que construir un modelo para cada una de ellas.

2. Objetivos

El objetivo general de este trabajo es comparar la precisión de las estimaciones para subunidades de gestión empleando:

- Solo datos de campo (Inventario tradicional).
- Datos de campo, información auxiliar LiDAR y modelos de regresión considerando la posible variabilidad entre las distintas subunidades para las que se generan estimaciones.

3. Metodología

3.1. Áreas de estudio

El trabajo se ha desarrollado en dos zonas de estudio diferentes de forma que se pudiesen analizar las mejoras que la incorporación de datos LiDAR puede ofrecer en inventarios realizados en condiciones diferentes.

3.1.1. Valle de las Caderechas (Burgos)

La primera zona de estudio comprende los términos municipales de Rucandio, Aguas Cándidas y Cantabrana, correspondiéndose con un grupo amplio de Montes de Utilidad Pública (nº 67, 68, 95, 96, 97, 514, 519, 524, 592, 593, 599 y 699 del Catálogo de Burgos) que suman un total de 2.336,44 ha. Se trata de un valle bastante cerrado, con pendientes fuertes, que discurre aproximadamente en dirección NO-SE. La elevación oscila entre los 700 m y los 1.170 m sobre el nivel del mar.

La vegetación está dominada por pinares de *Pinus pinaster* Aiton, salpicados de un latizal de *Quercus faginea* Lam, y en algunas ocasiones otras especies de frondosas. La mayor parte de los pies de *Pinus pinaster* Aiton, tienen diámetros normales comprendidos entre los 10 y los 50 cm con troncos sinuosos. La distribución general de pies puede considerarse multidiamétrica. Los árboles menores de 10 cm son fundamentalmente frondosas (quejigo/encina) y apenas aparecen pies de grandes dimensiones (diámetros normales > 50 cm). Además de estas formaciones, en la zona aparecen en una proporción mucho menor, zonas de latizal del *Pinus pinaster* Aiton., zonas de pinar de *Pinus nigra* J.F. Arnold. y *Pinus sylvestris* L. así como manchas de *Quercus faginea* Lam., *Quercus ilex* L., mezclas de ambas especies y formaciones de ribera. La extensión de cada una de estas zonas es bastante reducida por lo que fueron eliminadas del presente estudio. El área final considerada en el análisis es de 1.365,5 ha.

En esta zona, se dispone de un vuelo LiDAR cuya densidad se homogeneizó a LiDAR a 2 puntos/m². Esta homogenización fue necesaria ya que el relieve en esta zona es muy abrupto y a que el vuelo se hizo a una altura constante, provocando grandes diferencias de densidad de puntos entre las zonas altas y las zonas bajas del valle. Una vez homogeneizada la densidad de puntos se procedió a calcular los tradicionales predictores LiDAR.

En total se midieron 202 parcelas de campo distribuidas de forma sistemática sobre el área de estudio. Las parcelas tienen 10 metros de radio y se georreferenciaron empleando un GPS diferencial de fase. En cada parcela se midió el diámetro y la altura de cada árbol. Empleando las ecuaciones de cubicación para la zona del inventario forestal nacional se obtuvieron los volúmenes de cada árbol de la parcela. Finalmente, para cada parcela se calculó la densidad (número de pies/ha), el volumen (m³/ha) y el área basimétrica (m²/ha).

Los montes del área de estudio están divididos en un total de 212 cantones cuyo tamaño medio es 11 ha. Dado el elevado número total de rodales y su pequeño tamaño es imposible analizar la variabilidad que existe entre este tipo de unidades. Como solución de compromiso, se procedió a agrupar los rodales en un total de 54 unidades homogéneas. Estas unidades son las áreas de interés para las que se generaron estimaciones en esta área de estudio.

3.1.2. Valle de la Fuenfría (Madrid)

El área de estudio se encuentra en el Valle de la Fuenfría, en el término municipal de Cercedilla (Madrid) dentro del M.U.P. nº 32 "Pinar y Agregados" de Madrid. Se trata de un valle con orientación Norte - Sur y cuyo intervalo altitudinal oscila entre los 1.250 metros sobre el nivel del mar y los 2.050 metros. Es una zona de 147,5 ha de pinar prácticamente monoespecífico de *Pinus sylvestris* donde esporádicamente aparecen individuos de *Quercus pyrenaica*.

Durante Julio de 2011, se realizó un vuelo LiDAR en la zona empleando un sensor ALS70 HP, con una configuración de vuelo que dio lugar a una densidad media de 11,6 puntos/m².

Se dispone de datos de campo de dos campañas de medición realizadas en 2013 y 2015 respectivamente. En total, se dispone de 140 parcelas circulares que se georreferenciaron empleando un GPS diferencial de fase. El protocolo de medición fue similar al empleado en Burgos, si bien en este caso, solo se midieron las alturas de 7 árboles por parcela. Las

mediciones de alturas se emplearon para ajustar un modelo Altura-Diámetro que posteriormente se empleó para estimar la altura de cada uno de los árboles para los que solo se había medido el diámetro. Durante la campaña de 2015, se obtuvieron barrenas de crecimientos que se emplearon para modelizar el crecimiento ocurrido en los dos últimos años y así sincronizar ambas mediciones a 2013.

Finalmente, en esta zona se delimitaron un total de 6 unidades equiparables a rodales. Estas 6 unidades son las áreas para las que se obtuvieron estimaciones. Es importante darse cuenta que estas unidades son equiparables a cantones, si bien la densidad de parcelas de inventario es muy superior a la que se emplearía en aplicaciones prácticas para obtener estimaciones para unidades similares.

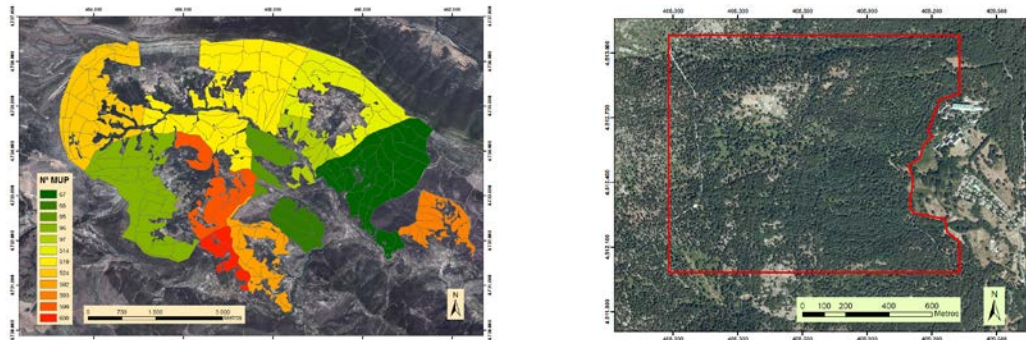


Figura 2. Zonas de estudio: Valle de las Caderechas en Burgos (izda.) y ladera de la Peña del Águila en el Valle de la Fuenfría, Cercedilla, Madrid (dcha.)

3.2. Inventario tradicional

Para cada una de las áreas de interés en cada una de las zonas de estudio se obtuvieron:

1. La media muestral de las parcelas para cada una de las variables de interés: \bar{y}
2. La desviación típica de cada variable de interés: σ_y
3. El coeficiente de variación de la media muestral: $CV_{tradicional} = \frac{\sigma_y/\sqrt{n}}{\bar{y}}$
4. Intervalos de confianza para la media muestral: $IC_{tradicional} = \bar{y} \pm 1,96 \cdot \sigma_y/\sqrt{n}$

3.3. Inventario LiDAR

La estimación con datos LiDAR se basó en suponer que la variable de interés sigue un modelo cuya expresión general es:

$$y_{i,j} = \boldsymbol{\beta} \cdot \mathbf{x}_{ij} + v_j + \varepsilon_{i,j}$$

Donde $y_{i,j}$ y $\mathbf{x}_{ij} = (1, x_{1,ij}, \dots, x_{p,ij})$ representan respectivamente el valor de la variable de interés y de los p predictores LiDAR en el j -ésimo pixel del área de interés i . $\boldsymbol{\beta}$ es un vector de dimensión $p+1$ de coeficientes de regresión. Los efectos aleatorios de las subunidades v_j se asume que tienen una distribución normal con media 0 y desviación típica σ_v^2 . Los errores del modelo se asumen que tienen una distribución normal con media cero pero que su varianza puede no ser constante. Se consideraron tres posibles modelos para la varianza del error:

- Varianza error constante: $\varepsilon_{i,j} \sim N(0, \sigma_e^2)$

- Varianza error proporcional: $\varepsilon_{i,j} \sim N(0, \sigma_e^2 \cdot X_i)$
- Varianza error exponencial: $\varepsilon_{i,j} \sim N(0, \sigma_e^2 \cdot e^{2 \cdot \delta \cdot X_i})$

Dado el elevado número de predictores LiDAR para cada área de estudio y variable considerada, se llevó a cabo un procedimiento de selección de modelos con las siguientes etapas: En primer lugar, se ajustaron los 5 mejores modelos con 1, 2, y hasta 3 variables. De este conjunto de 15 modelos se filtraron los aquellos que no cumplían las hipótesis iniciales del modelo general de regresión, es decir: (i) Ausencia de multicolinealidad: se aceptaron aquellos modelos en los que el Factor de Inflado de la Varianza era menor de 10. (ii) Independencia de los residuos: se aceptaron aquellos modelos en los que el p-valor de Durbin-Watson era mayor de 0,05. (iii) Normalidad de los residuos evaluada mediante el test de Shapiro-Wilk: p-valor mayor de 0,05. Después de estas consideraciones, se utilizó el BIC como criterio fundamental para elegir el mejor modelo.

Una vez que se eligió el mejor modelo lineal para cada variable dasométrica, se usaron las mismas variables independientes para ajustar un modelo lineal mixto (con efectos aleatorios para las zonas de interés) con ordenada en el origen aleatoria. En Burgos, se consideraron además modelos mixtos con ordenada en el origen y pendiente aleatoria. Se compararon los modelos de efectos fijos y los modelos mixtos correspondientes a una misma variable dasométrica empleando un test de razón de verosimilitudes y el mejor modelo de acuerdo a este test fue seleccionado para cada una de las variables.

Por tanto, en el caso de aquellas variables para las cuales el modelo lineal de efectos fijos fue significativamente mejor, la expresión resulta de eliminar el término v_j del modelo. Es decir, el modelo queda de la forma:

$$y_{i,j} = \beta \cdot x_{ij} + \varepsilon_{i,j}$$

Finalmente se procedió a elegir el mejor modelo para la varianza del error. Para cada uno de los modelos seleccionados en el paso anterior, se obtuvieron los modelos con varianza de los errores constante, proporcional y exponencial. Estos tres modelos se compararon entre si empleando como criterio estadístico el AIC. El modelo con mejor AIC fue finalmente seleccionado para generar las estimaciones de cada variable.

Una vez seleccionado el modelo en el que se basó la estimación para cada una de las variables se procedió a calcular los siguientes parámetros:

1. EBLUP o valor estimado para cada área de interés: \hat{y}_{LIDAR}
2. El error cuadrático medio del EBLUP: $RMSE_{LIDAR}$ (RAO & MOLINA, 2015)
3. El coeficiente de variación del EBLUP: $CV_{LIDAR} = \frac{RMSE_{LIDAR}}{\hat{y}_{LIDAR}}$
4. Intervalos de confianza para el EBLUP: $IC_{LIDAR} = \hat{y}_{LIDAR} \pm 1,96 \cdot RMSE_{LIDAR}$

Es importante tener en cuenta que cuando los efectos aleatorios de las áreas de interés no son significativos el EBLUP y su RMSE se simplifican notablemente. En BREIDENBACH et al. (2016) se pueden consultar las expresiones del EBLUP y su RMSE en el caso de que los efectos aleatorios de las áreas de interés no sean significativos.

Finalmente, el aumento en la precisión de las estimaciones basadas en LiDAR con relación a las estimaciones realizadas empleando solo datos de campo se midió a través del parámetro Δ_{eff} .

$$\Delta_{eff} = \frac{CV_{tradicional} - CV_{LiDAR}}{CV_{tradicional}}$$

4. Resultados

Los mejores modelos seleccionados han sido el modelo lineal mixto con ordenada en el origen aleatoria (para las variables G en Burgos y las variables V y N en Madrid) y el modelo lineal de efectos fijos (para las variables V y N en Burgos y G en Madrid). Los parámetros de modelos elegidos se muestran en la Tabla 1.

Se ha comparado gráficamente la utilización del modelo lineal normal con el modelo mixto para aquellas variables en las que el modelo mixto supone una mejora en la estimación. Tal como se muestra en la Figura 3, en el caso del modelo mixto, las observaciones (parcelas de muestreo) correspondientes a un rodal se acercan más a los valores predichos.

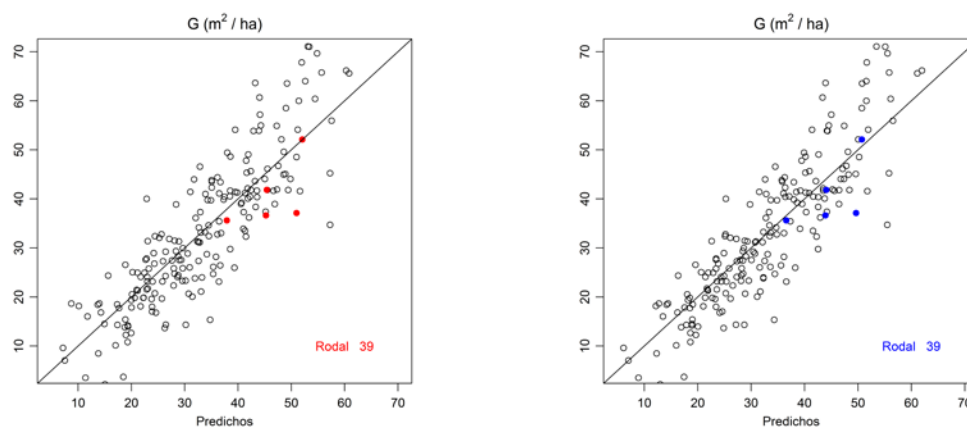
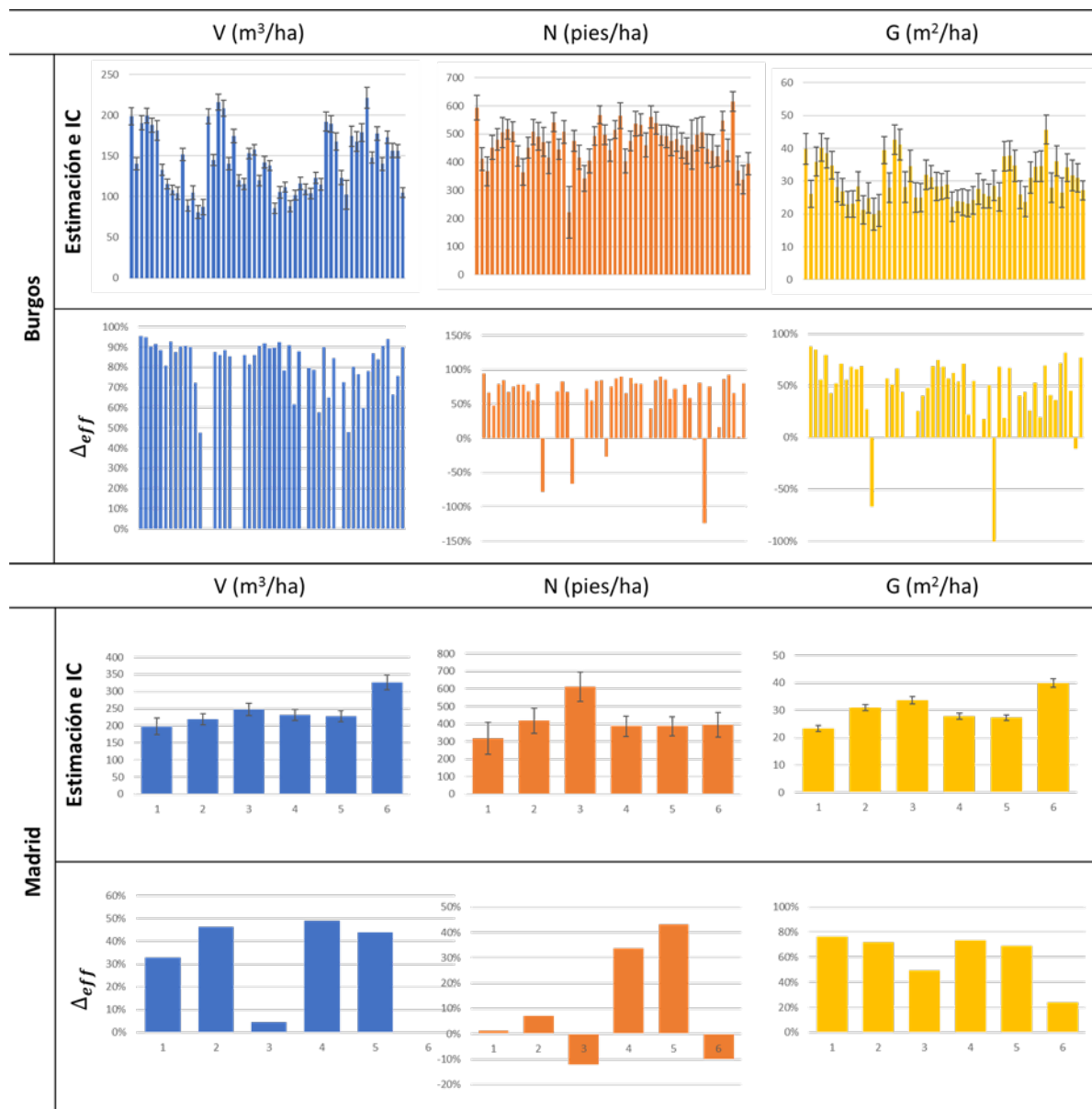


Figura 3. Ejemplo de un rodal de Burgos en los que el modelo mixto (dcha.) supone una mejora frente al modelo que sólo incorpora el efecto fijo (izda.) para la estimación del área basimétrica

1 Tabla 1. Modelos seleccionados para las diferentes variables dasométricas.

Zona	Variable	Modelo	Coefficientes	Estimaciones	σ_v	σ_e	$var(\varepsilon_{ij})$
Burgos	V (m ³ /ha)	$y_{i,j} = \beta \cdot x_{ij} + \varepsilon_{i,j}$	Intercept	-61,80			$\sigma^2 \cdot e^{0,36 \cdot Elv_mean}$
			Elv_mean	46,46			
			Elv_P10	-308,77			
	N (p/ha)	$y_{i,j} = \beta \cdot x_{ij} + \varepsilon_{i,j}$	Intercept	552,43			$\sigma^2 \cdot e^{0,008 \cdot PercR1_Abv36}$
			PercR1_Abv36	10,08			
			Elv_P10	-1191,91			
	G (m ² /ha)	$y_{i,j} = \beta \cdot x_{ij} + v_j + \varepsilon_{i,j}$	Intercept	11,29			$\sigma^2 \cdot Elv_mean$
			Elv_mean	4,83	2,36	3,51	
			Elv_AAD	7,12			
Madrid	V (m ³ /ha)	$y_{i,j} = \beta \cdot x_{ij} + v_j + \varepsilon_{i,j}$	Intercept	-69,28			$\sigma^2 \cdot Elv_P40^{2,5}$
			Elv_P40	18,68	13,68	2,28	
			PecRTot_AbvMed	3,68			
	N (p/ha)	$y_{i,j} = \beta \cdot x_{ij} + v_j + \varepsilon_{i,j}$	Intercept	72,77			$\sigma^2 \cdot PecR1_Abv3^{2,5}$
			Elv_P70	-22,45	55,96	0,87	
			PecRe1_Abv3	9,38			
	G (m ² /ha)	$y_{i,j} = \beta \cdot x_{ij} + \varepsilon_{i,j}$	R1_AbvMed	0,01			$\sigma^2 \cdot Elv_P25$
			Intercept	-3,39			
			Elv_P25	1,17			
			PecRTot_AbvMed	0,77			

2



3

4

Figura 4. Predicciones (EBLUPs) e intervalos de confianza para las estimaciones en los diferentes rodal

5 El principal resultado del presente trabajo (Figura 4) es que el empleo de datos LiDAR
 6 permite reducir de forma significativa los errores de estimación para áreas de interés como
 7 rodales o grupos de rodales. La reducción de la incertidumbre medida a través del parámetro Δ_{eff}
 8 alcanza para algunas variables valores próximos al 90%. Los valores medios del parámetro Δ_{eff}
 9 en Burgos, son 82,23% para V, 58,84% para N y 38,59% para G. Para Madrid estos valores fueron
 10 29,43%, 10,53% y 57,53%. La menor reducción de la incertidumbre en Madrid se explica por el
 11 elevado número de parcelas empleado en esta zona. Es importante observar, que Δ_{eff} no siempre
 12 mejora, es decir, en áreas de interés muy homogéneas, la estimación con datos de campo puede
 13 ser mejor. En estos casos, podría reemplazarse la estimación basada en datos LiDAR por la
 14 estimación de campo sin causar perjuicio.

15

16

17 Atendiendo solo a los errores de estimación obtenidos con cada una de las dos técnicas
 18 empleadas puede decirse que, en el caso de Burgos (Tabla 2). La incorporación de información
 19 auxiliar LiDAR permite reducir los errores relativos promedio para las áreas de interés consideras
 de un 40,55% a un 6,06% en promedio para V. Para G los errores relativos del inventario



20 tradicional son en promedio del 33,93% mientras que la incorporación de la información auxiliar
 21 LiDAR reduce el promedio del error del error relativo al 15,09%. Para N la estimación basad en
 22 datos de campo tiene un error relativo medio de 35,81% mientras que empleando datos LiDAR el
 23 error relativo medio es del 15,09%. Es importante observar, como en este caso, con una densidad
 24 de muestreo de una parcela cada 6,75 ha la incorporación de información auxiliar LiDAR permite
 25 obtener precisiones en torno al 15% o mejores, de forma que permitirían alcanzar las precisiones
 26 requeridas en ciertos trabajos de inventario con un menor esfuerzo de trabajo en campo.

27
 28 En Madrid (Tabla 2), al incorporar información auxiliar LiDAR, el error relativo medio para las áreas
 29 de interés consideradas pasa de un 12,07% a un 7,93% para V, de un 20,42% a un 17,45% para
 30 N y de un 12,01% a un 3,95% para G. En este caso, salvo para N, la gran densidad de parcelas
 31 empleadas permite que sólo empleando datos de campo se obtengan precisiones aceptables para
 32 aplicaciones prácticas. Sin embargo, no es común disponer de esta densidad de parcelas tan
 33 elevada. Se comprueba, que la incorporación del LiDAR mejora la estimación de modo que en
 34 escenarios como el de Madrid (Inventarios para pequeños montes) el LiDAR permite reducir el
 35 trabajo de campo. Incorporar información LiDAR supone por otro lado un aumento de los trabajos
 36 de gabinete, de modo que para este tipo de aplicaciones es cuestionable si la mejora en la
 37 estimación o la posible reducción del trabajo de campo que podría obtenerse al usar datos LiDAR
 38 compensa el mayor tiempo de procesado y análisis que debe realizarse en oficina.

40 *Tabla 2. Aumento promedio de la eficiencia al incorporar datos LiDAR y errores relativos medios al emplear inventario*
 41 *tradicional o inventario LiDAR.*

Zona	Variable	Promedio Δ_{eff}	Promedio Error relativo tradicional	Promedio Error relativo LiDAR
Burgos	V (m ³ /ha)	82,23%	40,55%	6,06%
	N (p/ha)	57,94%	35,18%	9,99%
	G (m ² /ha)	38,59%	33,93%	15,09%
Cercedilla	V (m ³ /ha)	29,43%	12,07%	7,93%
	N (p/ha)	10,53%	20,42%	17,65%
	G (m ² /ha)	57,53%	12,01%	3,96%

42 5. Discusión

43 Sin duda, una de las mayores dificultades de la inventariación con LiDAR es la estimación en
 44 superficies pequeñas (cantones, rodales) y el cálculo de su error asociado. Esto se debe a que el
 45 modelo se ajusta para la totalidad de la zona de estudio y luego se desea predecir a una escala
 46 inferior. La solución más inmediata sería la construcción de modelos de regresión diferentes para
 47 cada subunidad (cantón o rodal, dependiendo de la precisión que se desee). Sin embargo, esto no
 48 suele ser factible debido a que no se posee un número suficiente de parcelas que permita ajustar
 49 modelos de regresión.

50
 51 El empleo de modelos mixtos y estimadores EBLUP aparecen en la literatura como una de las
 52 principales alternativas para afrontar este problema (RAO, 2015). Distintos autores (BREIDENBACH
 53 & ASTRUP, 2012; GOERNDT et al., 2011; MAURO et al., 2016) han comprobado la utilidad del
 54 EBLUP en trabajos de inventario forestal asistidos con información de sensores remotos. Los
 55 resultados obtenidos en el presente estudio también indican que el empleo del EBLUP para la
 56 estimación de variables forestales asistida con datos LiDAR es la mejor alternativa cuando se
 57 necesitan obtener estimaciones a nivel de subunidad (área de interés). El empleo de estimadores
 58 sintéticos de regresión, puede acarrear importantes problemas de sesgo, especialmente cuando la
 59 variabilidad entre subunidades es elevada.

63 Curiosamente, aunque uno de los principales estudios sobre estimación en áreas pequeñas,
64 en el que se detalla cómo obtener estimadores basados en modelos mixtos procede del campo de
65 la teledetección (BATTESE et al., 1988), la aplicación de técnicas de estimación para áreas
66 pequeñas en inventarios forestales asistidos con información auxiliar procedente de sensores
67 remotos, ha sido muy reducida e intermitente (BREIDENBACH. et al., 2008; NOTHDURFT et al.,
68 2009; GOERNDT et al., 2011; BREIDENBACH & ASTRUP, 2012).

69
70 Sin embargo, pese a que una gran cantidad de técnicas de estimación para áreas pequeñas
71 basadas en modelos mixtos eran conocidas a finales de los años 90, el primer estudio, completo,
72 en el que se aplicaron estas técnicas para predecir variables forestales a partir de datos ALS fue
73 realizado en 2011 (GOERNDT et al., 2011). Hasta dicha fecha, prácticamente todos los trabajos en
74 los que se han obtenido estimaciones para variables forestales a partir de datos ALS, se han
75 realizado empleando estimadores sintéticos basados en modelos de regresión, en los que no se
76 consideraban efectos aleatorios a nivel de subpoblación.

77
78 Los modelos mixtos son una generalización del modelo lineal de regresión clásico que
79 contempla la posible existencia de observaciones correlacionadas o con variabilidad heterogénea,
80 vinculadas a la presencia de factores aleatorios tanto a nivel de rodal como a nivel de píxel. Por
81 tanto, modelizan efectos aleatorios que responden a una organización jerárquica de la población
82 (rodales, píxeles). La varianza del efecto aleatorio a nivel de subunidad recoge la variabilidad entre
83 las subunidades, mientras que la varianza del efecto aleatorio a nivel de píxel recoge la
84 variabilidad dentro de cada subunidad no explicada por el modelo. De esta forma, permite tener en
85 cuenta la variabilidad que puede existir entre áreas de interés (rodales, cantones...), derivada del
86 hecho que existan diferentes calidades de estación o tipos de masa que no sea posible capturar
87 únicamente con la variación propia de los estadísticos del LiDAR.

88
89 En este sentido, hay que resaltar que los resultados de este trabajo indican que esta
90 variabilidad no siempre es significativa, como ha resultado ser para las variables V y N en Burgos y
91 G en Madrid. Pero hay otras variables en las que el modelo sí supone una mejora significativa (G
92 en Burgos y V y N en Madrid), por lo que de forma general se aconseja comprobarlo siempre a la
93 hora de abordar el inventario forestal LiDAR.

94
95 Un aspecto novedoso del estudio realizado es que se emplean modelos mixtos a nivel de
96 elemento poblacional con heterocedasticidad. Cuando se estiman variables dasométricas, suele
97 ocurrir, que la varianza residual aumenta con el valor de la variable a estimar. La alternativa
98 generalmente empleada para abordar la heterocedasticidad, ha sido emplear transformaciones de
99 la variable respuesta (ANDERSEN et al., 2005; GONZÁLEZ-FERREIRO et al., 2012). Aunque el
100 empleo de transformaciones de la variable respuesta es una práctica habitual, generalmente
101 conlleva una pérdida de linealidad que dificulta obtener estimaciones por agregación de los
102 valores asociados a los píxeles. En el contexto de la estimación en áreas pequeñas, (Slud and
103 Maiti, 2006) han tratado este problema. En dicho estudio se contempla la estimación basada en
104 modelos tipo Fay-Herriot en los que la variable respuesta ha sido transformada Logarítmicamente.
105 El estudio de Molina, (2009), considera el caso general de obtener estimadores insesgados
106 empleando modelos mixtos de regresión multivariante con transformación logarítmica de la variable
107 respuesta, así como una aproximación al MCP (mean crossed product error) de dichos estimadores
108 del que se deriva, como caso particular el MSE.

109
110 Por último, se debe reseñar que la comparación entre los errores del inventario tradicional y
111 los errores del inventario LiDAR es difícil debido a que son técnicas distintas (muestreo vs
112 regresión) (SÄRNDAL et al., 1978). Las regresiones se basan en hacer estimaciones a partir de
113 modelos que, una vez validados, se asumen como correctos para una determinada población
114 (*model-based approach*). Por su parte, en el muestreo (*design-based approach*) no se asume
115 ningún modelo para la población; en este caso lo que importa es la forma de extraer la muestra del
116 total poblacional. Sin embargo, a través del parámetro Δ_{eff} se ha podido cuantificar la reducción



117 de los errores de estimación del inventario LiDAR frente al inventario tradicional a nivel de áreas
 118 de interés (rodales o grupos de rodales).

119

120 6. Conclusiones

121

122 Las principales conclusiones de este trabajo son:

123

124 1. El empleo de estimadores EBLUP permite usar información auxiliar LiDAR en inventario
 125 forestal, haciendo posible obtener estimaciones de variables forestales insesgadas y de
 126 gran precisión en unidades de gestión equiparables a cantones.

127 2. La variabilidad entre áreas de interés no siempre es significativa y en ocasiones no es
 128 necesario emplear modelos mixtos. En cualquier caso, el uso de modelos de efectos fijos
 129 deber realizarse sobre la base de que no existe variabilidad entre áreas de interés.

130

131 7. Bibliografía

132

133 ANDERSEN, H. E.; MCGAUGHEY, R. J.; REUTEBUCH, S.E. 2005. Estimating forest canopy fuel
 134 parameters using LIDAR data. *Remote Sens. Environ.* 94, 441-449.

135

136 BATTESE, G.E.; HARTE, R.M.; FULLER, W.A. 1988. An Error-Components Model for Prediction of
 137 County Crop Areas Using Survey and Satellite Data. *J. Am. Stat. Assoc.* 83, 28-36.

138

139 BREIDENBACH, J.; KUBLIN, E.; MCGAUGHEY, R.; ANDERSEN, H.E.; REUTEBUCH, S. 2008. Mixed-
 140 effects models for estimating stand volume by means of small footprint airborne laser scanner
 141 data. *Photogramm. J. Finl.* 21, 4-15.

142

143 BREINDENBACH, J.; ASTRUP, R. 2012. Small area estimation of forest attributes in the Norwegian
 144 National Forest Inventory. *European Journal of Forest Research.* 1-13.

145

146 BREINDENBACH, J.; MCROBERTS, R.E.; ASTRUP, R. 2016. Empirical coverage of model-based
 147 variance estimators for remote sensing assisted estimation of stand-level timber volume. *Remote
 148 sensing of environment.* 173: 274-281.

149

150 GOERNDT, M.E.; MONLEON, V.J.; TEMESGEN, H. 2011. A comparison of small-area estimation
 151 techniques to estimate selected stand attributes using LiDAR-derived auxiliary variables. *Canadian
 152 Journal of Forest Research.* 41, 1189-1201.

153

154 GONZÁLEZ-FERREIRO, E.; DIÉGUEZ-ARANDA, U.; MIRANDA, D.; 2012. Estimation of stand variables
 155 in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry.*

156

157 MAURO, F.; MOLINA, I.; GARCÍA-ABRIL, A.; VALBUENA, R.; AYUGA-TÉLLEZ, E. 2016. Remote sensing
 158 estimates and measures of uncertainty for forest variables at different aggregation levels.
 159 *Environmetrics.* 27: 225-238.

160

161 NOTHDURFT, A.; SABOROWSKI, J.; BREIDENBACH, J. 2009. Spatial prediction of forest stand
 162 variables. *European Journal of Forest Research.* 128, 241-251.

163

164 RAO, J.N.K.; MOLINA, I. 2015. *Small Area Estimation.* Wiley. Hoboken, New Jersey.

165

166 SÄRNDAL, C.E.; THOMSEN, I.; HOEM, J.M.; LINDLEY, D.V.; BARNDORFF-NIELSEN, O.; DALENIUS, T.,
 167 1978. Design-Based and Model-Based Inference in Survey Sampling. *Scand. J. Stat.* 5, 27-52.