



7º CONGRESO FORESTAL ESPAÑOL

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

26 - 30 junio 2017 | Plasencia
Cáceres, Extremadura

7CFE01-086

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

© Sociedad Española de Ciencias Forestales

Integrando datos LiDAR, información satelital y parcelas del Inventario Forestal Nacional español en la predicción de variables de inventario

FERNANDEZ-LANDA, A.¹, TOMÉ MORÁN, J.L.¹, SANDOVAL ALTELARREA V.J.² y VALLEJO BOMBIN, R.²

¹ Agresta S. Coop. C/ Duque de Fernán Núñez, 2, 1º 28012. Madrid.

² Ministerio de Agricultura, Alimentación y Medio Ambiente, Calle Gran Vía de san Francisco 4, 28005 Madrid

Resumen

Uno de los mayores retos del Inventario Forestal Nacional Español es incorporar de forma progresiva nuevas tecnologías sin perder la información histórica acumulada ni su esencia de “inventario continuo en el tiempo”. El presente estudio combina datos LiDAR del Plan Nacional de Ortofotografía Aérea, datos de campo propios del IFN4 e imágenes satelitales de media resolución para estimar de forma espacialmente continua las principales variables de inventario en masas puras y mixtas de *Pinus sylvestris* L. y *Fagus sylvatica* L. en La Rioja. Los modelos ajustados de área basimétrica y volumen poseen errores similares a los publicados por otros autores en áreas menores utilizando vuelos y parcelas específicas. Sin embargo, la variable número de pies presenta mayores errores que los encontrados en la literatura. El estudio muestra como la utilización de información ya capturada supone una oportunidad para generar inventarios forestales continuos a escala regional reduciendo de forma drástica los costes de los inventarios forestales de alta resolución. Esta metodología podría tener gran aplicabilidad tanto en España como en otros países con planes nacionales de capturas de información LiDAR e inventarios forestales nacionales que trabajan con grandes bases de datos de parcelas de campo.

Palabras clave

Landsat, modelos, Random Forest, GLM

Introducción

Nuestro país es pionero en la captura de información LiDAR a escala nacional y España ya dispone de una primera cobertura LiDAR gracias a la incorporación de esta tecnología dentro del Plan Nacional de Ortofotografía Aérea (PNOA). Esta información, a pesar de que no se captura con fines exclusivamente forestales, aporta un gran valor nuestro sector ya que permite describir de forma espacialmente continua y con elevado grado de detalle la estructura del bosque.

Además, el volumen de información satelital disponible de forma abierta está creciendo de manera exponencial. En 2008 se abre el archivo de imágenes Landsat que en la actualidad dispone de más de un millón de escenas desde 1972. La Agencia Espacial Europea en el ámbito de la Misión Copernicus ha puesto en órbita en los últimos 2 años varios satélites cuyas imágenes están disponibles de forma abierta en internet. Las nuevas misiones satelitales que actualmente se encuentran en desarrollo, algunas de ellas diseñadas exclusivamente para medir los bosques (como la misión Biomass de la ESA), van a generar a futuro todavía más volumen de información de gran valor para nuestro sector. Dentro de la información abierta y disponible a día de hoy, los citados sensores satelitales ópticos de resolución media Landsat y Sentinel-2 aportan información complementaria al LiDAR a una resolución adecuada permitiendo comprender mejor la composición y distribución espacial de las principales especies forestales.

Estas nuevas tecnologías y fuentes de información pueden suponer un complemento valioso al Inventario Forestal Nacional (IFN), permitiendo generar una cartografía de nuestras masas forestales de cada vez más detalle y resolución. El IFN, entendido como un sistema de seguimiento del estado de nuestros bosques, tiene ante sí el reto de aprovechar e incorporar estas tecnologías y fuentes de

información sin perder su capacidad de análisis de la evolución histórica de nuestros recursos forestales, ni su esencia de “inventario continuo en el tiempo”.

En este contexto, el presente trabajo desarrolla una primera experiencia de integración de vuelos LiDAR nacionales de baja densidad de puntos, imágenes satelitales Landsat y parcelas permanentes del cuarto IFN para generar cartografía de alta resolución de las principales variables de inventario forestal en masas puras y mixtas de pino silvestre y haya en La Rioja.

1. Objetivos

El objetivo principal del estudio es analizar la combinación tanto de información satelital como LiDAR con parcelas del Cuarto Inventario Forestal Nacional para estimar de forma espacialmente continua las principales variables de inventario forestal.

2. Metodología

Área de estudio e información de campo

Para generar y evaluar la cartografía de inventario forestal de alta resolución se eligieron dos áreas de piloto (con un total de 16.000 ha, Figura 1) con presencia de masas puras y mixtas de pino silvestre y haya en las Sierras de Cameros y Cebollera (municipios de Ortigosa de Cameros, El Rasillo, Villoslada de Cameros y Lumbreras).

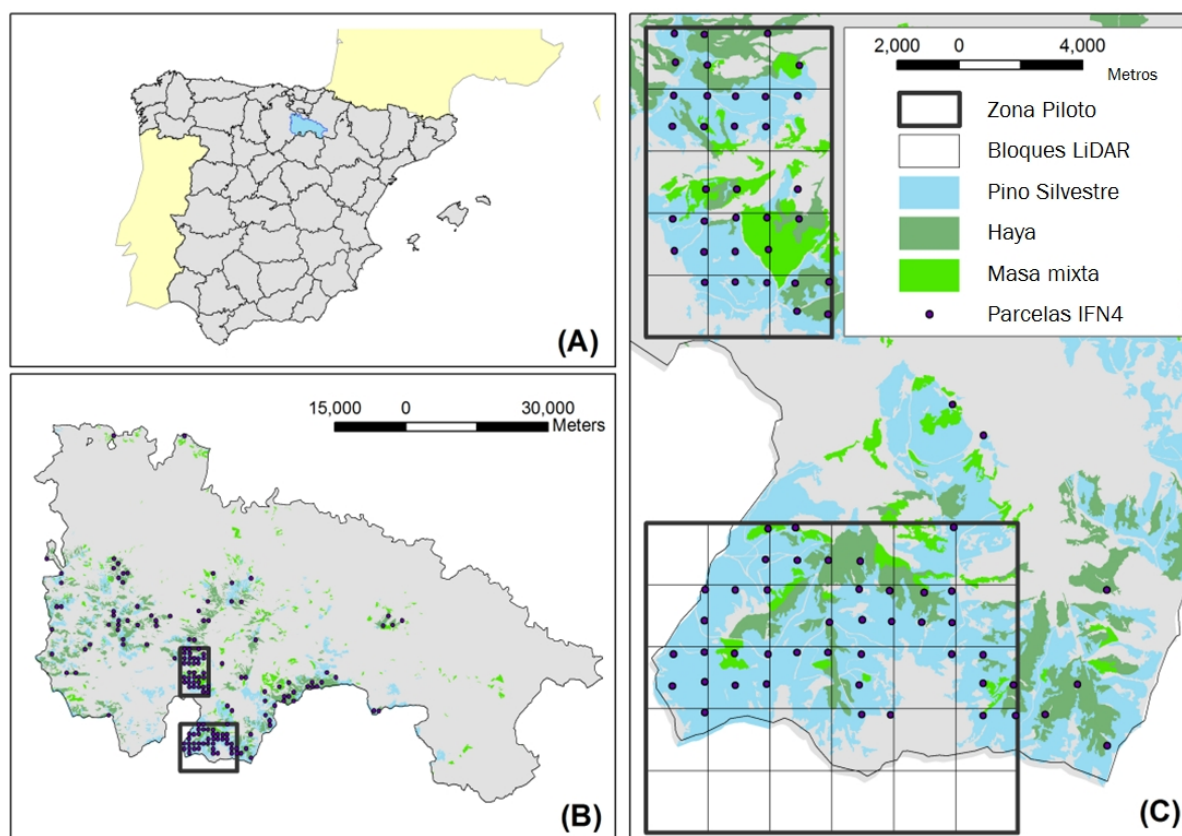


Figura1. Localización del área de estudio (A y B), de las parcelas del IFN4 seleccionadas y de las áreas piloto (B y C)

Se ha trabajado con la base de datos de parcelas de campo del 4º Inventario Forestal Nacional (IFN4) en la provincia de La Rioja. El trabajo de campo del IFN4 en esta provincia se realizó entre los meses de Junio de 2011 y Agosto de 2012. Un total de 160 parcelas con *Pinus sylvestris* L. y *Fagus*

sylvatica L. como especies dominantes fueron seleccionadas cubriendo toda la variabilidad estructural de las masas de estas especies en la provincia. Estas parcelas fueron recolocadas manualmente para mejorar sus coordenadas en base a i) la información continua de alturas de vegetación del LiDAR, ii) los croquis de las parcelas realizados por los equipos de campo, iii) la ortofotografía aérea del PNOA y iv) la altura total, especie y localización de cada uno de los pies medidos en la parcela. Las parcelas fueron desplazadas una media de 8,4 metros en el proceso de recolocación. En las últimas campañas de campo el IFN ha incorporado equipos GPS con precisión submétrica de cara a mejorar la integración de las parcelas del IFN4 y la información LiDAR.

Los pies medidos en las parcelas del IFN con una distancia mayor a 14.1 metros del centro de la parcela no fueron tenidos en cuenta con el objeto de fijar la superficie máxima de la parcela del IFN en 625 m². Esta superficie equivale a píxeles de 25 metros de lado donde son aplicados los modelos de predicción de variables. De esta forma las parcelas del IFN quedan con tres coronas (de 5, 10 y 14.1 metros), para cada uno de los pies medidos fue aplicado el factor de expansión correspondiente y referir así todas las variables a la hectárea.

La muestra de parcelas fue dividida de forma aleatoria en dos submuestras (Tabla 1), una para el ajuste de los modelos y otra para su validación. De esta forma ninguna de las parcelas utilizadas en el proceso de ajuste fue utilizada en el proceso de validación.

Tabla 1. Resumen de la base de datos de parcelas del IFN seleccionada

| Estrato | N parcelas | Tipo muestra | Estadísticos | G (m ² /ha) | V (m ² /ha) | N (pies/ha) |
|-----------|------------|---------------|--------------|------------------------|------------------------|-------------|
| Silvestre | 56 | Entrenamiento | Media | 37,79 | 255,56 | 777,97 |
| | | | Máximo | 69,63 | 526,09 | 2945,47 |
| | | | Mínimo | 12,13 | 51,90 | 176,00 |
| | | | Desv. Est. | 12,85 | 116,65 | 570,35 |
| Silvestre | 23 | Validación | Media | 38,45 | 286,47 | 586,46 |
| | | | Máximo | 83,85 | 784,92 | 1116,62 |
| | | | Mínimo | 14,14 | 75,74 | 80,00 |
| | | | Desv. Est. | 18,27 | 172,74 | 280,47 |
| Haya | 61 | Entrenamiento | Media | 31,63 | 217,72 | 826,17 |
| | | | Máximo | 53,79 | 476,22 | 2022,20 |
| | | | Mínimo | 12,88 | 52,71 | 175,83 |
| | | | Desv. Est. | 8,69 | 85,89 | 457,42 |
| Haya | 20 | Validación | Media | 29,68 | 187,37 | 989,43 |
| | | | Máximo | 42,77 | 323,51 | 1974,87 |
| | | | Mínimo | 14,68 | 58,73 | 192,00 |
| | | | Desv. Est. | 7,29 | 72,45 | 550,00 |

Datos LiDAR

Para el presente estudio se ha utilizado la información LiDAR proveniente del PNOA con una densidad media de 0,5 pulsos/m² y un error cuadrático medio en altimetría inferior a 20 cm. La campaña de vuelo en La Rioja se desarrolló entre los meses de mayo y septiembre de 2010.

La información LiDAR fue procesada con el programa FUSION (McGaughey and Carson, 2003) del Servicio Forestal de los Estados Unidos. Para las dos áreas piloto se generó un modelo digital de elevaciones (MDE) de 2 metros de resolución a partir de los puntos clasificados como suelo, el MDE fue posteriormente utilizado para la asignar alturas sobre el suelo a todos los retornos clasificados como vegetación (normalización de la nube de puntos). A partir de la nube de puntos normalizada se calcularon 13 estadísticos de estructura de masa arbolada (Tabla 2) tanto para las parcelas de campo seleccionadas como en las áreas piloto de Cebollera y Cameros de forma continua en celdas de 625 m². Los retornos de vegetación con alturas inferiores a 2 metros, normalmente correspondientes a matorral, no fueron tenidos en cuenta en el cálculo de estos estadísticos.

Tabla 2. Estadísticos calculados a partir de los retornos de vegetación por encima de 2 metros

| | | | | | | | |
|-------------------|---------------------|-------------------|---------------------------|-----------------------|--------------------------|---------------------|------------------|
| Media | Desviación estándar | Varianza | Coefficiente de variación | Rango intercuartílico | Kurtosis | Percentil del 1% | Percentil del 5% |
| Percentil del 25% | Percentil del 50% | Percentil del 75% | Percentil del 95% | Percentil del 99% | Fracción Cobida Cubierta | Ratio libre de copa | |

El Ratio Libre de copa (CRR) es un estadístico relacionado con la longitud libre de copa de la masa a partir de observaciones de las alturas (Parker y Russ, 2004) y se define como:

$$CRR = \frac{(H_{media} - H_{mínimo})}{(H_{máximo} - H_{mínimo})}$$

Información Satelital

Como fuente de información satelital se utilizó una imagen del satélite Landsat 5 correspondiente al mes de julio de 2011. Se trata por tanto de una fecha intermedia entre la captura de los datos LiDAR (de Mayo a Septiembre de 2010) y los trabajos de levantamiento de las parcelas de campo por parte de los equipos del IFN4 (realizados entre Junio de 2011 y Agosto de 2012).

La escena Landsat se encuentra geoméricamente corregida, por tanto, únicamente se realizó una validación visual de la geometría de la escena con respecto al modelo digital de superficie del LiDAR a 2 metros de resolución y la ortofotografía aérea del PNOA. Se realizó una corrección radiométrica y atmosférica (Chander et al., 2009) de la escena Landsat transformando la imagen de valores digitales (proporcional a la radiancia recibida por el sensor) en valores de reflectancia (porcentaje de radiación reflejada por la superficie).

Clasificación de imagen y mapa de Vegetación

Para la clasificación de la imagen satelital se ha utilizado Random Forest (en adelante RF), una técnica de aprendizaje automática agrupada (Breiman 2001) que se ha convertido en una técnica muy extendida en trabajos de cartografía de vegetación y de predicción de variables forestales. RF ha sido usado y recomendado para la clasificación de coberturas de la tierra por numerosos autores (p.ej. Waske y Braun 2009; Rodríguez-Galiano et al. 2012). A diferencia de otros métodos paramétricos, RF es poco restrictivo y facilita la combinación de múltiples fuentes de información (imágenes multi e hiperespectrales, datos LiDAR y RADAR, datos climáticos, modelos digitales del terreno, etc.) (Rodríguez-Galiano et al. 2012).

Para el ajuste del modelo RF de clasificación de la imagen Landsat se han digitalizado regiones de interés homogéneas en función de las clases de cobertura de la tierra consideradas (Tabla 3). Se

ha utilizado tanto la ortofotografía aérea del PNOA como la información de las parcelas de campo del 4º IFN para la digitalización de estas regiones de entrenamiento.

Tabla 3. Leyenda de la clasificación de la imagen Landsat 5

| Id clase | Clase | Id clase | Clase |
|----------|-------------------|----------|---------|
| 1 | Pino silvestre | 6 | Encinar |
| 2 | No arbolado | 7 | Urbano |
| 3 | Hayedo | 8 | Nubes |
| 4 | Agua | 9 | Sombras |
| 5 | Rebollar-robledal | | |

Se han usado 13 variables predictoras apiladas en una imagen en la clasificación. Las bandas de la 1 a la 6 corresponden con la información multispectral del satélite bandas (Blue, Green, Red, NIR, SWIR-1 y SWIR-2), como banda 7 se ha utilizado el Normalized Difference Vegetation Index (NDVI), mientras que las bandas 8-13 son derivadas del Modelo Digital de Elevaciones (Elevation, Slope, Aspect, Curvature, Plan curvature y Profile curvature).

La importancia de las variables predictoras en la clasificación ha sido estimada con el estadístico *Mean Decrease Accuracy*. El modelo RF ajustado también permite generar un raster de calidad o incertidumbre de la clasificación a la misma resolución de la imagen de partida (en este caso 30 metros). Este raster contiene el porcentaje de árboles de decisión que han votado por la clase finalmente seleccionada para cada uno de los píxeles.

Se ha estimado la precisión de usuario, la precisión del productor y la precisión total de la clasificación a partir de una muestra independiente de validación. Esta muestra de validación ha sido generada de manera aleatoria en la zona de estudio asignando a cada punto una clase de referencia en base a interpretación visual de las ortofotos del PNOA.

Ajuste y validación de modelos

Para la predicción de las principales variables forestales se han ajustado modelos lineales generalizados (GLM por sus siglas en inglés). Los GLM permiten trabajar con distintas funciones de distribución (además de la distribución normal), siendo por tanto métodos más flexibles que los modelos lineales generales. Los GLM también permiten describir relaciones no lineales entre las variables por medio de la función vínculo.

Las variables independientes ajustadas con los modelos GLM fueron área basimétrica (G), Volumen (V), número de pies (N) y los parámetros a y b que describen la distribución diamétrica por medio de una función Weibull bi-paramétrica (a_W y b_W).

Se probaron ajustes con funciones de distribución "gaussian" y "gamma" y funciones de vínculo "inverse", "log", "identity" y "power". La función de vínculo "power" se probó con valores comprendidos entre -2 y 2. Para el ajuste de los modelos lineales generalizados se utilizó el paquete de R glm (Knudson C. 2015). y glm2 (Marschner I. 2014).

Se ajustaron modelos hasta obtener en todos los parámetros un nivel de significación del 95%. Para la selección de variables se utilizó el método paso a paso por medio del paquete MASS de R (Venables & Ripley, 2002). Se evitaron modelos que tuvieran una fuerte colinealidad entre las variables predictoras, evitando seleccionar modelos que con factores de inflación de la varianza superior a 10. Para la comparación y selección de modelos se utilizaron el criterio de información de

Akaike, el coeficiente de determinación (pseudo- R^2) y los valores absolutos y relativos del sesgo, el error absoluto medio (EAM), el error cuadrático medio (RMSE). La comparación de la bondad de ajuste de los modelos a la muestra de validación independiente se realizó calculando el sesgo, EAM y RMSE en valores absolutos y relativos.

Se ajustaron modelos GLM por separado para cada estrato (hayedo y pinar) y posteriormente se ajustaron modelos utilizando toda la base de datos incluyendo el estrato como una variable categórica en el modelo.

3. Resultados

Mapa de vegetación

La precisión del mapa de vegetación obtenida en la validación independiente fue de 0,93, las clases pinar y hayedo obtuvieron una precisión de usuario de 0,97 y 0,8 respectivamente. En cuanto a la precisión de productor la clase pinar obtuvo 0,89 y la clase hayedo 0,93. La banda 7 (el índice de vegetación NDVI) posee una importancia alta en la clasificación tanto del pinar como del hayedo, siendo la banda más importante con gran diferencia sobre las demás en la clasificación del hayedo. Para la clase pinar el infrarrojo cercano (NIR) aparece con mayor importancia seguido de cerca en importancia por el infrarrojo medio (SWIR-1) y el NDVI (Figura 2).

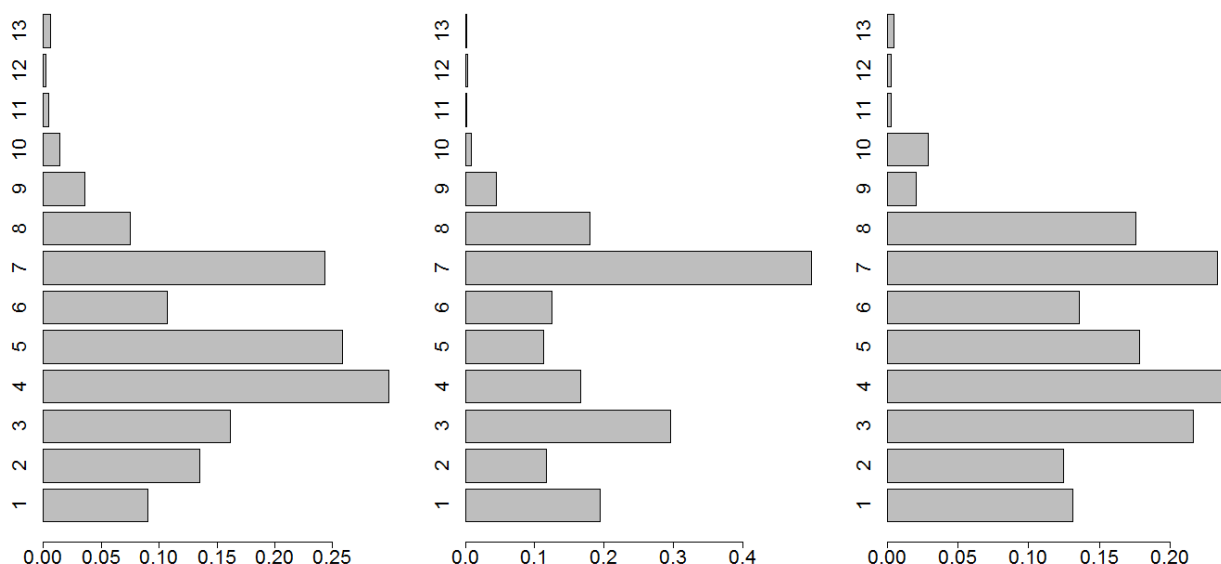


Figura 2. Importancia de las 13 bandas consideradas (eje y del gráfico) en el modelo RF. Se ha considerado el estadístico MeanDecreaseAccuracy (eje x) para la estimación de la importancia de las variables. A la izquierda la importancia de las variables para la clase pino silvestre, en el centro para la clase hayedo y a la derecha un promedio de la importancia de las variables considerando todas las clases.

La banda más importante proveniente del modelo digital de elevaciones es la elevación, la encontramos en quinto lugar en importancia cuando consideramos todas las clases.

Los resultados de la clasificación obtenidos muestran una gran coherencia con el Mapa Forestal de España (MFE). Las teselas definidas como masas mixtas de silvestre y haya en el MFE se definen como una mezcla de píxeles de estas dos especies en la clasificación de la imagen Landsat (Figura 3A y 3B) aportando información espacialmente explícita de la distribución de especies dentro de las teselas definidas como mixtas en el MFE.

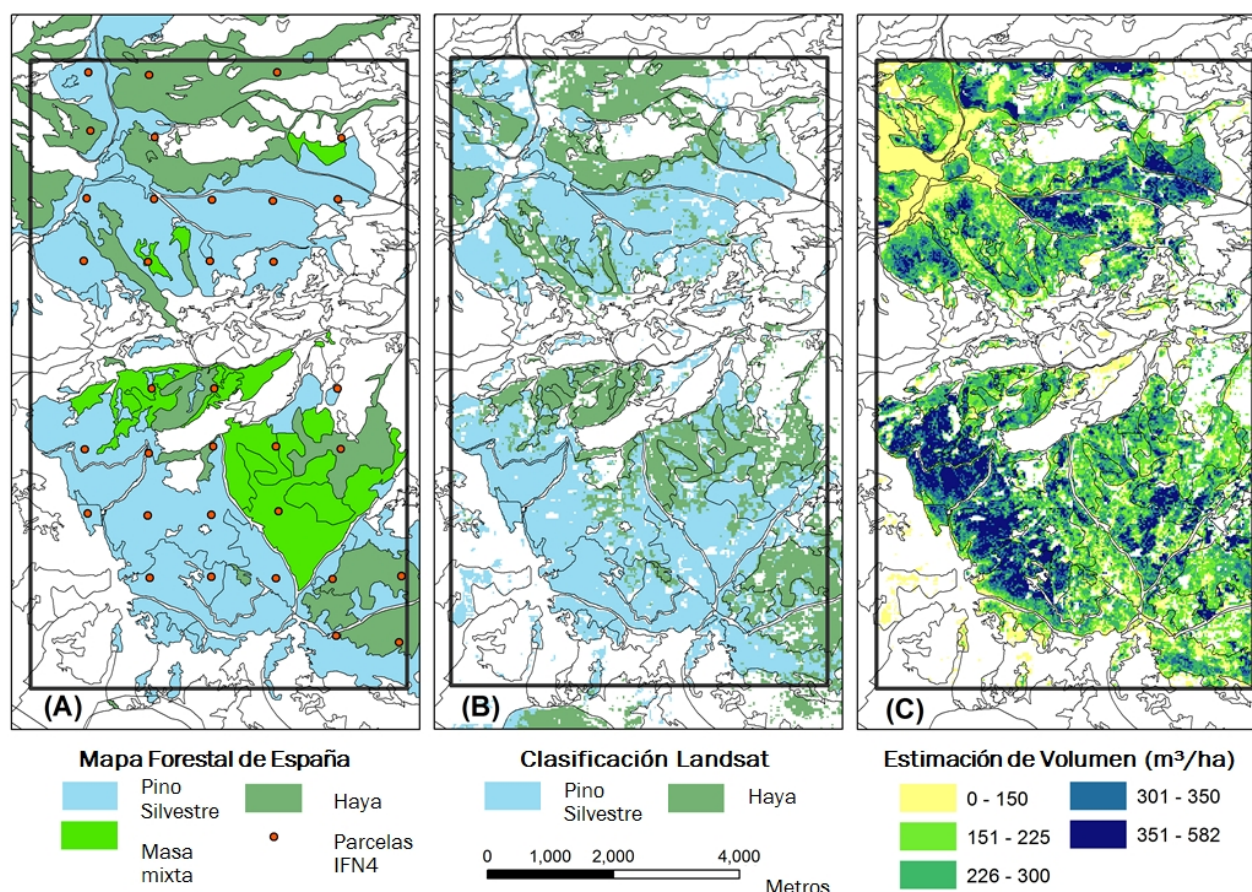


Figura 3. Mapa forestal de España para el área piloto de Cameros (A), resultados de la clasificación de Landsat (B) y de la estimación en Volumen (C) para la misma área.

Ajuste y validación de modelos

En todos los casos el criterio de información de Akaike seleccionó modelos con función de distribución “gamma” frente a “gaussian”. Los ajustes realizados en el estrato de pinar resultaron con mayores pseudo- R^2 en todas las variables ajustadas que en el caso de los ajuste de estrato de haya. Los modelos GLM para ambas especies generaron en pseudo- R^2 intermedias entre el estrato de pinar y el de hayedo. Las funciones vínculo utilizadas tras el proceso de selección de modelos fueron “power”, “identity”, “log”, “inverse” dependiendo de la base de datos y variables ajustada (Tabla 4).

Si analizamos el RMSE resultante de la fase de ajuste del modelo vemos que el estrato de pinar obtuvo mejores modelos en G y en el parámetro a de Weibull y el estrato de hayedo obtuvo RMSE más bajos en V, N y el parámetro b de Weibull.

En el caso de los ajustes GLM realizados con toda la base de datos la variable categórica “estrato” fue significativa para todos los modelos ajustados (Tabla 4), lo que refuerza la importancia de una buena estratificación por especies en inventarios forestales con LiDAR.

La validación de los modelos con la muestra independiente de parcelas del IFN muestra muy buenos resultados para los ajustes GLM que incorporan el estrato como variable cualitativa (Tabla 5). Estos modelos son los que obtuvieron menor sesgo y RMSE en el proceso de validación para la mayoría de las variables y especies analizadas (Tabla 5) y fueron utilizados para obtener una cartografía continua de las variables de inventario en las 2 áreas piloto consideradas (Figura 3C).

1

Tabla 4. Modelos GLM seleccionados en la fase de ajuste para cada una de las variables y especie

| Sp | Variable | Nombre modelo | Variables independientes | Función distribución | Función vínculo | pseudo R ² | Sesgo | EAM | RMSE | sesgo relativo | eam relativo | RMSE relativo |
|-------|----------|---------------|---|----------------------|-----------------|-----------------------|--------|---------|---------|----------------|--------------|---------------|
| Silv | G | G_glm_Silv | variance, IQ, CRR, LFCC | gamma | power (0,8) | 0,58 | -0,003 | 2,369 | 3,323 | 0,0 | 6,3 | 8,8 |
| Silv | V | V_glm_silv | P50, CRR, IQ, stddev | gamma | Identity | 0,75 | -0,969 | 45,200 | 66,703 | -0,4 | 17,7 | 26,1 |
| Silv | N | N_glm_silv | P95, LFCC | gamma | Log | 0,59 | 10,074 | 265,605 | 391,678 | 1,3 | 34,1 | 50,3 |
| Silv | a_W | a_glm_silv | variance, P25, P75, LFCC | gamma | Log | 0,53 | -0,032 | 4,894 | 6,168 | -0,1 | 16,8 | 21,2 |
| Silv | b_W | b_glm_silv | IQ, P75, LFCC | gamma | power (2) | 0,32 | 0,002 | 1,243 | 1,528 | 0,1 | 33,0 | 40,6 |
| Haya | G | G_glm_haya | CRR, P75 | gamma | power (2) | 0,39 | -0,077 | 5,659 | 6,847 | -0,2 | 17,8 | 21,5 |
| Haya | V | V_glm_haya | IQ, P25, CRR | gamma | Identity | 0,64 | -1,435 | 43,216 | 53,640 | -0,7 | 19,8 | 24,5 |
| Haya | N | N_glm_haya | stddev, CV, P01, P50, CRR | gamma | Log | 0,40 | -4,461 | 274,083 | 361,115 | -0,5 | 33,4 | 44,0 |
| Haya | a_W | a_glm_haya | P75 | gamma | Inverse | 0,44 | 0,000 | 5,150 | 6,850 | 0,0 | 20,8 | 27,6 |
| Haya | b_W | b_glm_haya | CRR, LFCC | gamma | Log | 0,29 | -0,002 | 0,561 | 0,745 | -0,1 | 23,2 | 30,8 |
| Todas | G | G_glm_estrato | Estrato:CRR, LFCC, IQ | gamma | Log | 0,48 | -0,015 | 6,140 | 8,097 | 0,0 | 17,8 | 23,4 |
| Todas | V | V_glm_estrato | IQ:Estrato, P25, CRR: Estrato, LFCC | gamma | power(0.4) | 0,67 | 0,084 | 47,973 | 63,349 | 0,0 | 20,3 | 26,9 |
| Todas | N | N_glm_estrato | CV, P95:Estrato, LFCC | gamma | Log | 0,44 | 4,245 | 284,394 | 390,551 | 0,5 | 35,4 | 48,6 |
| Todas | a_W | a_glm_estrato | Estrato, IQ:Estrato, P25:Estrato, LFCC:Estrato | gamma | Log | 0,51 | -0,011 | 4,997 | 6,597 | 0,0 | 18,7 | 24,7 |
| Todas | b_W | b_glm_estrato | stddev , P25, LFCC:Estrato | gamma | Inverse | 0,42 | -0,003 | 0,928 | 1,205 | -0,1 | 30,4 | 39,5 |

2

3

Tabla 5. Resultados validación independiente de los modelos.

| Especie | Variable | Modelo | Sesgo | eam | RMSE | sesgo relativo | EAM relativo | RMSE relativo |
|-----------|----------|-------------------------|--------|--------|--------|----------------|--------------|---------------|
| Silvestre | G | G_glm_silv | -1.93 | 9.27 | 11.59 | -5.03 | 24.10 | 30.15 |
| Silvestre | G | G_glm_estrato | -0.71 | 9.23 | 11.44 | -1.85 | 24.02 | 29.76 |
| Silvestre | N | N_glm_silv | -98.00 | 215.98 | 263.18 | -16.71 | 36.83 | 44.88 |
| Silvestre | N | N_glm_estrato | -95.98 | 209.19 | 265.74 | -16.37 | 35.67 | 45.31 |
| Silvestre | V | V_glm_silv | -13.62 | 88.15 | 127.06 | -4.76 | 30.77 | 44.35 |
| Silvestre | V | V_glm_estrato | 13.08 | 73.29 | 105.57 | 4.57 | 25.58 | 36.85 |
| Haya | G | G_glm_haya | -1.39 | 4.49 | 5.18 | -4.47 | 14.44 | 16.66 |
| Haya | G | G_glm_estrato | -1.02 | 4.19 | 4.77 | -3.43 | 14.12 | 16.06 |
| Haya | N | N_glm_haya | 104.38 | 472.00 | 553.74 | 11.79 | 53.33 | 62.57 |
| Haya | N | N_glm_estrato | -16.38 | 36.09 | 42.84 | -8.74 | 19.26 | 22.87 |
| Haya | V | V_glm_haya | -18.53 | 37.80 | 44.47 | -9.00 | 18.36 | 21.60 |
| Haya | V | V_glm_estrato | 116.04 | 445.35 | 542.07 | 11.73 | 45.01 | 54.79 |
| Todas | G | G_glm_silv y G_glm_haya | -1.68 | 7.04 | 9.18 | -4.66 | 19.54 | 25.48 |
| Todas | G | G_glm_estrato | -0.85 | 6.89 | 8.98 | -2.49 | 20.04 | 26.12 |
| Todas | N | N_glm_silv y N_glm_haya | -15.90 | 64.73 | 97.75 | -6.21 | 25.26 | 38.14 |
| Todas | N | N_glm_estrato | 2.64 | 319.03 | 417.67 | 0.34 | 41.22 | 53.97 |
| Todas | V | V_glm_silv y V_glm_haya | -3.87 | 335.05 | 423.87 | -0.50 | 43.08 | 54.50 |
| Todas | V | V_glm_estrato | -0.62 | 55.99 | 82.55 | -0.26 | 23.29 | 34.34 |

4

5

6

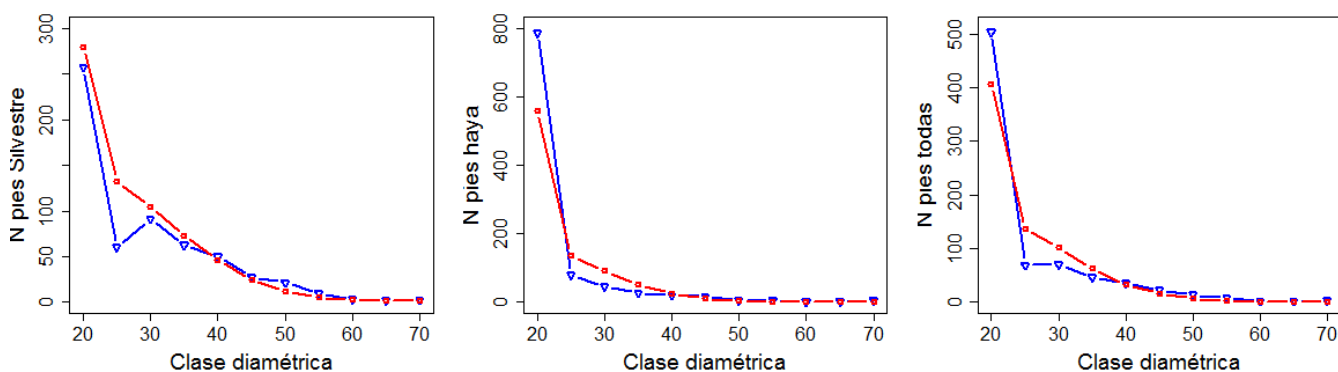
7

8

9

10

Las distribuciones diamétricas real y predicha a partir de la función Weibull en la muestra de validación poseen formas similares, describiendo ambas una estructura de masa irregular (Figura 4). La muestra de validación representa una masa irregular teniendo en cuenta que la selección de la muestra de validación se ha realizado de forma aleatoria entre todas las parcelas seleccionadas en la provincia y que por tanto posee parcelas con diferentes estados de desarrollo y estructuras de masa.



11

12

13

Figura 4. Comparación de la distribución diamétrica real (línea azul) con la distribución diamétrica predicha (línea roja) en la muestra de validación. La clase diamétrica 20 incluye todos los pies entre 7,5 cm y 22,5 cm.

14

15

16

17

El análisis del sesgo por clase diamétrica en la muestra de validación muestra valores altos para clases diamétricas de 5 cm, cercanos al 100 % en algunos casos. Se observa una reducción del sesgo relativo al agrupar y reducir el número de clases diamétricas consideradas. Considerando sólo dos clases diamétricas, pies mayores y menores de 32,5 cm, el sesgo relativo se reduce a -1% y 5%.

18 4. Discusión

19
20 El presente estudio muestra como la estratificación necesaria en un inventario LiDAR puede ser
21 generada con una resolución adecuada por medio de la clasificación de imágenes satelitales Landsat
22 combinadas con otras variables auxiliares. La precisión obtenida en la clasificación de imagen es alta,
23 indicando que la incorporación de este tipo de imágenes en inventarios forestales en grandes
24 superficies, en especial en inventarios con tecnología LiDAR, puede suponer una mejora en la
25 resolución de la cartografía de especies principales y por tanto, una mejora en las estimaciones de las
26 variables de inventario forestal.

27
28 La utilización de variables derivadas del modelo digital de elevaciones en la fase de
29 clasificación de especies principales puede mejorar los resultados de la cartografía, sobre todo
30 cuando se trabaja en ámbitos territoriales amplios (provincia o región). Varios autores han detectado
31 la importancia de la elevación en la clasificación de coberturas (Dewan and Yamaguchi 2009;
32 Rodríguez-Galiano et al. 2012). Por un lado, la elevación (y todas sus posibles derivadas) son
33 variables que obtenemos de forma sencilla con precisión y elevada resolución a partir de la
34 información LiDAR y además, por otro lado, son variables que juegan un papel determinante en
35 muchos procesos ecológicos y de ecología del paisaje siendo un buen estimador de otras variables
36 ambientales como la temperatura y la precipitación. Por este motivo es de esperar que la elevación y
37 otras las variables derivadas del MDT ganen importancia en la clasificación cuando se trabaje a
38 escala regional y aumente la variabilidad ecológica del área de estudio.

39
40 Los RMSE de los modelos GLM ajustados para el área basimétrica (entre 3,3 m²/ha para
41 pinares puros y 8,1 m²/ha para masas mixtas) se encuentran en el rango de los obtenidos por otros
42 autores. González-Ferreiro et al (2012) y Stephens et al. (2007) obtuvieron RMSE en torno a 8 m²/ha
43 para *Pinus Radiata*. Otros autores como González-Olabarria et al (2012) en masas de *Pinus nigra* y
44 *Pinus Pinaster* en el Ibérico central obtuvieron un RMSE de 6,89 m²/ha, o Holmgren (2004) en masas
45 boreales consiguió un RMSE de 2,7 m²/ha.

46
47 Analizando los RMSE obtenidos para la variable volumen (RMSE comprendidos entre 53,64 en
48 haya y 66,7 m³/ha en pinar) observamos valores similares a los encontrados en otros trabajos.
49 Hollaus et al (2007) trabajando en los Alpes austriacos obtuvieron un RMSE de 96,8 m³/ha,
50 Gonzalez-Ferreiro et al (2012) obtuvieron valores entre 76,9 y 94,1 m³/ha para *Pinus Radiata* en
51 Galicia, Næsset (2004) generó modelos para esta variable con RMSE comprendidos entre 32,9 y 67,8
52 m³/ha en masas boreales, o González-Olabarria et al (2012) que para masas de pinar en España
53 obtuvieron un RMSE de 52,46 m³/ha.

54
55 El cuanto al número de pies, los RMSE de los modelos GLM ajustados para esta variables se
56 encuentran entre 361,1 pies/ha y 391,68 pies/ha (entre 44% y 50,3 % en valores relativos). Estos
57 resultados son sensiblemente peores a los obtenidos en otros trabajos a nivel de monte (1000 ha
58 aprox) por González-Olabarria et al (2012) en el Ibérico central (Soria) con un RMSE de 206,98
59 pies/ha (29.4 %) o los obtenidos por Næsset (2002) en el sureste de Noruega con precisiones
60 relativas en esta variable entre 16,9–22,2%.

61
62 Las precisiones obtenidas en el área basimétrica y el volumen se encuentran dentro de los
63 rangos de otros trabajos, incluso considerando aquellos estudios cuyo ámbito geográfico mucho más
64 reducido, mil o pocos miles de ha. Estos resultados están en la línea de los obtenidos por Nord-Larsen
65 & Schumacher (2012) para variables similares empleando un vuelo LiDAR nacional y parcelas de
66 radio variable del IFN Danés. En cambio, las precisiones obtenidas en el número de pies son
67 sensiblemente inferiores en este trabajo a las alcanzadas en estudios similares. Esto puede ser
68 debido principalmente al efecto del radio variable de las parcelas del IFN, que afecta de forma más
69 acusada en la estimación de N que de G y V, errores altos en la estimación del número de pies de las

70 clases diamétricas más bajas pueden suponer errores bajos en la estimación de volumen y área
71 basimétrica.

72
73 En cuanto a la estimación de las distribuciones diamétricas, la validación independiente muestra
74 errores altos para clases diamétricas de 5 cm, los resultados mejoran sensiblemente cuando
75 trabajamos con agrupaciones de diámetros más amplias. La estimación de distribuciones de
76 diámetros a partir de información LiDAR posee desviaciones que se deben no sólo a los errores en la
77 estimación de los parámetros a y b de Weibull sino también a los errores en la estimación del número
78 de pies, variable cuya bondad de ajuste suele ser más baja cuando se estima con datos LiDAR.

79 80 **5. Conclusiones**

81
82 Los errores obtenidos en la fase de ajuste y de validación de los modelos de área basimétrica y
83 volumen son similares a los registrados por otros autores (usando un vuelo específico y parcelas de
84 campo específicas). Se observan errores mayores en la variable número de pies (y por tanto también
85 en su distribución por clases diamétricas) que los encontrados en la literatura, que pueden ser
86 explicados por la influencia de la metodología de parcelas de radio variable en esta variable.

87
88 La incorporación de información satelital de media resolución e información LiDAR proveniente
89 de vuelos nacionales en el IFN u otros inventarios en grandes superficies puede contribuir a mejorar
90 la resolución y a describir la variabilidad espacial de las principales variables de masa forestal.

91
92 Estos resultados refuerzan el interés del IFN en incorporar la tecnología LiDAR. Está previsto
93 hacer una prueba piloto en Extremadura y Canarias (con intención de continuar en el resto de
94 provincias y CCAA) combinando información de los estratos del MFE25, valores dasométricos de las
95 parcelas de campo del IFN4 e información procesada del último vuelo LiDAR disponible para expandir
96 de manera continua los valores dasométricos en los estratos del MFE utilizando una metodología
97 similar a la anteriormente expuesta. Para facilitar esta tarea se han incorporado equipos GPS con
98 precisión submétrica, de cara al levantamiento de los centros de parcelas del IFN a partir de
99 Extremadura. Por otro lado, está previsto realizar un nuevo inventario forestal en las masas
100 productivas del norte de España, en el que la información de alturas y fracción de cubierta
101 proporcionada por el último vuelo LiDAR disponible será de gran ayuda a la hora de confeccionar los
102 estratos del MFE.

103 104 **6. Agradecimientos**

105
106 Desde el equipo redactor queremos dar las gracias a todo el equipo que contribuye a que el
107 Mapa Forestal y el Inventario Nacional Forestal por su colaboración y la confianza depositada en
108 nosotros. Y a Juan Carlos Ojeda del IGN por su apoyo en todo lo relacionado con el PNOA.

109 110 **7. Bibliografía**

- 111
112 BREIMAN, L. (2001): Random Forests. Machine Learning 45 (1), pp 5–32
- 113
114 CHANDER G., MARKHAM, B.L., HELDER, D.L.. 2009. Summary of current radiometric
115 calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sensing of
116 Environment 113 (2009) 893–903
- 117
118 DEWAN M A., YAMAGUCHI Y., 2009. Using remote sensing and GIS to detect and monitor land
119 use and land cover change in Dhaka Metropolitan of Bangladesh during 1960-2005.
120 Environmental Monitoring and Assessment. 150, 237-249.

121

- 122 GONZÁLEZ-FERREIRO E., DIÉGUEZ-ARANDA U., MIRANDA D. 2012. Estimation of stand
123 variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry*,
124 Vol. 85, No. 2, 2012
125
- 126 GONZÁLEZ-OLABARRIA J. R., RODRÍGUEZ F., FERNÁNDEZ-LANDA A., MOLA-YUDEGO B. 2012.
127 Mapping fire risk in the Model Forest of Urbión (Spain) based on airborne LiDAR
128 measurements. *Forest Ecology and Management* 282 (2012) 149–156
129
- 130 HOLLAUS, M., WAGNER, W., MAIER, B. AND SCHADAUER, K. 2007. Airborne laser scanning of
131 forest stem volume in a mountainous environment. *Sensors*. 7, 1559–1577.
132
- 133 HOLMGREN J. 2004. Prediction of tree height, basal area and stem volume in forest stands
134 using airborne laser scanning. *Scandinavian Journal of Forest Research*. Volume 19, Issue 6,
135 pages 543-553
136
- 137 KNUDSON C. 2015. glmm: Generalized Linear Mixed Models via Monte Carlo Likelihood
138 Approximation. R package version 1.0.3. <http://CRAN.R-project.org/package=glmm>
139
- 140 MCGAUGHEY, R. J., CARSON, W.W., 2003. Fusing LIDAR data, photographs, and other data
141 using 2D and 3D visualization techniques. In: *Proceedings of Terrain Data: Applications and*
142 *Visualization–Making the Connection*, October 28–30, 2003. pp. 16–24.
143
- 144 MARSCHNER I. 2014. GLM2: Fitting Generalized Linear Models. R package version
145 1.1.2. <http://CRAN.R-project.org/package=glm2>
146
- 147 NÆSSET E. 2002. Predicting forest stand characteristics with airborne scanning laser using a
148 practical two-stage procedure and field data. *Remote Sensing of Environment* 80 88– 99
149
- 150 NÆSSET, E. 2004 Practical large-scale forest stand inventory using a small-footprint airborne
151 scanning laser. *Scand. J. Forest. Res.* 19, 164–179.
152
- 153 NORD-LARSEN T, SCHUMACHER J. 2012. Estimation of forest resources from a country wide
154 laser scanning survey and national forest inventory data. *Remote Sensing of Environment*.
155 119:148-157
156
- 157 PARKER AND RUSS. 2004. The canopy and stand development assessing forest canopy
158 structure and complexity with near-surface altimetry. *Forest Ecology and Management* 189,
159 307-315
160
- 161 RODRÍGUEZ-GALIANO VF, GHIMIRE B, ROGAN J, CHICA-OLMO M, RIGOL-SÁNCHEZ JP. 2012.
162 An assessment of the effectiveness of a random forest classifier for land-cover classification.
163 *ISPRS Journal of Photogrammetry and Remote Sensing* 67: 93–104.
164
- 165 VENABLES W. N. Y RIPLEY B. D. (2002) *Modern Applied Statistics with S*. Fourth Edition.
166 Springer, New York. ISBN 0-387-95457-0
167
- 168 WASKE B, BRAUN M. 2009. Classifier ensembles for land cover mapping using multitemporal
169 SAR imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 64 (5):450–457.