



Estimación de la biomasa potencial
disponible de *Pinus radiata* para
usos energéticos en Bizkaia a partir
de datos LiDAR aerotransportados.
Propuesta metodológica

TESIS DOCTORAL
2017

Autor: José Ignacio Moro Deordal

Director: Dr. Eugenio Ruiz Urrestarazu

DEPARTAMENTO DE GEOGRAFÍA, PREHISTORIA Y ARQUEOLOGÍA

A Nerea y a mis hijas Leire y Maia. Sin
vuestra ayuda y comprensión no hubiera
podido terminar este reto.

Contenido

INTRODUCCIÓN, HIPÓTESIS, OBJETIVOS Y PLANTEAMIENTO METODOLÓGICO	1
1. Introducción	2
1.1. La valorización energética de la biomasa	6
1.2. La biomasa forestal para uso energético en la Comunidad Autónoma de Euskadi y Bizkaia.....	14
2. Hipótesis y objetivos	22
3. Planteamiento metodológico	24
3.1. Estimación de la Biomasa Forestal Potencial Total	26
3.2. Estimación de la Biomasa Forestal Potencial Accesible.....	37
3.3. Estimación de la Biomasa Forestal Potencial Disponible.....	42
ANTECEDENTES Y ESTADO DE LA CUESTIÓN	43
4. La estimación de la biomasa en el contexto de la teledetección	44
4.1. Fundamentos de la teledetección	44
4.2. La teledetección laser	47
4.3. El LiDAR Aerotransportado	49
4.3.1. Características técnicas.....	51
4.3.2. Formato de los datos	58
4.3.3. Productos derivados de datos LiDAR brutos	61
4.3.4. Formas de onda completa y sistemas de retorno discreto.....	67
4.3.5. Patrones de distribución de la huella LiDAR.....	70
4.3.6. Adquisición de datos LiDAR.....	72
4.4. El LiDAR en los estudios forestales	78
4.4.1. LiDAR Aerotransportado y estimación de parámetros forestales a escalas regional y local	82
4.4.2. Enfoques basados en área.....	86
4.4.3. Enfoques de análisis a nivel de árbol individual	91
4.5. El LiDAR en el estudio de la biomasa forestal	96
4.5.1. Estimación de la biomasa. Definiciones y Términos	96
4.5.2. Modelos alométricos de biomasa.....	100
4.5.3. Métodos de relación de componentes y factores de expansión de la biomasa.....	103

4.5.4.	LiDAR satelitales y estimaciones de la biomasa a escala global	105
4.5.5.	Métricas y modelos de predicción de biomasa forestal derivadas de datos LiDAR.....	108
EL ÁREA DE ESTUDIO		123
5.	El área de estudio: delimitación y presentación general	124
5.1.	El pino radiata en la CAE y en Bizkaia	130
6.	OBTENCIÓN DE LAS VARIABLES DE LOS MODELOS DE ESTIMACIÓN	136
6.1.	Tratamiento de los datos de campo de la biomasa forestal...	136
6.1.1.	El Inventario Nacional Forestal.....	137
6.1.2.	El cuarto inventario forestal nacional en el País Vasco	143
6.1.3.	Los datos de campo del IFN4 en Bizkaia	150
6.1.4.	Cálculo de la biomasa de las parcelas del IFN4	156
6.1.5.	Comprobación del posicionamiento de las parcelas de campo	162
6.2.	Tratamiento de los datos LiDAR	166
6.2.1.	Obtención y procesamiento de las métricas	169
7.	METODOLOGÍA ESTADÍSTICA	180
7.1.	Análisis de correlaciones	180
7.2.	Modelos de regresión lineal	181
7.2.1.	Valoración de la capacidad explicativa y de la exactitud de la predicción.....	185
7.2.2.	Supuestos del análisis de regresión lineal	189
7.2.3.	Métodos de selección de variables	202
7.3.	Modelos de regresión no lineales	216
7.4.	Métodos de validación de los modelos	218
8.	ANÁLISIS DE LOS RESULTADOS DE ESTIMACIÓN.....	223
8.1.	Análisis de la normalidad	225
8.2.	Selección de variables	232
8.2.1.	Modelos con métodos <i>Stepwise</i>.....	235
8.2.2.	Modelos con Métodos de Mejores Subconjuntos	239
8.3.	Selección final de modelos	258
8.4.	Tratamiento de los valores anómalos	261

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

8.4.1.	Estimación de modelos sin parcelas anómalas	272
8.5.	Cartografiado de las estimaciones y del grado de incertidumbre	275
8.5.1.	Incertidumbre de las predicciones	276
8.6.	Discusión de los resultados de los modelos de estimación...	280
CAPÍTULO X		285
CARTOGRAFÍA E INVENTARIO		285
9.	CARTOGRAFIA E INVENTARIO DE LA BIOMASA	286
9.1.	Cartografía e inventario a nivel provincial de la Biomasa Forestal Potencial Total y el CO ² almacenado.	286
9.2.	Estimación de la Biomasa Forestal Potencial Accesible.....	289
9.2.1.	Limitantes y condicionantes relacionados con la pendiente:	289
9.2.2.	Limitantes y condicionantes relacionadas con la distancia a desembosque:.....	291
9.2.3.	Limitantes y condicionantes jurídico-administrativos:..	295
9.2.4.	Coeficiente de recogida.	304
9.3.	Biomasa Forestal Accesible. Cartografía e inventario a escala provincial, comarcal y municipal.....	304
10.	CONCLUSIONES	310
Referencias		314

ÍNDICE DE FIGURAS

Procesos de valorización energética de la biomasa	7
Esquema de los posibles sistemas logísticos en últimas claras y cortas finales en terreno en pendiente. (Cabrera et al., 2011)	41
Sistema LiDAR aerotransportado (Matti Maltamo et al., 2014) ..	51
Retornos en vegetación de los pulsos láser (Matti Maltamo et al., 2014).....	53
Forma de onda de respuesta LiDAR (Matti Maltamo et al., 2014)	54
Huella LiDAR.....	55
Modelo Digital del Terreno (Matti Maltamo et al., 2014)	65
Modelo Digital de Superficies (Matti Maltamo et al., 2014) ...	67
Forma de onda completa.....	68
Respuesta de suelo y vegetación sin sustrato arbustivo	69
Huella de LiDAR satelital (Matti Maltamo et al., 2014)	71
Líneas de vuelo LiDAR aerotrasnportado (Matti Maltamo et al., 2014)	75
Posibles escenarios de interceptación de la nube de puntos LiDAR del dosel forestal. Fuente: (Suárez et al., 2005)	84
Distribución y estructura de la vegetación mostrada por clasificación de una nube de puntos LiDAR aerotransportado.	85
Componentes de biomasa de árbol (Matti Maltamo, Erik Næsset y Jari Vauhkonen, 2014)	97
Resumen de los pasos de la metodología de masa/área (Matti Maltamo et al., 2014)	113
Evolución de la superficie forestal en Bizkaia Fuente: Inventarios Forestales	131
Distribución del <i>Pinus radiata</i> en formaciones de Latizal 7 Fustal en Bizkaia (Elaboración propia a partir del IFN4) .	134
Estadillo de levantamiento de parcelas del IFN4.....	142
Estadillo de levantamiento de parcelas del IFN4.....	143
Parcelas con <i>P. radiata</i> del IFN4	153
Parcelas incluidas y eliminadas	154
Cantidad de biomasa estimada por parcela a partir de las distintas ecuaciones alométricas ((Canga et al., 2013; Montero González et al., 2006; Pérez-Cruzado, 2008)	161
"Parcelas virtuales en las 8 direcciones".....	163
Diferencia entre el valor medio del percentil 95 de altura de la parcela original y el de cada una de las parcelas desplazadas	165

Hojas LiDAR cubriendo la CAE, Bizkaia, la superficie ocupada por el P. radiata en Latizal/Fustal y parcelas de campo seleccionadas: (elaboración propia)	167
Densidad de puntos LiDAR: (elaboración propia)	168
Métricas LiDAR explicadas en representaciones 3D (McGaughey, 2016)	172
Esquema de cálculo del porcentaje de cobertura con una altura de corte (McGaughey, 2016)	174
Métricas LiDAR de densidad explicadas en representaciones 3D	177
Métricas LiDAR de proporción de retornos por estrato	178
Modelo MARS (Fuente: http://media.salford-systems.com/pdf/spm7/IntroMARS.pdf)	211
Validación Cruzada (Método de retención) By Joan.domenech91 (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons	219
Validación cruzada de K iteraciones con K=4 By Joan.domenech91 (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons	220
Validación cruzada dejando uno fuera (LOO) By Joan.domenech91 (Own work) [CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons	221
Coeficiente de correlación de Pearsons entre variables	228
Importancia de los predictores para Montero (Método Stepwise)	236
Importancia de los predictores para Canga (Método Stepwise) .	237
Importancia de los predictores para Pérez-Cruzado (Método Stepwise)	238
Importancia de los predictores para Montero (Mejor subconjunto R^2)	241
Importancia de los predictores para Canga (Mejor subconjunto R^2)	242
Importancia de los predictores para Pérez-Cruzado (Mejor subconjunto R^2)	242
Estadísticos de los coeficientes para Montero (Mejor subconjunto Prevención del sobreajuste)	244
Importancia de los predictores para Canga (Mejor subconjunto Prevención del sobreajuste)	244
Importancia de los predictores para Pérez-Cruzado (Mejor subconjunto Prevención del sobreajuste)	245
Selección de variables para Montero (Métodos Ridge y LASSO) .	248
Selección de variables para Canga (Métodos Ridge y LASSO) ...	249

Selección de variables para Pérez-Cruzado (Métodos Ridge y LASSO)	250
Graficas de normalidad, homocedasticidad y linealidad de los residuos (MARS)	260
Graficas de normalidad, homocedasticidad y linealidad de los residuos (Regresión Lineal)	261
Gráficas de valores anómalos	266
Ejemplo de parcelas con datos anómalos	272
Graficas de normalidad, homocedasticidad y linealidad de los residuos (Regresión Lineal)	274
Graficas de normalidad, homocedasticidad y linealidad de los residuos (MARS)	274
Gráfico de la incertidumbre de las predicciones	279
Mapa de productividad forestal (Palomares y Serrano, 2000) Elaboración propia	290
Aminoramiento de la Biomasa Forestal Potencial Accesible. Factor distancia	294
Limitantes y condicionantes de la Reserva de la Biosfera de Urdaibai (Elaboración propia)	299
Limitantes y condicionantes de los Parques Naturales de Urkiola y Gorbeia (Elaboración propia)	300
Mapa de limitantes y condicionantes de las Zonas de Especial Conservación de Ordunte y Armañón (Elaboración propia) ...	302
Mapa de limitantes y condicionantes del Plan Territorial Sectorial Agroforestal (Elaboración propia)	303

ÍNDICE DE TABLAS

Tabla 1 Tasas de productividad forestal de la CAE.....	15
Tabla 2 Costes de producción de la biomasa	17
Tabla 3 Valor de los parámetros de la función (a y b), del coeficiente de determinación ajustado (R2 adj) y el error estándar de la estimación (SEE) para cada especie y fracción de biomasa.....	34
Tabla 4 Valores modulares de las distintas fracciones de biomasa (kg de materia seca) en España para <i>P. radiata</i> ..	35
Tabla 5 Valores de la clasificación de puntos LiDAR.....	60
Tabla 6 Componentes de biomasa del árbol	97
Tabla 7 Variación de superficies y densidades de <i>P. radiata</i>	133
Tabla 8 Estratos forestales del IFN4	147
Tabla 9 Factores de expansión según Dn.	155
Tabla 10 Ecuaciones alométricas de <i>P. radiata</i> de Pérez-Cruzado	156
Tabla 11 Ecuaciones alométricas de <i>P. radiata</i> de Montero Gonzalez	157
Tabla 12 Ejemplo de cálculo de la biomasa en una parcela del IFN4	158
Tabla 13 Biomasa de las parcelas seleccionadas para el estudio	159
Tabla 14 Pruebas de normalidad de la variable Percentil 95 de altura en las parcelas desplazadas.	164
Tabla 15 Parámetros de configuración del sensor LiDAR.....	166
Tabla 16 Métricas LiDAR relacionadas con la altura y la estructura vertical del dosel extraídas a partir de los datos LiDAR con el programa FUSION	171
Tabla 17 Métricas LiDAR relacionadas con la densidad de cubierta del dosel extraídas a partir de los datos LiDAR con el programa FUSION.....	173
Tabla 18 Transformaciones Box-Cox	191
Tabla 19 Modelos de regresión curvilínea	217
Tabla 20 Codificación de la variables del estudio.....	224
Tabla 21 Pruebas de normalidad de las variables. (En rojo variables que no han superado al menos un test de normalidad)	225
Tabla 22 Variables descartadas por falta de significación estadística en el C. de Correlación de Pearsons.....	228
Tabla 23 Grado de ajuste y error de los modelos de regresión con una sola variable.....	229

Tabla 24 Grado de ajuste y error de los modelos de regresión con una sola variable. Variables log-transformadas.....	231
Tabla 25 Método Stepwise.....	236
Tabla 26 Estadísticos de los coeficientes para Montero (Stepwise).....	236
Tabla 27 Estadísticos de los coeficientes para Canga (Stepwise).....	237
Tabla 28 Estadísticos de los coeficientes para Pérez-Cruzado (Stepwise).....	238
Tabla 29 Método mejores subconjuntos. Criterio R^2	241
Tabla 30 Estadísticos de los coeficientes para Montero (Mejor subconjunto R2).....	241
Tabla 31 Estadísticos de los coeficientes para Canga (Mejor subconjunto R2).....	242
Tabla 32 Estadísticos de los coeficientes para Pérez-Cruzado (Mejor subconjunto R2).....	243
Tabla 33 Método mejores subconjuntos. Criterio de prevención del sobreajuste.....	244
Tabla 34 Importancia de los predictores para Montero (Mejor subconjunto Prevención del sobreajuste).....	244
Tabla 35 Estadísticos de los coeficientes para Canga (Mejor subconjunto Prevención del sobreajuste.....	245
Tabla 36 Estadísticos de los coeficientes para Pérez-Cruzado (Mejor subconjunto Prevención del sobreajuste.....	245
Tabla 37 Resultados de ajuste de los métodos Rigde y Lasso	247
Tabla 38 ajuste y error de los modelos MARS.....	252
Tabla 39 Variables y Funciones de Base de los Modelos MARS	253
Tabla 40 Coeficientes y VIF de las funciones de base de los modelos MARS.....	254
Tabla 41 Resumen del rendimiento de los modelos.....	256
Tabla 42 Rendimiento de los modelos seleccionados.....	258
Tabla 43 Estadísticos de los coeficientes y colinealidad de los modelos seleccionados.....	258
Tabla 44 Test de validación de los supuestos estructurales de los modelos seleccionados.....	260
Tabla 45 Parcelas con valores atípicos.....	263
Tabla 47 Test de validación de los supuestos estructurales de los modelos seleccionados sin valores anómalos.....	273
Tabla 48 Coeficientes y VIF de las funciones de base de los modelos seleccionados sin anómalos.....	273
Tabla 49 Ecuaciones finales.....	273

Tabla 50 Zonas de productividad de acuerdo con los valores alcanzados por el índice de Paterson (Palomares y Serrano, 2000)	290
Tabla 51 Valores de minoración de la biomasa en función de la distancia al desembosque	292
Tabla 52 Condicionantes y limitantes en la Reserva de la Biosfera de Urdaibai	297
Tabla 53 Condicionantes y limitantes en los Parques Naturales de Urkiola y Gorbeia	300
Tabla 54 Condicionantes y limitantes en las Zonas de Especial Conservación de Armañón y Ordunte	301
Tabla 55 Inventario de la Biomasa Forestal Potencial Disponible a escala provincial, comarcal y municipal	307

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

CAPITULO 1
INTRODUCCIÓN, HIPÓTESIS, OBJETIVOS Y
PLANTEAMIENTO METODOLÓGICO

1. Introducción

La asunción por parte de las sociedades occidentales de la importancia que el abastecimiento energético suponía para la reproducción de sus modos de vida impactó con fuerza en los años primeros 70 del pasado siglo al hilo de la denominada crisis del petróleo. La evolución de los precios del petróleo y la distribución geográfica de las reservas de energía han condicionado posteriormente, durante más de tres décadas, las opciones energéticas de los países desarrollados. En un primer momento, el aumento generalizado de los precios del petróleo llevó a la búsqueda de energías alternativas, entre las cuales se encontraban las renovables: solar, eólica, biomasa, hidráulica... (Domínguez Bravo, 2002). Sin embargo, el incumplimiento de las expectativas de crecimiento de los precios del petróleo y los elevados precios de las energías renovables a corto plazo (inversión en I+D+I) propició un descenso en el interés del impulso de éstas durante las décadas de los 80 y 90 (Esteban Pascual, Luis S, 2004).

A finales de los noventa, las políticas energéticas, tanto europeas como estatales, comienzan a derivar hacia cuatro líneas estratégicas, coincidiendo, por un lado, con la ratificación de los procesos de Globalización de los Mercados (p.e. Creación de la Organización Mundial del Comercio, OMC en 1995) y el resurgimiento del neo-liberalismo político e ideológico; y por otro, con la aceptación internacional de la necesidad de reducción las

emisiones de gases invernadero protocolizada en el compromiso de Kioto de 1997: a) la liberalización de los mercados, b) la garantía del suministro, c) el desarrollo de las infraestructuras de interconexión y d) la reducción de emisiones contaminantes.

Respecto a las energías renovables (EERR) ya en el Libro Blanco de la Comisión Europea sobre la energía renovable se planteaba la necesidad de fomentar activamente las energías renovables en Europa, justificando su importancia, en buena medida, como colaboradoras necesarias para la consecución de esas líneas estratégicas, además de poner en valor sus potencialidades como fuerzas motrices de otras de carácter socio-económico y territorial: a) fomentar el desarrollo regional y la creación de empleo; b) promover el desarrollo industrial de las tecnologías de Energías Renovables y posicionamiento estratégico a nivel mundial (exportación).

En el caso del España, la convergencia con estos objetivos ha estado condicionada por cuatro características propias del sector energético, que, en gran medida, podrían calificarse como debilidades:

a) Consumo energético por unidad de producto interior bruto mayor que la media europea (UE-15 y UE27).

b) Elevada dependencia energética, propiciada fundamentalmente por la escasa presencia de yacimientos de energía primaria fósil, agravadas por una escasa interconexión física con el resto de Europa; y un excesivo protagonismo de la logística

naval y la consiguiente vulnerabilidad frente a perturbaciones de las vías de comunicación marítimas.

c) Infracapacidad e inadecuada estructura del refino de petróleo -importamos gasoil y exportamos gasolina (14.000 y 1.800 kt respectivamente de media entre 2004 y 2008, (EUROPEAN COMMISSION'S DIRECTORATE-GENERAL FOR ENERGY, 2011) -, así como una insuficiente capacidad de almacenamiento de gas y de plantas de generación eléctrica reversible.

d) Emisiones de gases de efecto invernadero por encima de la media de los países de la EU15 firmantes del protocolo de Kioto, explicadas fundamentalmente por el crecimiento de la generación eléctrica, del transporte y del tratamiento y eliminación de residuos (un 2% anual desde 1990 a 2008). (MMARM, 2011, IDAE, 2010).

Las estrategias implementadas desde las distintas administraciones para hacer frente a estos retos pueden resumirse en tres vías de actuación:

a) Liberalización y transparencia de los mercados de la energía: básicamente mediante medidas de carácter normativo - no exentas de vaivenes difícilmente explicables - que han originado, desde su puesta en marcha y a lo largo de sus respectivos recorridos, intensos debates, principalmente en torno al sector eléctrico.

b) Modernización de la red de transporte y distribución (p.e. un incremento del 2% anual 2000-2010 del nº de km de la red de 400 kV) (REE, 2010), acompañado por mejoras en las interconexiones internacionales (especialmente con Francia y Portugal), así como la promoción del ahorro y la eficiencia energéticos.

c) El desarrollo de las energías renovables (apoyadas fundamentalmente en mecanismos regulatorios feed-in tariff (BOE, 2007)). Estímulos cuyo objeto era desarrollar una planificación energética que, al igual que ha sucedido con los intentos de regularización de los mercados, han dejado entrever, a menudo, improvisación y contradicciones ante coyunturas económicas diferentes: Planes de Energías Renovables (IDAE, 1999; IDAE, 2005; IDAE, 2010); en el caso de la Comunidad Autónoma de Euskadi (CAE) de las denominadas Estrategias Energéticas de Euskadi (3E) (EVE, 2006) y 3E 2020.

En este sentido, conviene destacar que, a pesar de que el territorio español presentaba factores endógenos, que podría calificarse de fortalezas, con respecto a la mayoría de los países de la UE, como son unas condiciones geográficas muy favorables para el desarrollo de este tipo de energías y una coyuntura de fuerte crecimiento de la demanda energética, tan sólo la eólica, con un 20% de la potencia instalada y una cobertura de la demanda de entorno al 14% parece haber cumplido con las expectativas de

crecimiento que auguraba a las EERR la planificación energética.

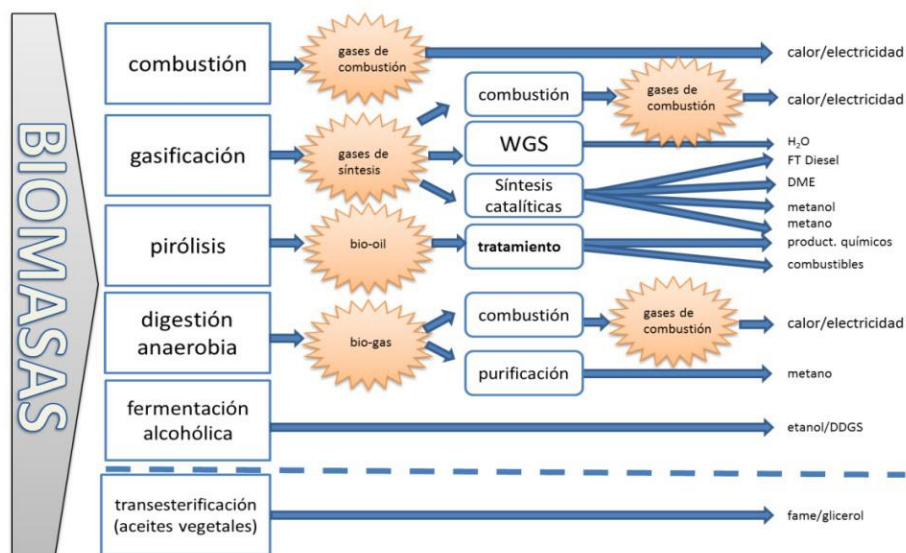
En el caso de la CAE, la generación de energía de origen renovable ha tenido como protagonista casi absoluta (al margen de la hidroeléctrica, que, a pesar de provenir de una fuente renovable, su origen y evolución es muy anterior a este proceso) a la eólica (con una media del 30% en generación, y en los últimos años con la fuerte irrupción de la termoeléctrica de origen renovable con un porcentaje del mix renovable del 30% desde 2007. (Fuente: Ente Vasco de la Energía).

Como se señalaba más arriba las normas que más impacto han tenido en la implantación y desarrollo de las EER en su conjunto, tanto en España como en la CAE, son, en un primer momento el Real Decreto 661/2007, de 25 de mayo, por el que se regula la actividad de producción de energía eléctrica en régimen especial y el RD 1/2012 de 27 de enero, que anulaba el anterior en materia de incentivos económicos para la producción de energía eléctrica a partir de energías renovables.

1.1. La valorización energética de la biomasa

El termino biomasa hace referencia en este contexto a toda materia orgánica originada en un proceso biológico, espontáneo o provocado, utilizable como fuente de energía. Desde el punto de vista energético, agrupa al conjunto de sustancias orgánicas de origen vegetal, animal o derivadas de cualquier proceso de transformación natural o artificial de las mismas, que pueden producir energía al

ser sometidas a determinados procesos físicos, químicos o biológicos. En esta definición no se incluye a los combustibles fósiles y sus derivados debido a que, aunque también tuvieron un origen biológico hace millones de años, no pueden considerarse como renovables (Fernández, 2003; Jarabo, Fernández, Trujillo, Elortegui y Pérez, 1999).



Fuente: Gloria Etxebarria (EVE). Seminario técnico «Situación actual de la biomasa y tendencias de futuro: recursos tecnológicos, mercados y políticas». Donostia, 22-23 de 2010. Elaboración propia

Figura 1 Procesos de valorización energética de la biomasa

Una diferencia particular de la biomasa sólida respecto a otras fuentes de energía renovable es su potencial de almacenamiento, lo que la convierte en un stock energético que se renueva con la suficiente rapidez para el uso humano (Jarabo et al., 1999; Puig Sales, D.L. 1985). Sin embargo, para que la biomasa sea considerada como un recurso renovable, su explotación debe ser más acompañada que la velocidad de regeneración (Domínguez Bravo, 2002). Otras

características significativas de la biomasa sólida como recurso energético son su diversidad y su escasa concentración espacial, lo que implica un gran número de procesos: cultivo (en el caso de cultivos energéticos), extracción, logísticos y de transformación para su aprovechamiento final, entre otros. (Richardson, Björheden, Hakkila, Lowe y Smith, 2006).

En cuanto a la valorización energética de la biomasa, ésta se produce básicamente a través de tres grandes tipos de procesos: Procesos termoquímicos (Combustión, Gasificación, Pirolisis); Procesos bioquímicos (Digestión anaerobia, Fermentación alcohólica); Procesos químicos (Transesterificación).

Estos procesos implican diversas y muy diferentes combinaciones de materias primas y procesos de conversión, que dan lugar a otras tantas aplicaciones energéticas, biocombustibles y subproductos químicos (Figura X).

En el caso de la biomasa forestal su utilización se restringe básicamente a procesos de combustión, gasificación y pirolisis que dan lugar a aplicaciones eléctricas, térmicas y de cogeneración (producción conjunta de energía eléctrica y térmica) (IDAE, 2005) (EVE, 2006).

La biomasa forestal para fines energéticos se obtiene básicamente a partir de tratamientos sobre el vuelo arbóreo de las masas forestales existentes. No debemos olvidar que la valorización energética de estas masas forestales entra en competencia directa con otros usos potenciales,

especialmente el maderero. Los tratamientos pueden clasificarse como:

- Tratamientos que no suponen la corta de árboles, como es el caso de podas, trasmoches, olivaciones, etc.
- Tratamientos que suponen corta de parte de los pies de las masas, cortas intermedias o de mejora de las masas, clareos (en los primeros estadios de la masa), claras, resalveos, etc.
- Cortas finales de las masas forestales al final del turno de aprovechamiento.

Tanto las podas, olivaciones, etc. como los clareos no tienen un fin comercial, sino que son operaciones destinadas a la formación y mejora de las masas de modo que permiten un mayor y mejor desarrollo de la masa en pie y se fomenta la autoprotección de las mismas frente a riesgo de enfermedades, plagas e incendios, por tanto, no compite con otros usos como el maderero y por ello se puede considerar como biomasa de restos.

En los demás tratamientos habría que diferenciar por un lado los riberones (punta de corta de copas) y ramas y ramillas, de los fustes propiamente dichos con una vocación tradicionalmente maderera. Mientras que los primeros se pueden considerar que tienen una orientación claramente energética, en el caso de los fustes dependerá del diámetro de los árboles; en el caso de cortas finales y 2ª claras el fuste, dada su dimensión, tendrá una vocación de uso

claramente maderero, sin embargo, puede ocurrir que en ciertas especies las dimensiones que alcanza el fuste en el momento de su 1ª clara pueda entrar en competencia de uso maderero versus energético. Así pues, y poniendo una limitación al diámetro de los pies para distinguir un uso final u otro se podrían considerar los siguientes tipos de biomasa aprovechable:

- Biomasa de restos con punta delgada de 7-7,5 cm, en la cual se incluyen riberones, ramas y ramillas procedentes de los distintos tratamientos y árboles completos con diámetro inferior a 7,5 cm (clareos).
- Biomasa de restos competitiva, compuesta por riberones, ramas y ramillas procedentes de los distintos tratamientos y árboles completos con diámetro inferior a 20 cm (clareos y algunas claras).
- Biomasa de árbol completo, compuesta por árboles enteros procedentes de masas en pie con o sin aprovechamiento actual.

Las tecnologías de explotación energética de la biomasa requieren además de una serie pre-tratamientos, necesarios para su utilización: reducción granulométrica (astillado, triturado, molienda...), de reducción de la humedad (secado natural o forzado) y densificación (pelletizado o briquetado), etc...

Las aplicaciones térmicas han supuesto tradicionalmente el sistema más extendido para aprovechar energéticamente los residuos forestales (utilización de leñas). En la

actualidad se han desarrollado gran número de mejoras en el desarrollo de sistemas de combustión que permiten su utilización tanto en aplicaciones en el sector doméstico (de carácter individual, como colectivo y centralizado - "District heating" -) como en el sector industrial (aplicaciones en hornos, en secaderos y en calderas industriales (Lomas, Urbano, Merino y Camarero, 2001; Spliethoff, 2010)).

La producción de electricidad mediante el uso de biomasa forestal y de biomasa en general se ha desarrollado durante los últimos años, siendo muy notables los esfuerzos en innovación tecnológica llevados a cabo. La biomasa forestal suple la materia prima de dos tipos de plantas de producción de electricidad: centrales de biomasa y centrales de co-combustión (quema de los residuos conjuntamente con combustibles fósiles, lo que supone una disminución del uso de éstos), siendo estas últimas las que alcanzan una mayor potencia de generación eléctrica (Lomas et al., 2001; Spliethoff, 2010).

La cogeneración - producción conjunta de energía térmica y eléctrica - se plantea cuando una empresa presenta consumos térmicos y eléctricos importantes, aunque no es una tecnología exclusiva de este sector, ya que aparece cada vez con mayor frecuencia en el sector residencial, comercial y de servicios. Esta tecnología presenta como gran ventaja la consecución de rendimientos globales superiores a los sistemas de producción de energía térmica o eléctrica por separado. Requiere de inversiones importantes y suele instalarse en empresas generadoras de

residuos de biomasa en grandes proporciones (industrias de transformación de la madera, papeleras...) por lo que la contribución de los residuos forestales en esta aplicación es escasa (Lomas et al., 2001; IDAE, 2005).

En el Real Decreto 661/2007 se reglamentaban las tarifas y primas a los productores de energía en función de la fuente energética (organizadas en grupos y subgrupos), además de otros factores como la potencia de la instalación y el tiempo trascurrido desde su puesta en marcha. En el caso que nos ocupa, las biomásas sólidas de origen forestal se agrupaban bajo los epígrafes:

b.6.1.a "Cultivo energético agrícola": Biomasa, de origen agrícola, producida expresa y únicamente con fines energéticos, mediante las actividades de cultivo, cosecha y, en caso necesario, procesado de materias primas recolectadas. Según su origen se dividen en: herbáceos o leñosos);

b.6.1.b "Cultivo energético forestal": biomasa de origen forestal, procedente del aprovechamiento principal de masas forestales, originadas mediante actividades de cultivo, cosecha y en caso necesario, procesado de las materias primas recolectadas y cuyo destino final sea el energético.

b.6.3 "Residuos de aprovechamientos forestales y otras operaciones selvícolas en las masas forestales espacios verdes": Biomasa residual producida durante la realización de cualquier tipo de tratamiento o aprovechamiento selvícola en masas forestales,

incluidas cortezas, así como la generada en la limpieza y mantenimiento de los espacios verdes.

b.8.1 "Biomasa procedente de instalaciones industriales del sector agrícola": Residuos de la producción de aceite de oliva y aceite de orujo de oliva. Residuos de la producción de aceitunas. Residuos de la extracción de aceites de semillas. Residuos de la industria vinícola y alcoholera. Residuos de industrias conserveras. Residuos de la industria de la cerveza y la malta. Residuos de la industria de producción de frutos secos. Residuos de la industria de producción de arroz. Residuos procedentes del procesado de algas. Otros residuos agroindustriales.

b.8.2 "Biomasa procedente de instalaciones industriales del sector forestal": Residuos de las industrias forestales de primera transformación. Residuos de las industrias forestales de segunda transformación (mueble, puertas, carpintería). Otros residuos de industrias forestales. Residuos procedentes de la recuperación de materiales lignocelulósicos (envases, palets, muebles, materiales de construcción,).

Quedaban excluidas del Decreto: 1) los "Combustibles Fósiles", incluyendo la turba, y sus productos y subproductos. 2) "Residuos de madera": a) Tratados químicamente durante procesos industriales de producción b) Mezclados con productos químicos de

origen inorgánico c) De otro tipo, si su uso térmico está prohibido por la legislación. 3) "Cualquier tipo de biomasa o biogás contaminado con sustancias tóxicas o metales pesados". 4) Papel y cartón. 5) Textiles. 6) Cadáveres animales o partes de los mismos, cuando la legislación prevea una gestión de estos residuos diferente a la valorización energética.

Las tarifas y primas de referencia más altas eran las correspondientes a las categorías b.6.1 (cultivos energéticos) para plantas de más de más de 2MW de potencia nominal, mientras que las biomásas provenientes de residuos de aprovechamientos forestales y otras operaciones silvícolas tienen, de media, un 27% menos de prima por kw/h y tarifas reguladas un 20% menores. Todas estas tarifas y primas desaparecieron con el RD 1/2012 menguando aún más las posibilidades reales de valorización energética de la biomasa forestal, especialmente la relacionada con la producción eléctrica.

1.2. La biomasa forestal para uso energético en la Comunidad Autónoma de Euskadi y Bizkaia

La percepción de la existencia de fortalezas y debilidades en lo que respecta a la puesta en valor de la biomasa forestal para usos energéticos varía, como no podía ser de otra manera, entre los distintos agentes intervinientes en los procesos de valorización: productores de materia prima, promotores, comercializadores, administraciones, grupos ecologistas...

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Para los productores de materia prima, - en el caso de Bizkaia básicamente pequeños propietarios forestales - la percepción se centra en los factores de producción. Los aspectos positivos pasan por la consideración de que las existencias y productividad de la biomasa leñosa son muy importantes en términos cuantitativos, aunque la producción real de los bosques resulta muy variable debido a la existencia de suelos muy diversos, diferentes tipos de espesura y especies que genéticamente tienen tasas de crecimiento y producción muy distintas (Tabla 1).

Tabla 1 Tasas de productividad forestal de la CAE

Volumen de madera CAE			Crecimiento de madera CAE			
	Tipo de masa forestal	Volumen (m ³)		Tipo de masa forestal	Producción anual (m ³)	Crecimiento m ³ /ha/año
ALAVA	Coníferas	5.644.380	ALAVA	Coníferas	330.588	8,37
	Fronosas	9.788.965		Fronosas	264.248	2,59
	Todas las especies	15.433.345		Todas las especies	594.835	4,20
GIPUZKOA	Coníferas	13.803.288	GIPUZKOA	Coníferas	1.174.556	16,72
	Fronosas	6.685.903		Fronosas	302.771	5,58
	Todas las especies	20.489.191		Todas las especies	1.477.327	11,86
BIZKAIA	Coníferas	14.997.221	BIZKAIA	Coníferas	1.383.555	16,12
	Fronosas	3.896.749		Fronosas	375.533	8,38
	Todas las especies	18.893.970		Todas las especies	1.759.088	13,46
CAE	Coníferas	34.444.889	CAE	Coníferas	2.888.699	14,77
	Fronosas	20.371.617		Fronosas	942.552	4,69
	Todas las especies	54.816.506		Todas las especies	3.831.251	9,66

Productividad forestal CAE	
Bosques procedentes de regeneración natural:	Especies autóctonas: 2-5 m ³ /ha/año.
	Especies alóctonas: 5-10 m ³ /ha/año.
Bosque procedente de regeneración artificial: Bosque cultivado.	Especies de crecimiento lento: 4-6 m ³ /ha/año.
	Especies de crecimiento medio: 6-20 m ³ /ha/año.
	Especies de crecimiento rápido: 15-30 m ³ /ha/año.

Fuente: Inventario Forestal CAE 2005. Elaboración propia

Sin embargo, la compleja orografía del territorio, encarece notablemente los costes de producción (replantación, tratamientos selvícolas y explotación). Por ejemplo, para Bizkaia, la estimación de los productores de los costes de explotación (tala, saca, desembosque y transporte) para productos teóricamente aprovechables para usos energéticos, así como las diferencias entre precios

de mercado energético/costes de explotación se pueden resumir en la Tabla 2.

Tomando como referencia los precios de la madera entregada en centros de consumo a marzo del 2010 (tras el paso del ciclón "Xhyntia") en los que la madera de trituración se pagaba a 30,00 €/tm, la madera delgada a 40,00 €/tm y la madera gruesa entre 55,00 - 70,00 €/tm, y a pesar de que los precios de madera delgada y gruesa han sufrido históricamente fuertes fluctuaciones, la valorización energética de las biomásas forestales no resultaba especialmente atractiva para los productores.

A esto hay que unirle la fuerte implantación de una industria que hace uso del producto maderable que entra en competencia directa con la valorización energética: (industria papelera, industria del tablero, etc.). Ante esta perspectiva, los productores apostaban por una declaración de cultivos energéticos de algunos tipos de cortas que podrían contribuir a conseguir mejores precios en el mercado energético y hacerlo más competitivo. Otro de los factores que consideraban ineludible es el de la distribución geográfica de las plantas de generación y centros de consumo (para el caso de la valorización térmica) con el fin de minimizar la repercusión del transporte en los costes finales de la materia prima. (ARRIETA, 2010).

El diagnóstico de la administración ponía el foco del problema en la estructura del mercado de la biomasa, un mercado a su juicio, inmaduro, que crea incertidumbre sobre

el aseguramiento del suministro de biomasa, en su momento debido, a las insuficientes retribuciones de la electricidad generada por parte del RD 661/2007 y más aún a partir de la entrada en vigor del RD 1/2012.

La contestación social es también un argumento que se esgrime a menudo como barrera para el desarrollo de este sector. Las administraciones locales, especialmente los municipios, colaboradoras necesarias para asegurar una correcta distribución espacial de la generación, tampoco parecen haber apostado decididamente por impulsar este tipo de plantas, especialmente las de menor tamaño, centrándose casi exclusivamente en la promoción de aplicaciones térmicas de tipo "District heating" (ETXEBARRIA, 2010).

Tabla 2 Costes de producción de la biomasa

Costes de Producción				
TIPO DE PRODUCTO	TIPO DE CORTA	COSTES (€/Tm)		
Arbol entero	C. Intermedia	> 60,00		
Restos de corta	C. Final	> 60,00		

Diferencia entre precios y costes de producción				
TIPO DE MADERA	TIPO DE CORTA	Pr.Un (€/Tm)	COSTES (€/Tm)	Diferencia (€/Tm)
Mad. Trituración (apeas)	C. Intermedia	30,00	50,00	-20,00
	C. Final	30,00	30,00	0,00
Mad. Delgada	C. Intermedia	40,00	40,00	0,00
	C. Final	40,00	25,00	15,00
Mad. Gruesa	C. Final	62,50	22,00	40,50

La perspectiva de los promotores pone hincapié en las reticencias existentes de parte del sector forestal a la

apertura al mercado energético, lo que dificulta que las plantaciones forestales existentes puedan ser valorizadas energéticamente en forma de biomasa. Esto dificulta la madurez de un mercado de la biomasa que permita garantizar el adecuado suministro a medio y largo plazo a las plantas de biomasa. Por otro lado, en muchos casos, la cadena logística de suministro no está definida con certeza, ni en cuanto a los costes asociados, ni en cuanto a sus componentes. Por otro lado, la dispersión de la propiedad de la biomasa complica la consecución de contratos con garantía de suministro. (DÍAZ, 2009).

La valorización energética de la biomasa forestal en Euskadi presentaba a finales de la primera década de los 2000 unas perspectivas de desarrollo muy poco halagüeñas, especialmente en lo que se refería a la producción de energía eléctrica y la co-generación a pesar de las potencialidades del territorio. Una estructura de la propiedad basada en pequeñas explotaciones, con una orografía que dificulta y encarece los costes de producción, unido a la existencia de una industria de transformación de la madera que entra en competencia directa con la de destino energético y a los insuficientes incentivos para la generación de energía renovable de origen forestal, desincentivaba la involucración de los productores en el proceso de valorización energética de la biomasa forestal, dejando al mercado sin la garantía de suministro imprescindible para su maduración.

La aprobación de la Estrategia Energética de Euskadi 2020 plantea incrementar el aprovechamiento de las energías

renovables un 87% para alcanzar en el año 2020 los 905.000 tep, lo que significaría una cuota de renovables en consumo final del 14%. Otro de los objetivos fundamentales de esta estrategia es el de lograr que en 2020 no se superen los niveles de consumo de energía primaria del año 2008 mediante la intensificación de las actuaciones en eficiencia energética. Esto supondría alcanzar un ahorro de 1.050.000 tep anuales para el año 2020 mejorando la intensidad energética final un 22% (datos del 2012).

Aumentar la participación de la cogeneración y las renovables para generación eléctrica de forma que pasen del 18% en el año 2010 al 38% en el 2020 contribuyendo a la mitigación del cambio climático mediante la reducción de 2,5 Mt de CO₂ son compromisos que forman parte de este desiderátum.

Estos objetivos han tenido, de momento, una tímida traducción en políticas públicas concretas. Caben destacar algunas medidas impulsadas desde la administración autonómica y las forales que, han consistido básicamente en ayudas para promocionar el uso de la biomasa, casi de manera exclusiva para fines térmicos, en distintos ámbitos, tanto públicos como privados.

En este sentido habría que destacar el convenio firmado entre HAZI (sociedad pública del Gobierno vasco para impulsar la competitividad y la sostenibilidad del sector primario, alimentario y del medio rural y litoral vasco), el Ente Vasco de la Energía y distintas corporaciones locales para promocionar la implantación de redes de

suministro de calor, de centrales de producción térmica y otros proyectos de instalaciones energéticas de menor entidad, siempre a escala local. Además, se han firmado acuerdos para desarrollar proyectos relacionados como la logística del recurso, como Centrales Logísticas de Biomasa para gestionar los recursos madereros disponibles en los montes de determinados municipios.

Hoy día pocos son los que ponen en cuestión que el aprovechamiento de la biomasa forestal ofrece ventajas que van más allá de la mejora de la autosuficiencia energética o de la contención del balance de CO². Puede resultar de gran ayuda para una gestión sostenible de los montes de utilidad pública tal y como se constata en los países del norte de Europa donde el uso de la esta fuente renovable para fines energéticos está muy extendido. Mediante la limpieza de los montes se asegura su futuro en condiciones sostenibles y se evitan plagas e incendios.

Además, supone una oportunidad para la generación de proyectos que pueden crear empleo estable en el ámbito rural, en la medida en que estas labores de obtención del y gestión del recurso de la biomasa las realicen personas vinculadas a los propios municipios que la utilizarán en sus sistemas de generación de calor.

El uso de recursos locales puede resultar también especialmente útil para generar una industria en torno a las energías limpias y renovables. Un desarrollo que mediante el uso de tecnologías energéticas eficientes cree un sector productivo generador de riqueza y empleo local,

tanto en las labores forestales para la obtención de la biomasa como en la puesta en marcha, la logística o el mantenimiento de las instalaciones energéticas (calderas, etc.).

2. Hipótesis y objetivos

El objetivo general de esta tesis es ofrecer una aproximación a las potencialidades energéticas existentes a escala provincial, comarcal y municipal en el Territorio Histórico de Bizkaia a partir de una propuesta metodológica para la estimación de la Biomasa Forestal Potencial Disponible (BFPD) de las masas forestales existentes de una de las principales especies forestales de repoblación del territorio como es el *Pinus radiata* D. Don (en adelante *P. radiata*).

A su vez, se pretende que los resultados que pudieran surgir del desarrollo de la metodología aquí descrita puedan emplearse tanto en la gestión forestal sostenible, como a la planificación territorial, por ejemplo, mediante la elaboración de Planes Estratégicos de aprovechamiento de biomasa para usos energéticos, habida cuenta de la incidencia que la planificación energética tiene en la ordenación territorial, circunstancia que alcanza su máximo exponente cuando se trata de energías renovables.

Se pretende, por tanto, desarrollar un modelo de estimación de la masa y de su distribución geográfica aplicable a escala local y regional que proporcione una cartografía suficientemente precisa y actualizable sobre los recursos de BFPD, contribuyendo así a resolver una de las barreras más importantes en la puesta en valor energética de este

recurso: su conocimiento preciso en volumen y localización.

Este objetivo se fundamenta en la hipótesis de que, por un lado, es posible estimar, con una precisión aceptable y controlada en términos de exactitud, el volumen de la BFPD por unidad de superficie y su distribución espacial mediante el uso de fuentes de datos provenientes de sensores remotos y trabajo de campo. Que, a través de modelos predictivos, que incluyan variables predictivas/explicativas relacionadas con la altura, la densidad y la estructura vertical de las masas forestales, es posible extrapolar dichos valores a la totalidad del espacio forestal ocupado por la especie objeto de análisis.

Esta hipótesis descansa a su vez en que el hecho de que la estimación de variables forestales en general, y de la biomasa, en particular, viene siendo una línea de trabajo e investigación fructífera en el campo de la teledetección y que los resultados obtenidos para especies semejantes y territorios biogeográficamente similares al del estudio han sido ampliamente aceptados por la comunidad científica.

Aunque la biomasa no puede ser medida directamente mediante datos provenientes de sensores remotos (imágenes, datos LiDAR etc...) multitud de estudios demuestran que la información de parámetros relacionados con la altura y la densidad de la vegetación, extraíbles a partir de datos LiDAR, puede ser relacionada con la cantidad de biomasa por unidad de superficie (t/ha) (Collins, Parker y Evans,

2004; Dong et al., 2003; Latifi et al., 2015; Muukkonen y Heiskanen, 2007; Sorin C Popescu, 2007; Zhao, Popescu y Nelson, 2009a).

3. Planteamiento metodológico

Tradicionalmente las estimaciones de biomasa se han realizado principalmente mediante 2 métodos: directos e indirectos. Los métodos directos o destructivos, suponen la corta de individuos para el posterior pesado de los fustes, ramas y hojas directamente, determinando posteriormente su peso seco (Montero González, Ruiz-Peinado y Muñoz, 2006). Los métodos indirectos o no destructivos se apoyan en datos ya inventariados. Los datos que se emplean más habitualmente son el diámetro del tronco a la altura del pecho, o diámetro normal, y la altura de los árboles situados en las parcelas muestrales. Posteriormente se estiman volúmenes, áreas basimétricas y otro tipo de variables dasométricas.

Entre los métodos indirectos existen varias aproximaciones: una parte de los datos de volumen de madera con corteza (m³/ha) a los que se aplican factores de expansión de la biomasa. El factor de expansión es la ratio entre la biomasa aérea total y la biomasa del fuste. Se emplean cuando no hay información detallada para de las distintas variables que caracterizan un árbol.

Una vez obtenida la biomasa en verde como el producto del volumen inventariado por el factor de expansión, el siguiente paso es estimar la cantidad de biomasa seca, que

se obtiene al multiplicar la cantidad de biomasa húmeda por la densidad promedio específica de las especies (Brown, 1997).

Otra aproximación pasa por la aplicación de técnicas de análisis estadístico en base a variables comúnmente adquiridas en los inventarios forestales: diámetro a la altura del pecho, altura de árbol, altura dominante o el área basal (Canga et al. 2013). Se basan en el principio de alometría: a escala de una población existe una relación estadística entre las medidas de un solo individuo que permite predecir una medida de un árbol (biomasa) en función de otra medida (por ejemplo, diámetro).

Uno de los principales inconvenientes de los inventarios forestales es que conllevan una gran inversión de tiempo y dinero. Para aminorar este problema se han ido desarrollado técnicas basadas en la Teledetección, que permiten la obtención de datos de manera rápida y precisa en zonas extensas (Johansen, Phinn y Witte, 2010). Entre estas técnicas está cobrando cada vez más importancia el LiDAR (Light Detection and Ranging) aerotransportado. Esta metodología se basa en la emisión de un láser que barre la superficie terrestre desde un avión o un helicóptero registrando rebotes de cualquier elemento que encuentra en su trayectoria, registrando nubes de puntos georreferenciadas que permiten caracterizar el dosel vegetal.

Así pues, a pesar de que la biomasa no puede ser medida directamente mediante imágenes de satélite o los datos

proporcionados por los Escáneres Láser Aerotransportados (ELA), la información sobre reflectividades de las cubiertas y la altimétrica proporcionada por éstas, sí puede ser relacionada con la biomasa estimada a partir de trabajo de campo, tal y como se ha demostrado en las numerosas y variadas líneas de trabajo y de investigación que se vienen desarrollando en este ámbito.

La metodología para estimar el volumen de Biomasa Forestal Potencial Disponible (BFCD), objetivo principal de este estudio, pasa por establecer un planteamiento en varias fases:

3.1. Estimación de la Biomasa Forestal Potencial Total

La Biomasa Forestal Potencial Total (BFPT) representa el volumen (t/ha) de biomasa forestal arbolada existente en las parcelas pobladas por la especie objeto de esta tesis. Desde la perspectiva de la valorización energética la BFPT puede cuantificarse y presentarse de varias maneras diferentes, en peso seco, a un determinado grado de humedad (habitualmente la humedad en el momento de la corta) y en toneladas equivalente de petróleo (tep).

No se ha querido olvidar en este trabajo otros servicios prestados a los ecosistemas por los bosques, además del potencial abastecimiento energético¹ : como elemento de control de la erosión, como regulador del clima y de calidad del aire, de regulación hídrica; otros servicios de abastecimiento como la madera, alimentos silvestres

¹ Los servicios de los ecosistemas son los beneficios que las personas obtenemos de la naturaleza y que contribuyen al mantenimiento de nuestro bienestar y calidad de vida.

(setas, caza, frutos etc...); de recreación, estético... y sin duda uno de los más importantes, tanto a nivel local como global, como es el de actuar como almacén de CO². Este es además el que tiene una relación más directa y fácil de estimar a partir de la biomasa forestal. Por ello, a pesar de que el objetivo principal de esta tesis es el de estimar las potencialidades energéticas de la biomasa existente de pino radiata en Bizkaia, se ha considerado oportuno y relevante poder incorporar este dato que contribuye a dar más valor aún, si cabe, a las plantaciones de especies forestales de esta especie.

Para poder realizar estas estimaciones resulta imprescindible recurrir a una metodología que permita aprovechar los datos disponibles, todos ellos de carácter indirecto. Por un lado, al dato estimado de biomasa aérea existente en las parcelas de muestreo obtenidos en la fase de trabajo de campo del Inventario Forestal Nacional (IFN) (variable dependiente). Datos que, como veremos, se obtienen por procedimientos indirectos mediante el empleo de ecuaciones alométricas. Por otro se recurrirá a un conjunto de métricas que se considera que resultarán capaces de comportarse como variables predictoras, y que serán extraídas de la fuente anteriormente citada: los datos LiDAR Aerotransportados.

Una vez extraídas, procesadas y depuradas estas variables se ha de estudiar la relación existente entre la variable dependiente y los distintos tipos de variables predictoras para poder establecer uno o varios modelos predictivos.

La primera dificultad que nos encontramos para poder realizar ese estudio surge del hecho de que en el apeo de las parcelas muestrales de los IFN no se ha venido haciendo ningún tipo de cálculo directo, ni de estimación de la biomasa forestal existente en cada parcela. Por ello, antes de nada, resulta imprescindible aproximarnos a estos valores a partir de estimaciones fiables de los valores de biomasa total de individuos tipo a partir de datos dendrométricas como el diámetro normal (DN) y la altura (h) recogidas, estas sí, en los sucesivos IFN a nivel de árbol durante el apeo de las parcelas.

Este estudio exigiría, como se ha comentado arriba, de un trabajo de campo previo en el que se hiciera el apeo, mediante técnicas destructivas, de un conjunto de individuos muestrales y el posterior desarrollo de ecuaciones alométricas que permitieran establecer la relación matemática existente entre los pesos de biomasa seca de cada individuo y una o más variables dendrométricas representativas. Este trabajo de campo resulta inabordable para un estudio de estas características en términos económicos, logísticos y temporales por lo que se ha optado por utilizar los datos proporcionados por tres estudios:

Por un lado, los datos recogidos en un estudio de gran alcance, recogidos en (Montero González et al., 2006) en el que se presenta la biomasa fijada en las diferentes partes del árbol para las principales especies forestales españolas y el CO² fijado en las mismas, que para el caso de la biomasa total de la especie *P. radiata* se obtuvieron valores de R² de 0,977 y un RMSE de 0,19 (en unidades ln),

y valores de R^2 de entre 0,669 y 0,976 para las distintas fracciones de biomasa, con un RMSE de entre 0,2 y 0,62.

Por otro, habida cuenta de que los modelos pueden presentar grandes variaciones locales, se ha optado por calibrar los modelos de estimación a partir de ecuaciones alométricas desarrolladas por (Canga, Diéguez-Aranda, Elias y Cámara, 2013) basadas en un estudio realizado en Asturias, zona con unas características bioclimáticas y edáfico-ecológicas muy semejantes al área de estudio de esta tesis.

Por último, se ha considerado el modelo de cálculo de biomasa a nivel de árbol realizado en la "Unidad de Xestión Forestal Sostenible" de Pérez-Cruzado, 2008, que se encuentra en el informe provisional del "Proyecto para la realización de acciones encaminadas a evaluar y promover la utilización de biomasa forestal para la producción de energías renovables en el País Vasco. U.P.M. - BASOLANAK ELKARTEA."

En el estudio de (Canga et al., 2013) la muestra estaba compuesta por 27 árboles, pertenecientes a la especie *P. radiata*, repartidos en 5 rodales diferentes. Aplicaron un método del muestreo destructivo semejante al empleado por Montero et al 2006, obteniendo submuestras para cada uno de los diferentes componentes de la biomasa (madera, corteza, y corona), que una vez secas fueron pesadas. Posteriormente, desarrollaron modelos alométricos para cada fracción de la biomasa utilizando como predictores el diámetro al cuadrado (d^2), o el producto de la altura por el diámetro al cuadrado ($d^2 \cdot h$). Los valores para el

coeficiente de determinación para cada uno de los componentes, oscilaba entre 0,77 y 0,92 con un RMSE de entre 27,40 kg y 110,2 kg, en el caso de los componentes estudiados por separado. Finalmente, el modelo que evaluaba la biomasa total, aportaba valores R² de 0,9376 y un RMSE de 139 kg, teniendo en cuenta que la cantidad de biomasa seca medida en el laboratorio era de 605.2 kg.

El modelo de Pérez Cruzado, 2008 se basa en seis parcelas de Galicia con un abanico de diámetros entre 9,5 y 73,2. Para el ajuste de las ecuaciones obtenidas se utilizó una metodología de ajuste simultáneo empleando el método de mínimos cuadrados conjuntos (Seemingly Unrelated Regressions, SUR) para asegurar la condición de aditividad, y se tuvo en cuenta además la falta de homogeneidad de la varianza que habitualmente presentan este tipo de relaciones, para cuya corrección se empleó la metodología del ajuste ponderado.

A continuación, se explica, a modo de ejemplo de este tipo de metodologías de cálculo de ecuaciones alométricas, la base metodológica del modelo de (Montero González et al., 2006) como ejemplo los resultados detallados del estudio de para el *P. radiata*:

a.- En primer lugar, se lleva a cabo un muestreo destructivo realizado en una muestra de varios árboles de diferentes diámetros. Dicha estimación, para cada especie, incluye árboles pertenecientes a una masa de densidad media y con una calidad de estación característica de esa especie. La muestra debe estar

formada por árboles con un desarrollo normal para la especie, sanos, equilibrados, excluyéndose pies debilitados, mal conformados o que afectados por efecto borde. La muestra debía ser de al menos tres pies por clase diamétrica para calcular la biomasa aérea, y de un pie por clase diamétrica para calcular la biomasa radical. De los árboles elegidos se medía el diámetro normal en pie y la altura total tras el apeo. Una vez apeados se desraman y se separan en diferentes fracciones de biomasa.

b.- Posteriormente se corta, trocea y se pesan todas las fracciones del árbol (Valentine, Tritton y Furnival, 1984). A continuación, se pesan las fracciones seleccionadas y se estima el peso total del árbol. Una vez conocidos los pesos de las distintas fracciones se ajusta un modelo matemático que relaciona los pesos de biomasa seca con una o más variables representativas de estos árboles. A pesar de que existen gran variedad de modelos para estimar la biomasa de un árbol o de sus fracciones, todas las ecuaciones proceden de alguna de las siguientes funciones matemáticas (Diéguez Aranda et al., 2003):

- Lineal: $b = b_0 + b_1 \cdot x_1 + \dots + b_j \cdot x_j$
- No lineal: $b = b_0 \cdot x_1^{b_1} \cdot x_2^{b_2} \cdot \dots \cdot x_j^{b_j}$
- Logarítmica: $\log b = b_0 + b_1 \cdot \log x_1 + b_2 \cdot \log x_2 + \dots + b_j \cdot \log x_j$

donde:

b = biomasa seca total o de alguno de los componentes del árbol.

x_j = variables dendrométricas explicativas.

b_j = parámetros del modelo

Las variables explicativas más utilizadas son el diámetro normal, el diámetro normal al cuadrado, la altura total, el producto entre el diámetro normal al cuadrado y la altura, la edad y la altura de copa viva (Parresol, 1999). La ecuación más comúnmente utilizada es la alométrica (potencial) que relaciona la biomasa seca total o la biomasa seca de alguno de los componentes del árbol con el diámetro normal. Esta ecuación tiene esta forma:

$$b = a \cdot d^\beta$$

Esta ecuación al no ser lineal, debe linealizarse tomando logaritmos en ambos términos de la ecuación, resultando una ecuación de la forma:

$$\ln b = \alpha + \beta \cdot \ln d$$

donde:

b = biomasa seca total o de alguno de los componentes del árbol.

d = diámetro normal.

α, β = parámetros estimados para el modelo.

\ln = logaritmo neperiano.

c.- El siguiente paso consiste en determinar la materia seca. A partir de varias muestras de biomasa de cada fracción calculan, en laboratorio, el porcentaje de humedad. La determinación del peso seco de las muestras se realiza aplicando el porcentaje de humedad a los pesos en verde tomados en campo.

d.- A continuación, se calculan los valores modulares de biomasa a partir de las mencionadas funciones de regresión para cada fracción de biomasa, que relacionan su peso seco con su diámetro normal. En el caso del estudio de Montero González eligieron la ecuación alométrica, linealizada como modelo logarítmico, ya que, aunque en algunos casos el coeficiente de determinación sea menor que en otros modelos, daba mejores resultados al aplicarla en las clases diamétricas menores. Como variable independiente eligieron el diámetro normal, para una mejor aplicación de estas funciones a los datos del Inventario Forestal Nacional (IFN). Esta función (logarítmica) simplifica los cálculos, y además, en la mayoría de los casos, incrementa la bondad del análisis estadístico. Sin embargo, la transformación logarítmica introduce un sesgo (subestimación) en los cálculos (Baskerville, 1972). Para la eliminación de este sesgo, el resultado final debe ser multiplicado por un factor de corrección (CF), calculado a partir del error estándar de la estimación (SEE). Este valor del coeficiente de corrección es calculado según la siguiente expresión (Sprugel, 1983):

$$CF = e^{-\frac{SEE^2}{2}}$$

Por tanto, desde la ecuación logarítmica se llegó a obtener los valores modulares de biomasa (kg de materia seca) con la re-transformación y aplicación del coeficiente de corrección, utilizando los parámetros que aparecen recogidos en la tabla 3.

$$\begin{aligned} \ln b &= a + b \cdot \ln d \\ \Downarrow \\ b &= CF \cdot A \cdot d^b \end{aligned}$$

donde: $A = e^a$

Tabla 3 Valor de los parámetros de la función (a y b), del coeficiente de determinación ajustado (R² adj) y el error estándar de la estimación (SEE) para cada especie y fracción de biomasa

Biomasa	Parámetros		R ² adj	SEE
	a	b		
<i>P. radiata</i>				
BT	-2,61093	2,48739	0,977	0,193270
BF	-3,02878	2,56358	0,976	0,200075
BR ₇	-10,5693	3,64861	0,710	0,525330
BR ₂₋₇	-4,12515	2,1173	0,746	0,615400
BR ₂	-3,53532	1,75877	0,669	0,616072
BA	-5,03445	2,05803	0,739	0,609518
Br	-2,78485	2,14449	0,939	0,309544

Leyenda de símbolos usados en la tabla 3, expresados en kg de materia seca:

BT = biomasa aérea total del árbol.

BF = biomasa del fuste.

BR₇ = biomasa de las ramas de diámetro mayor de 7 cm.

BR₂₋₇ = biomasa de las ramas de diámetro entre 2 y 7 cm.

BR₂ = biomasa de las ramas de diámetro menor de 2 cm.

BA = biomasa de las acículas.

BH = biomasa de las hojas.

Br = biomasa de la raíz.

Las funciones desarrolladas se aplicaron a los valores medios de las clases diamétricas que establece el Segundo Inventario Forestal Nacional (los cuales coinciden con los del IFN3), para obtener los valores modulares de biomasa. Estas clases van desde la clase 10 hasta la clase 70 cm, con una amplitud de 5 cm. Se incluyeron los pies menores (con diámetro menor de 7,5 cm) en la clase diamétrica de 5 cm y los pies mayores (con diámetro mayor de 72,5 cm) en la clase 70 (tabla 4).

Tabla 4 Valores modulares de las distintas fracciones de biomasa (kg de materia seca) en España para *P. radiata*

CD (cm)	Biomasa aérea					Biomasa radical	Biomasa total	
	Fuste	Ramas			Acículas			Total aérea
		R > 7 cm	R 2-7 cm	R < 2 cm				
5	2,8	—	0,5	0,5	0,2	4,1	2,0	6,1
10	17,6	—	2,5	2,0	0,9	23,0	9,0	32,0
15	50,4	0,6	6,0	4,1	2,0	63,0	21,6	84,6
20	105,8	1,6	11,0	6,8	3,7	128,9	39,9	168,9
25	187,4	3,7	17,6	10,0	5,8	224,6	64,5	289,1
30	298,2	7,1	25,9	13,8	8,5	353,5	95,3	448,8
35	440,9	12,5	35,7	18,0	11,6	518,7	132,6	651,3
40	617,8	20,2	47,1	22,7	15,2	723,0	176,6	899,6
45	831,1	30,9	60,2	27,7	19,3	969,1	227,3	1.196,5
50	1.082,7	45,1	74,8	33,2	23,8	1.259,5	285,0	1.544,5
55	1.374,2	63,5	90,9	39,0	28,8	1.596,5	349,6	1.946,0
60	1.707,4	86,7	108,7	45,2	34,2	1.982,2	421,3	2.403,5
65	2.083,7	115,4	128,0	51,7	40,1	2.418,9	500,2	2.919,1
70	2.504,4	150,4	148,8	58,5	46,4	2.908,6	586,3	3.494,9

(Montero González et al., 2006)

1.2.- Una vez establecida la fuente a partir de la cual se integrarán los valores de biomasa por unidad de superficie existente en las parcelas del Cuarto Inventario Forestal nacional (IFN4), el siguiente paso necesario para el cálculo de la BPPT de *P. radiata* pasa por obtener un conjunto de variables predictoras candidatas. Estas se extraerán de datos LiDAR aerotransportados que, como se presume, deberían estar correlacionadas con la biomasa y ser capaces de explicar su comportamiento a nivel de masa.

1.3.- El siguiente paso de esta fase implica buscar un método que permita relacionar de manera satisfactoria la información de biomasa estimada en las parcelas del IFN4 con las variables radiométricas. Se pretende en esta fase analizar diferentes métodos de extracción de la información radiométrica -que será relacionada con la de biomasa- con el objetivo de reducir al máximo la heterogeneidad de la muestra y, por tanto, establecer ajustes más robustos desde el punto de vista estadístico.

1.4.- Una vez elegido el método de extracción, el estudio se centrará en examinar el comportamiento predictivo de diferentes tipos de modelos de regresión a partir del análisis de su uso en casos similares en la literatura científica, seleccionando, en un primer nivel, aquellos que han demostrado explicar y predecir de manera más robusta este tipo de relaciones, para posteriormente seleccionar aquel o aquellos que demuestren un mejor comportamiento. Esto implica, en primer lugar, el estudio de las correlaciones de las variables predictoras entre sí

y posteriormente el análisis de los distintos tipos de modelos de regresión.

1.5.- El último paso de esta fase metodológica consistirá en la estimación, cartografiado e inventario de la Biomasa Forestal Potencial Total para la totalidad del Territorio Histórico de Bizkaia. Para ello se usará como base de referencia, dependiendo de los parámetros del modelo elegido, los datos LiDAR ajustados a celdas de la misma superficie que la de las parcelas de campo del IFN4 para ajustarlos a las masas de *P. radiata* que han sido cartografiadas simultáneamente a la realización del inventario. La cartografía resultante servirá de base para la siguiente fase del trabajo.

3.2. Estimación de la Biomasa Forestal Potencial Accesible

La Biomasa Forestal Potencial Accesible (BFPA) sería la resultante de aminorar la total considerando restricciones de índole ecológica, económica o de mecanización (figuras de protección, pendientes, productividad forestal, etc.); así como aplicar un coeficiente reductor (coeficiente de recogida²) debido a la eficiencia en la recogida, ya que se deben dejar restos en el terreno por razones ecológicas además de la posibilidad de no ser recogido el 100% del material (mayor dificultad por pendiente, etc.) En el

² Hace referencia a la eficiencia de la recogida de la biomasa, es decir, el porcentaje final de restos que se extrae de las masas tras cada tratamiento.

estudio Técnico de Evaluación de Energía de la Biomasa se fija este coeficiente en un 65% (Cabrera et al., 2011).

En esta fase se pretende determinar las zonas que resulten accesibles a la Biomasa Forestal Potencial Total estimada en el procedimiento anterior, delimitando las zonas más adecuadas para la extracción de biomasa forestal en función de la cantidad existente y de los factores que condicionan o limitan su explotación desde el punto de vista técnico como son:

Limitaciones por pendiente: Por razones económicas, de mecanización y de erosión es habitual que se limite el aprovechamiento a terrenos con pendiente igual o menor al 30% excepto en zonas de altas productividades (clases I y II del Mapa de Productividad Potencial Forestal, vid Figura X) donde los aprovechamientos pueden llegar hasta el 50% de pendiente en formaciones de *P. radiata* y otras especies como el *P. pinaster* y *Eucalyptus spp.* Dependiendo del sistema logístico³ se podría llegar a alcanzar hasta el 75%.

En la Cornisa Cantábrica, el 40% del transporte se lleva a cabo con biomasa bruta para su astillado o triturado fijo en planta o terminal logístico. Esta opción suele ser

³ Un sistema logístico define la serie de procesos (apeo, reunión, saca, adecuación y transporte) y maquinaria asociada (motosierra o multitaladora, peón o tractor recogedor, tractor agrícola o autocargador, empacadora, astilladora o trituradora y camión rígido, tráiler, camión con remolque o piso móvil) para la obtención de un producto final previamente definido a pie de central o parque de almacenamiento. De modo que un sistema logístico será diferente a otro según los procesos que compongan y la maquinaria con la que se lleve a cabo. Por ejemplo, el proceso de reunión puede ser necesario o no, y a su vez se puede realizar manual o mecanizadamente.

la más empelada para montes privados de menor tamaño, y para distancias menores a 30 km. Otro 40% se empaca y se transportan las pacas a planta o terminal logístico donde son tratadas hasta alcanzar los requerimientos del producto final. Esta opción es la que se suele emplear para distancias mayores y grandes consumidores a partir de montes privados. El 20% restante se suele tratar como astillado o triturado fijo en cargadero y transporte de astillas, especialmente en montes públicos o de cierto tamaño, con cargaderos disponibles en el propio monte o fincas cercanas (Cabrera et al., 2011).

Limitaciones relacionadas con la distancia a desembosque:

en la medida que pueden dificulten en extremo o hacer improductiva la extracción y el transporte de la biomasa forestal. Para ello se emplearán distintos tipos de análisis geoespacial mediante herramientas SIG.

Condicionantes jurídico-administrativos: refiere a los espacios protegidos y las normas sectoriales que establecen limitaciones a la explotación o determinados condicionantes a prácticas selvícolas en masas forestales de especies potencialmente aprovechables: En el caso de Bizkaia se considerarán los Parques Naturales de Gorbeia y Urkiola, las Zonas de Especial Conservación declaradas de Ordunte y Armañón y la Reserva de la Biosfera de Urdaibai. Las normativas sectoriales que tienen mayor incidencia en la gestión forestal son las Normas y Decretos Forales en materia de montes de la Diputación Foral de Bizkaia y el Plan Territorial Sectorial Agroforestal aprobado definitivamente por el Gobierno Vasco en 2014,

(Decreto foral 23/1990, de 8 de mayo, sobre Normas Provisionales de Evaluación de Impacto Ambiental por actuaciones agropecuarias y forestales que afectan al medio natural del Territorio Histórico de Bizkaia. BOB, 118 de 5 de junio de 1990., 1990; Decreto Foral 101/1994 de 20 de diciembre, por el que se determinan las distancias entre plantaciones forestales y fincas colindantes. BOB, 4 de 5 de enero de 1995, págs. 145-147., 1994; Norma Foral 3/1994, de 2 de junio de Montes y Administración de Espacios Naturales Protegidos. BOB, 123 de 28 de junio de 1994, págs. 7839-7873., 1994; Decreto Foral 52/95 de 30 de mayo, sobre Autorización de Repoblación Forestal en montes públicos no catalogados y particulares no protectores. BOB, 115 de 16 de junio de 1995, págs. 8677-8679., 1995; Decreto 177/2014, de 16 de septiembre, por el que se aprueba definitivamente el Plan Territorial Sectorial Agroforestal de la Comunidad Autónoma del País Vasco, 2014).

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

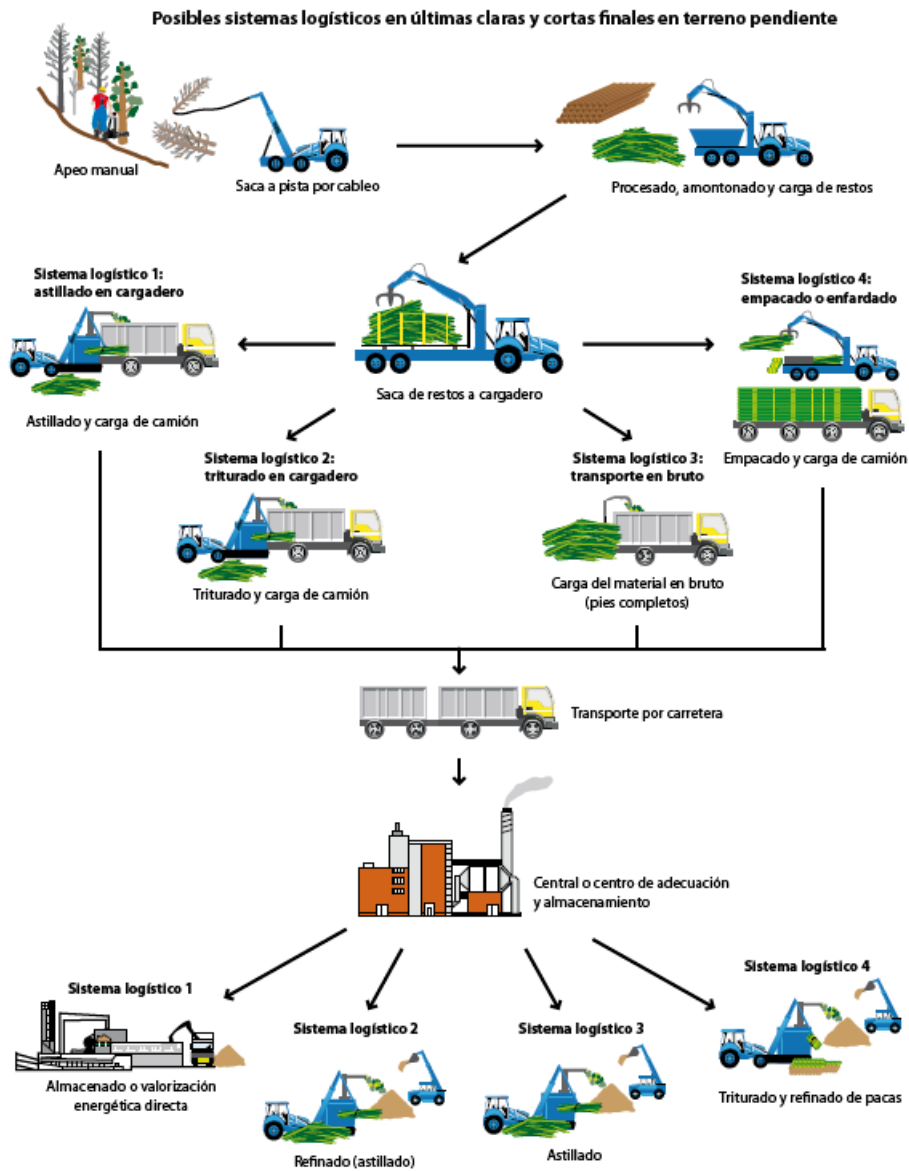


Figura 2 Esquema de los posibles sistemas logísticos en últimas claras y cortas finales en terreno en pendiente. (Cabrera et al., 2011)

Consideramos que sería deseable tener en cuenta otros factores que en la actualidad no tienen, desde el punto de vista normativo, determinaciones claras que condicionen o limiten la explotación forestal del *P. radiata*. Entre ellas cabría destacar la Red de Corredores Ecológicos de la CAE

en los que sería deseable su relevo con especies autóctonas que contribuirían a mejorar la conectividad ambiental (Gurrutxaga San Vicente y IKT S.A., 2005).

3.3. Estimación de la Biomasa Forestal Potencial Disponible

La Biomasa Forestal Potencial Disponible (BFPD) sería aquella biomasa accesible que no tiene un uso maderero, es decir, la biomasa accesible que no entra en competencia con las cortas madereras que se llevan a cabo en la actualidad.

El cálculo de esta biomasa se debería llevar a cabo en a partir de la biomasa accesible minorada por un coeficiente de cortas actuales de madera destinadas a otros usos, fundamentalmente madereros. (Cabrera et al., 2011) establecían un coeficiente de cortas del 41% para aquellas especies en las que el aprovechamiento de la biomasa no sea de árbol completo. Este sería el resultante del cociente entre las cortas actuales de madera y la disponibilidad de madera estimada en el IFN. Para el caso del *P. radiata* en la Cornisa Cantábrica establece este coeficiente en un 65% para aprovechamientos de biomasa de pies con diámetros inferiores a 7 cm.

CAPÍTULO 3 ANTECEDENTES Y ESTADO DE LA CUESTIÓN

4. La estimación de la biomasa en el contexto de la teledetección

Los objetivos de esta investigación exigen en primer lugar de una amplia revisión de la literatura científica relacionada con la aplicación de la teledetección en la estimación de la biomasa, la cual se enmarca en el más amplio contexto de la estimación de los denominados parámetros forestales continuos y la información biofísica. Alcanzar los objetivos formulados supone dedicar a esta labor un considerable esfuerzo recogido en este apartado.

Tras hacer alguna breve referencia básica a los fundamentos de la teledetección, se exponen alguna de las principales aplicaciones forestales. A continuación, se analizan estas en el contexto de los inventarios forestales y, finalmente, se exponen en algunas de las más citadas experiencias en relación con la estimación de biomasa.

4.1. Fundamentos de la teledetección

El término teledetección espacial se emplea para designar el conjunto de técnicas que permiten adquirir e interpretar imágenes de la superficie terrestre obtenidas desde sensores transportados en satélites, empleando para ello la energía electromagnética como medio para detectar y medir determinadas características y propiedades de los objetos (Chuvieco, 2008).

Un sistema de teledetección espacial se compone de: una fuente de energía, origen de la radiación electromagnética que capta el sensor; la cubierta terrestre, que recibe la radiación electromagnética y la refleja o emite conforme sus características físicas; un sistema sensor, que capta esa energía procedente de los objetos y que se encuentra instalado en un satélite; un sistema de recepción-comercialización, que almacena los datos adquiridos por el sensor y, finalmente, un intérprete o usuario, que analiza y convierte esos datos en información (Chuvienco, 2008).

La utilización de la energía electromagnética permite la observación y el análisis del territorio desde una perspectiva imposible con la simple utilización del ojo humano, aportando nuevas dimensiones de análisis. Aunque el espectro electromagnético es un sistema continuo, la teledetección lo separa, a efectos prácticos, en una serie de regiones en la cuales la radiación electromagnética presenta un comportamiento similar. Su denominación y rango espectral más aceptados son los siguientes: espectro visible (de 0,4 a 0,7 μm), infrarrojo próximo (de 0,7 a 1,3 μm), infrarrojo medio (de 1,3 a 8 μm), infrarrojo térmico (de 8 a 14 μm) y micro-ondas (a partir de 0,1 cm). (Chuvienco, 2008).

El criterio más empleado para la clasificación de los distintos sensores empleados en teledetección hace referencia al origen de la energía que captan. Así pues, existen dos tipos de sensores: activos y pasivos. Los sensores activos se caracterizan por emitir energía, fundamentalmente en la región de las micro-ondas, que es

reflejada por la superficie terrestre y es posteriormente captada por el sensor. El principal grupo de sensores de este tipo son los denominados sensores SAR (Synthetic Aperture Radar). Dentro de este tipo de sensores activos se encuentran los denominados LiDAR (Light Detection and Ranging) que operan en longitudes de onda que van desde los 500 a los 1500 nanómetros; de la luz ultravioleta a la infrarroja. Los sensores pasivos son aquellos que solamente recogen la energía electromagnética procedente de la superficie terrestre, ya sea ésta reflejada del Sol o emitida por ella en función de su temperatura (Chuvieco, 2008).

Otro criterio complementario para clasificar los sensores es la resolución espacial de las imágenes que proporcionan: baja, media y alta. La resolución espacial hace referencia al tamaño de píxel de la imagen de satélite, que es la unidad mínima de información en la imagen (Chuvieco, 2008). Existen diversos umbrales para incluir un sensor en uno de los tres grupos de resolución espacial aludidos, por lo que serán considerados sensores de baja resolución aquellos cuyas imágenes están formadas por píxeles de más de 100 m, de media resolución aquellos entre 10-100 m y de alta resolución los que proporcionan imágenes con píxeles inferiores a 10 m. La superficie de observación real de los distintos sensores (área de terreno registrada) está en relación directa con el tamaño del píxel; así, los sensores de alta resolución espacial que proporcionan píxeles de menos de 10 m recogen información simultánea de áreas pequeñas, mientras que los sensores de baja

resolución cubren en una sola imagen grandes extensiones de terreno (Chuvieco, 2008).

Por otro lado, los sensores también se caracterizan por su resolución espectral, radiométrica y temporal. La resolución espectral hace referencia al número y anchura de las bandas de información espectral que recoge el sensor; cuantas más sean y más estrechas, más sencilla resultará la caracterización espectral de las cubiertas registradas. Por su parte, la resolución radiométrica alude a la sensibilidad del sensor para codificar la información espectral recibida en niveles digitales (ND); una vez más, a mayor resolución radiométrica, más fácil será la identificación de cubiertas y la caracterización de las variables a estudiar en la imagen. Finalmente, la resolución temporal refiere a la frecuencia con la cual el sensor proporciona imágenes de una misma zona; cuanto menor sea el periodo de revisita, más fácil será la observación de determinados fenómenos caracterizados por su rápida evolución temporal (Chuvieco, 2008).

4.2. La teledetección laser

La base de todos los sistemas LiDAR es la emisión de un pulso de corta duración de la luz láser y la medición del tiempo transcurrido entre ésta y la detección de la parte de la luz reflejada captada por el sensor. Usando la velocidad de la luz, el tiempo transcurrido se convierte a la distancia desde el punto de emisión al objeto subyacente de que se reflejaba la luz. Debido a que la posición y la orientación del sensor se registran

utilizando un GPS y un INS, los datos LiDAR constituyen un conjunto de coordenadas 3D que representan la superficie escaneada desde el cual se reflejan los pulsos.

Los sistemas LiDAR utilizados en aplicaciones forestales suelen ser de huella estrecha, sistemas de retorno discreto que registran uno a unos los ecos de cada pulso emitido. El diámetro del haz de láser emitido en el suelo, caracterizado como el haz de huella, depende de la divergencia del haz y la altitud de vuelo. Los sistemas propietarios suelen tener diámetros de huella de 1 m o menos y son a menudo los preferidos para aplicaciones forestales porque facilitan vínculos precisos entre los datos de escaneado y los árboles individuales, parcelas, o masas forestales. En contraste, los sistemas de huella ancha tienen típicamente diámetros de la huella alrededor de 10s metros de diámetro. El espaciamiento de los pulsos en el suelo depende de parámetros como la altitud de vuelo y velocidad, frecuencia de repetición del pulso, el ángulo de visión, y el patrón de barrido. (Matti Maltamo, Erik Næsset y Jari Vauhkonen, 2014).

Los pulsos de láser emitidos generan diferentes tipos de datos mientras son reflejados desde el suelo y grabados en el sensor, dependiendo de la interacción con el dosel del bosque y si el sensor tiene funciones de forma de onda completa o es un sistema de retorno discreto. En términos simplificados, los sistemas de grabación de forma de onda completa digitalizan toda la secuencia de energía devuelta como una función de la altura, mientras que la mayoría de los sistemas empleados actualmente en aplicaciones

forestales son de retorno discreto; sistemas que graban de uno a unos varios (generalmente de cuatro cinco a) ecos discretos reflejados entre la superficie terrestre y la parte superior de la copa. Las características de los datos de los sistemas de retorno discreto almacenados dependen de factores tales como el algoritmo utilizado en tiempo real para desencadenar y grabar el eco de la señal de retrodispersión recibida en el sensor. Dependiendo del sensor, pueden incluso pueden ser grabadas informaciones adicionales tales como la intensidad de los ecos. (Matti Maltamo et al., 2014).

En estos últimos años, el LiDAR aerotransportado se ha implantado en trabajos altimétricos por su precisión y productividad. La tecnología se basa en un sensor láser sobre una plataforma móvil (helicóptero o avión) que barre el área de trabajo que emite pulsos regularmente y registra los rebotes o ecos que se producen sobre la superficie impactada.

4.3. El LiDAR Aerotransportado

La principal ventaja del láser escáner es el alto número de puntos obtenidos, así como la gran homogeneidad espacial que proporciona. Su principal desventaja, la dificultad de identificar el tipo de superficie sobre la que impacta. Para aminorar este problema la toma de datos LiDAR se suele acompañar con la toma de imágenes del área levantada.

El funcionamiento del LiDAR se basa en un sensor de barrido que emite pulsos láser y mide el tiempo que tardan en

llegar a la superficie terrestre y volver hasta el sensor. El sensor adquiere simultáneamente datos precisos sobre la dirección y posición en que la señal fue emitida y el eco recibido.

Los pulsos emitidos por los sensores LiDAR son, en principio, puntuales, pero a medida que se alejan del emisor el haz toma una forma cónica que, al producirse el impacto, la superficie detectada será una circunferencia dentro de la cual se halla el punto que se considera localizado.

Esta circunferencia se denomina "huella" y su tamaño depende de la altura de vuelo y de las características y disposición del sensor. Esto limita la altura de vuelo a un rango aproximado de entre 500 y 3.500 metros para evitar cometer errores demasiado grandes en la ubicación planimétrica de los puntos. Es por ello que con estos sistemas la precisión altimétrica es mayor que la planimétrica.

El Láser Escáner Aerotransportado (ALS) es un procedimiento de toma de datos que se compone de un sensor láser emplazado en una plataforma móvil que barre las superficies situadas entre el sensor y la topografía, que obtiene una nube de puntos de coordenadas de alta precisión.

El equipo está compuesto por un sensor láser, un sistema de navegación inercial (INS) y un receptor GPS en una plataforma móvil. Para obtener la máxima precisión se dispone en tierra una red GPS para trabajar de forma

conjunta con el receptor GPS de la plataforma (Fatoyinbo T., 2014).

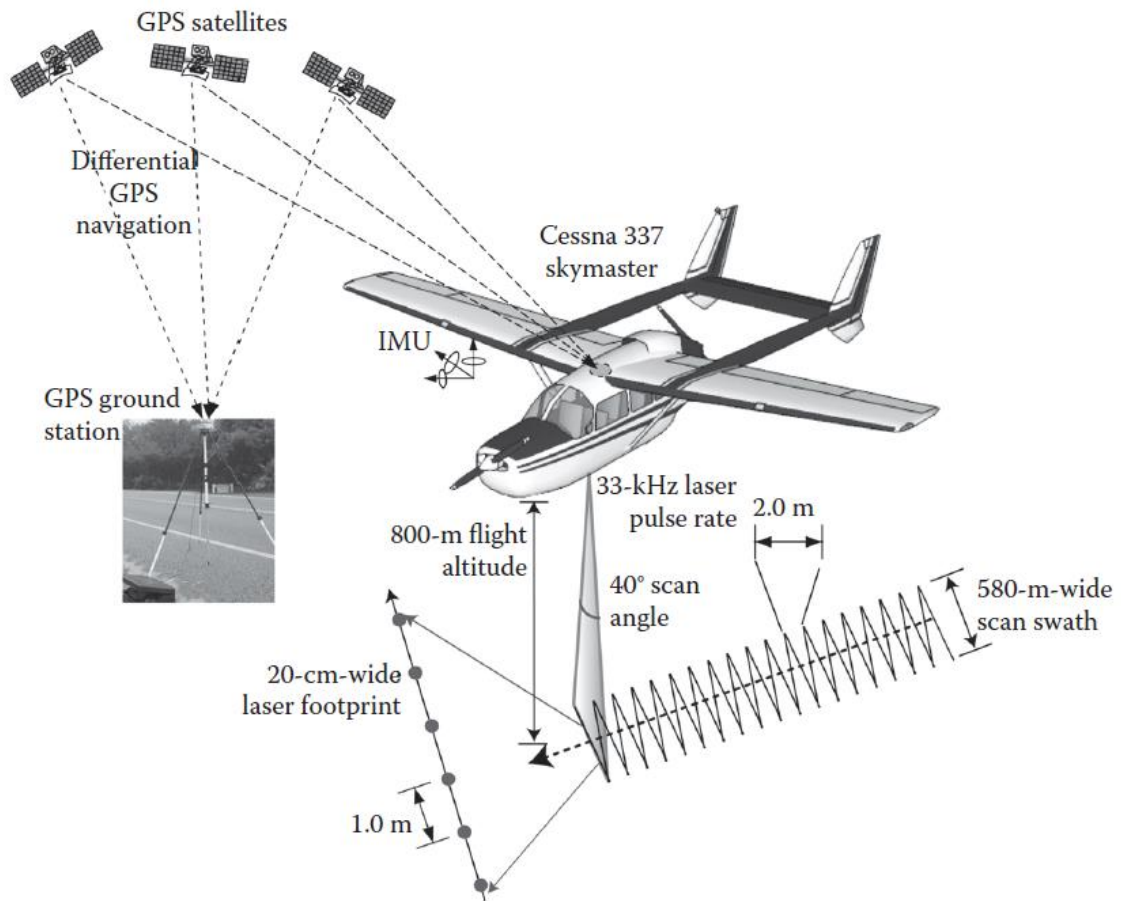


Figura 3 Sistema LiDAR aerotransportado (Matti Maltamo et al., 2014)

4.3.1. Características técnicas

Los parámetros más importantes que caracterizan un sensor LiDAR aerotransportado son:

Frecuencia de pulsos láser. Es el número de puntos por segundo que es capaz de emitir el sistema. Actualmente los sistemas llegan hasta 500 kHz siendo

comunes valores entre 100 y 250kHz. Este parámetro combinado con la velocidad del vuelo y el porcentaje de solape entre bandas de barrido determina la densidad de pulsos por metro cuadrado que se alcanzará. Los sistemas modernos pueden trabajar con varios pulsos simultáneamente, pudiendo emitir un nuevo pulso antes de haber recibido la información del rebote del anterior, consiguiendo así aumentar la densidad de pulsos.

Número de retornos o rebotes por pulso. Del reflejo de cada pulso laser emitido el sistema puede extraer entre uno y cinco o seis retornos individuales. Los sistemas actuales pueden ser configurados para extraer de cada pulso varios rebotes o retornos. Esto se consigue configurando el sistema para que diferencie los máximos relativos de la onda rebotada por cada pulso emitido.

En los sistemas multiretorno, la cantidad de rebotes recibidos para una misma densidad de pulsos emitidos variará dependiendo del tipo de superficie sobre la que impactan las señales. Un pulso emitido que rebote sobre una superficie desnuda como una carretera, un edificio (siempre que no sea su borde) o sobre pastizales bajos, concentrará toda la energía en un único máximo y el número de rebotes en esta zona será igual al de pulsos.

Los pulsos que impacten contra vegetación alta encontrarán en su camino antes de llegar al suelo,

varios obstáculos contra los que irá haciendo reflexiones parciales de la energía que contiene y por tanto sumará al total varios retornos por cada pulso emitido.

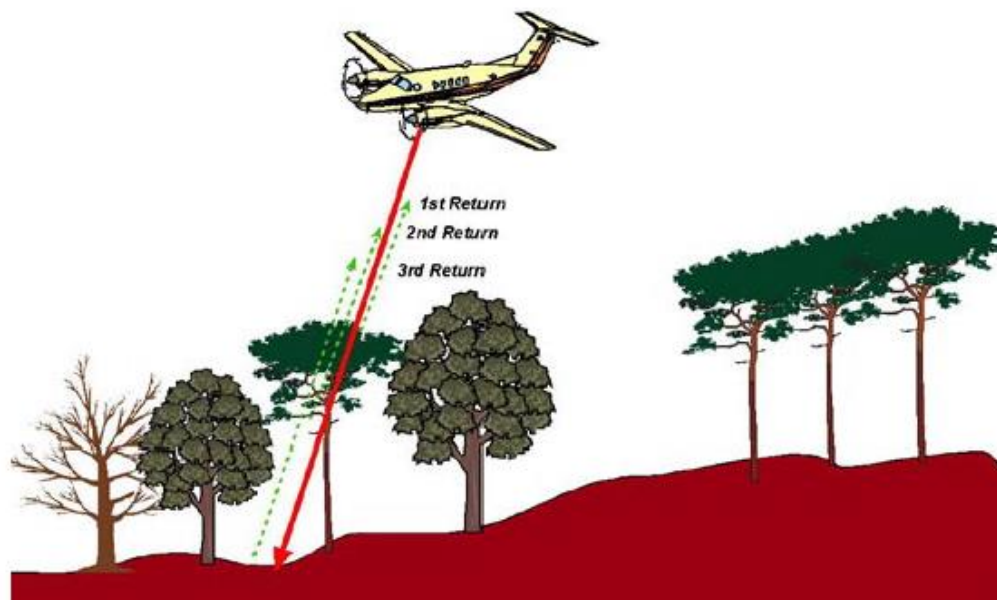


Figura 4 Retornos en vegetación de los pulsos láser (Matti Maltamo et al., 2014)

Esta propiedad es muy útil para el estudio de la estructura de las copas en masas forestales, pero hay que considerar que la distancia vertical que separa dos rebotes del mismo pulso no puede ser menor que la que necesita el sistema para poder discriminar dos máximos consecutivos en la forma de la onda de reflexión del pulso (Fatoyinbo T., 2014).

Esta distancia suele ser de dos o tres metros por lo que la vegetación baja de matorral tendrá una proporción mucho más bajas de segundos y posteriores retornos que la arbórea. Se debe tener en cuenta esta

distancia mínima si se pretende evaluar la altura de la base de las copas ya que la indeterminación resultante será al menos la de esta distancia.

Frecuencia del ciclo de escaneado. Número de ciclos por segundo del sistema que desvía la señal láser (espejo oscilante, polígono rotatorio...) para cubrir una franja de terreno. Se ajusta a la velocidad de vuelo siendo normales valores entre 10Hz y 120Hz.

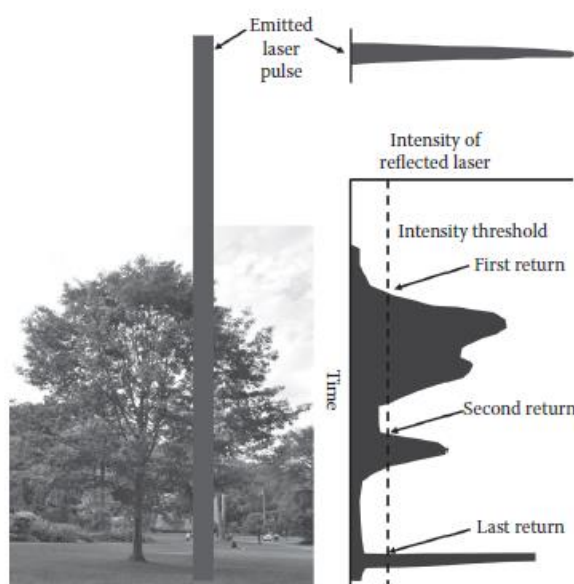


Figura 5 Forma de onda de respuesta LiDAR (Matti Maltamo et al., 2014)

Divergencia del haz (IFOV). La señal laser emitida por el sensor es inicialmente puntual, pero en el camino hacia la superficie se abre formando un cono cuya base tiene un diámetro que depende del ángulo de divergencia y de la distancia del emisor al objetivo. Los valores de este parámetro dependen del sistema utilizado y varía entre 0,1 y 1 miliradian (mrad). Los valores más frecuentes están entre 0,3 y 0,6 mrad.

Diámetro de la huella del haz sobre la superficie. Es directamente proporcional a la divergencia del haz y a la altura de vuelo. El diámetro de la huella es igual a la base del cono que se forma al abrirse el haz del pulso laser (p.e. con una divergencia de 0,3 mrad y una altura de vuelo de 2000 m. el diámetro de la huella del haz laser sería de 60 cm). Nominalmente se considera el diámetro de la base del haz sobre un plano perpendicular a su eje. Este diámetro aumenta con el semiángulo de apertura (β) y la pendiente del terreno. Por otro lado, la distribución de la energía del pulso no es uniforme en toda la extensión de la huella. Decrece radialmente desde el centro al borde según una distribución bidimensional Gaussiana (Gatziolis y Andersen, 2008; Ibáñez Martínez, 2008).

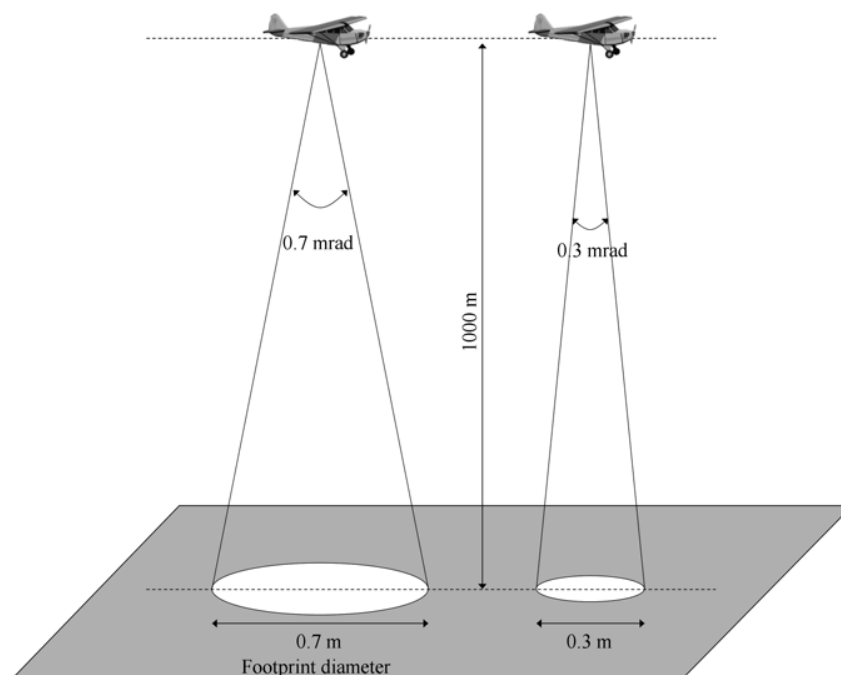


Figura 6 Huella LiDAR

Semiángulo de apertura (β). Ángulo del nadir a la máxima apertura de escaneado en la dirección perpendicular al vuelo. El error crece exponencialmente con este ángulo. El ancho de la banda escaneada depende de este ángulo y de la altura de vuelo. En los extremos de cada línea de barrido el espacio entre los centros de cada huella disminuye debido a que el dispositivo que desvía el rayo comienza a decelerar para cambiar de dirección para posteriormente recuperar la velocidad de trabajo.

Duración del pulso. El tiempo de emisión para cada pulso en nanosegundos (ns). Junto con los parámetros de discretización fija la mínima distancia entre dos retornos del mismo pulso.

Espaciamiento entre huellas. Es la distancia nominal entre los centros de las superficies cubiertas por dos haces consecutivos a lo largo y entre dos líneas de escaneado consecutivas. Junto con la divergencia del haz determina la resolución espacial de los datos LiDAR ya que fijan la superficie total cubierta por los haces. Es función de la altura de vuelo, frecuencia de pulsos y velocidad del avión.

Parámetros de discretización. Son las especificaciones del procesado de la energía rebotada para identificar los retornos individuales. Determinan la mínima energía para considerar un retorno y condicionan la distancia mínima entre dos retornos del mismo pulso (discretization tolerance).

El sensor receptor de la señal se encuentra solidario al emisor. Éste mide la potencia de retorno de la señal de modo continuo.

Cada sistema registra de manera distinta los retornos. Algunos registran sólo el último pulso, otros los dos, tres o cuatro primeros y el último, otros sólo el primero y el último, y los más modernos guardan toda la forma del retorno (sistemas de onda completa o fullwaveform).

Junto al sensor láser se coloca un sistema de navegación inercial (INS) que detecta los movimientos de alabeo, cabeceo y giro que pueda sufrir la plataforma y los compensa para evitar los errores de posicionamiento que pudieran ocasionar. Para calcular en tiempo real la posición del emisor en cada pulso el sistema se completa con un GPS. De la exactitud de los registros del GPS dependerá la precisión del trabajo posterior. Por ello, se utilizan receptores de doble frecuencia y se establece en tierra una red de aparatos similares situados en posiciones bien conocidas para utilizar sus registros en la corrección diferencial en tiempo real o en postproceso de los datos tomados por el receptor aéreo. Utilizando correcciones en tiempo real con un sistema GPS-RTK se consiguen precisiones altimétricas cercanas a los 15 centímetros. En postproceso se puede llegar a reducir el error altimétrico hasta hacerlo menor de 5 cm (Ibáñez Martínez, 2008).

4.3.2. Formato de los datos

Una vez combinados, tratados y aplicadas las necesarias correcciones a partir de los datos se obtiene una nube de puntos brutos. Cada uno de estos se corresponde con uno de los rebotes registrados y se almacena en un formato determinado.

Este formato se ajustará a las especificaciones de la ASPRS (American Society for Photogrammetry and Remote Sensing). La última fue aprobada en 2011 (Versión 1.4 de la especificación LAS). Este formato ".las" contiene datos binarios organizados en tres bloques: cabecera, registros de tamaño variable y propiamente los datos de puntos.

El **archivo cabecera** posee datos genéricos referidos al nº de puntos, nº de puntos para cada orden de rebote, coordenadas extremas del archivo, densidades medias de puntos, software, formato ".las" utilizado, proyección, metadatos y forma de onda.

El **bloque de registro de tamaño variable** contiene información del sistema de referencia, proyección y metadatos.

Los **campos de información de cada punto** son:

X, Y, y Z: Valores de las coordenadas almacenados como enteros largos. Mediante los datos de escala de cabecera estos datos se pueden convertir a valor real en coma flotante.

Intensidad: Entero que mide la potencia del pulso de retorno. Valor opcional y específico del sistema de adquisición por lo que su empleabilidad para establecer parámetros forestales es escasa ya que no son comparables entre vuelos o incluso entre sesiones distintas de un mismo vuelo o conjunto de datos.

Número del Retorno: El número del retorno es el orden que ocupa el dato en la secuencia de respuestas respecto a un único pulso de salida. Un pulso láser de salida puede tener muchos retornos y deben ser marcados por orden de llegada. El primer retorno tendrá el número 1, el segundo el dos y así sucesivamente hasta el máximo de retornos por pulso admitido por el sistema.

Número Total de Retornos (para un pulso emitido): El número máximo de retornos para un pulso dado. Los sistemas antiguos sólo admitían un rebote por pulso. Posteriormente aparecieron sistemas que distinguían el primer y último rebote para cada pulso. En la actualidad todos los sistemas discriminan múltiples retornos para cada pulso e incluso hay sensores que graban la información de la onda reflejada completa (fullwaveform).

Ángulo de escaneado: El ángulo que forma el eje del rayo emitido con el nadir del punto de emisión. Se codifica como positivo si es a la derecha del sentido de vuelo y como negativo si a la izquierda. Con ángulos de valor absoluto mayor de 30 se pueden cometer graves errores de localización.

Clasificación: Es un campo obligatorio que puede tener una serie de valores estándares propuestos por la ASPRS (tabla 5).

Tabla 5 Valores de la clasificación de puntos LiDAR

Valor	Descripción
0	Creados, nunca clasificados
1	No asignado
2	Suelo
3	Vegetación baja
4	Vegetación media
5	Vegetación alta
6	Edificios
7	Puntos bajos (ruido)
8	Clave de modelo
9	Agua
10	Reservado para ASPRS
11	Reservado para ASPRS
12	Puntos solapados
13-31	Reservado para ASPRS

Hora GPS: Se trata de la hora en formato punto flotante de doble precisión en la que se emitió el pulso y por tanto todos los rebotes del mismo pulso tienen este dato igual. Puede ser usado como identificador único de cada pulso.

Extremo de línea de barrido. Si el punto es el último de una línea de barrido antes del cambio de dirección del

dispositivo que desvía lateralmente el rayo emitido, este campo toma el valor 1. En caso contrario valdrá 0.

RGB. Cuando la captura de datos LiDAR se hace junto con la toma de fotografías aéreas, a cada punto se le puede asignar el valor de rojo, verde y azul del punto de la fotografía que coincide con su localización X-Y. De esta forma se pueden visualizar los puntos LiDAR como una nube 3D con colores reales.

4.3.3. Productos derivados de datos LiDAR brutos

El tratamiento de los datos LiDAR brutos tiene como objetivo la consecución de una gama diversa de productos de los cuales los más comunes son:

Nube de puntos brutos: Contiene la totalidad de los puntos obtenidos en los procedimientos descritos. Es habitual que existan valores anómalos que se pueden producir por irregularidades o desajustes en el proceso de adquisición de datos, por rebotes anómalos etc... Esto convierte en necesario un filtrado previo de los datos.

El filtrado de la nube de puntos requiere mucha intervención manual y por tanto puede resultar lento y caro si la calidad de los datos originales no es buena. Una adecuada planificación y ejecución en los procedimientos de toma de datos puede ahorrar mucho tiempo y dinero si se evitan la mayor parte de los errores que después habría que corregir en esta fase.

El usuario debe exigir al proveedor de datos los archivos de puntos ya depurados, puesto que es éste el que debe disponer de software adecuado para la realización de esta tarea de forma automática o semiautomática. Si no fuera así, el usuario final tendrá que realizar el filtrado y eliminación de los puntos erróneos de manera casi manual.

Normalmente la nube de puntos recibida por el usuario ha pasado por el proceso de clasificación por lo que cada punto tendrá asignado el valor de alguno de los grupos reseñados en la tabla 5.

La principal dificultad para el manejo de estos productos es su gran volumen. Dependiendo de la densidad de puntos por metros cuadrado, el tamaño de un archivo puede llegar a ser inmanejable. P.ej. para una densidad media de cuatro puntos por metro cuadrado se necesita aproximadamente un Megabyte por hectárea y, por tanto, superficies de 1.000 hectáreas darían lugar a ficheros de más de un Gigabyte.

Modelos Digitales del Terreno (MDT): Como se ha comentado arriba, una de las principales ventajas del LiDAR aerotransportado es que una parte considerable de los pulsos atraviesa la vegetación total o parcialmente llegando hasta el suelo. El número de rebotes que se reflejan en el suelo es variable según la fracción de cabida cubierta y la densidad de la vegetación. Esto permite seleccionar entre los rebotes aquellos que se entiende han llegado hasta el suelo y, por tanto, contienen información sobre las coordenadas y la cota de puntos del terreno con una precisión, de hasta 5 cm.

Los algoritmos de filtrado de puntos son variados y cada software utiliza su propio sistema para determinar qué puntos pertenecen al suelo. Normalmente en los procesos de filtrado se puede elegir el nivel de exigencia que se le impone a un punto para ser considerado como punto de suelo. Al aumentar la seguridad en el filtrado disminuye el número de puntos de suelo de forma que se puede llegar a carecer de datos en superficies más o menos extensas. En estos huecos la superficie es interpolada con la información de los puntos que los rodean. Esta carencia es más acusada cuanto más densa es la vegetación que cubre el terreno y menor la densidad de pulsos.

Se han publicado multitud de algoritmos de filtrado que, en general, se pueden clasificar en cuatro grupos:

Filtros basados en el cálculo TIN. Asumen que la superficie topográfica no presenta grandes discontinuidades nada más que ocasionalmente. Localizan los puntos mínimos considerados localmente y realizan una triangulación inicial que sirve como base para el cálculo del posterior modelo digital (Axelsson, 2000).

Filtros basados en interpolación. Comparan las elevaciones de los puntos y las estimadas usando varios métodos de interpolación. (Kraus y Pfeifer, 1998).

Filtros morfológicos. Se basan en la comparación y posición relativa de formas geométricas generadas en cada punto respecto a los circundantes (Vosselman, 2000).

Filtros basados en contornos activos. Utilizan modelos de formas activas para representar un contorno tridimensional que sirve de red para eliminar puntos no pertenecientes al suelo.

Se entiende por Modelo Digital del Terreno el modelo que se construye eliminando las alturas de la vegetación y las infraestructuras humanas, utilizando una malla que une los puntos clasificados como terreno y también la restitución de las líneas de ruptura.

Como se puede apreciar en la figura X, una vez eliminados los puntos no clasificados como terreno aparecen huecos en los que no hay información sobre la cota. En terrenos forestales estos huecos son más frecuentes cuanto más densa es la vegetación. Algunos procesos de generación de MDTs rellenan estos huecos con puntos de cota similar a la de los puntos que los rodean. En otros procesos se implementan algoritmos que asignan la cota de cada punto creado según la tendencia del terreno. Cuando no se rellenan con puntos los huecos se realiza una triangulación en ellos de forma que en el modelo final aparecen triángulos mucho mayores que los del resto de la malla.

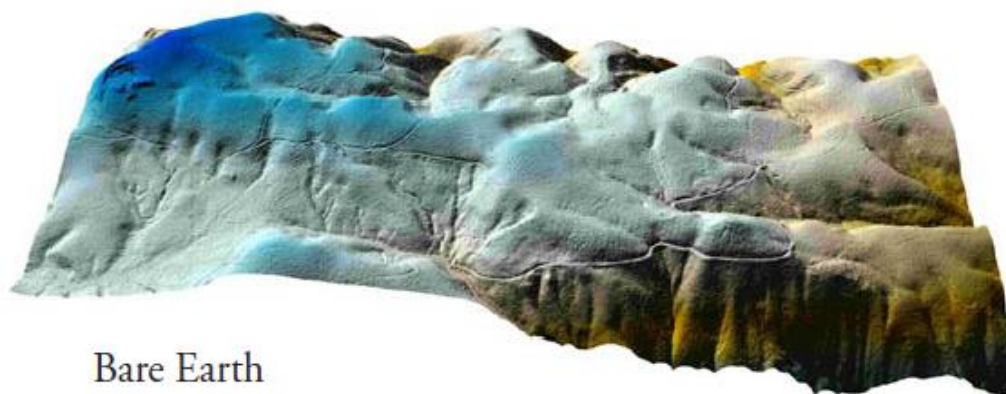


Figura 7 Modelo Digital del Terreno (Matti Maltamo et al., 2014)

El MDT es el principal producto derivado de los datos LiDAR y de su precisión dependerá en gran medida la fiabilidad de la información sobre parámetros forestales deducibles de la nube de puntos ya que se constituye en la referencia sobre la que se calcula la altura de cada punto de la nube. Todas las variantes de utilización de datos LiDAR para la gestión forestal utilizan la información contenida en los MDT que se deben construir previamente para poder referenciar la altura de cada punto y consecuentemente los errores cometidos en la confección del MDT se transmitirán a la altura de la vegetación localizada sobre cada celda de la malla del modelo (Kraus y Pfeifer, 1998; Vosselman, 2000).

Modelo Digital de Superficie (MDS) Se entiende por modelo digital de superficies el que se construye teniendo en cuenta las alturas de la vegetación y las infraestructuras humanas, utilizando una malla de elevaciones y la restitución de las líneas de ruptura. Para la confección

de los MDS sólo se tienen en cuenta los primeros rebotes de cada pulso.

Gráficamente se podría visualizar como un gran paño fino que se dejara caer sobre la superficie que forma el suelo y todos los elementos que hay sobre él.

Al rasterizar los puntos de primer rebote aparece un modelo que representa la superficie que contiene los puntos de máxima cota o cota media en cada celda. Esto ocasiona una superficie muy accidentada con muchos máximos y mínimos locales que no aportan información y pueden distorsionar el modelo.

Para evitarlo se usan filtros que suavizan la superficie eliminando gran parte de los extremos locales. En el proceso de filtrado se define una ventana móvil de $n \times m$ píxeles que recorre toda la superficie de trabajo. A cada píxel se le asigna el valor de la media o mediana de los de la ventana en la que se encuentra. De esta forma desaparecen los máximos y mínimos locales que distorsionan la forma de la superficie. Como contrapartida se incrementa la incertidumbre en la determinación de la altura de cada punto.

Es importante señalar que en el raster generado el valor de cada celda refleja la cota absoluta de cada punto y no su altura sobre el terreno (Kraus y Pfeifer, 1998).

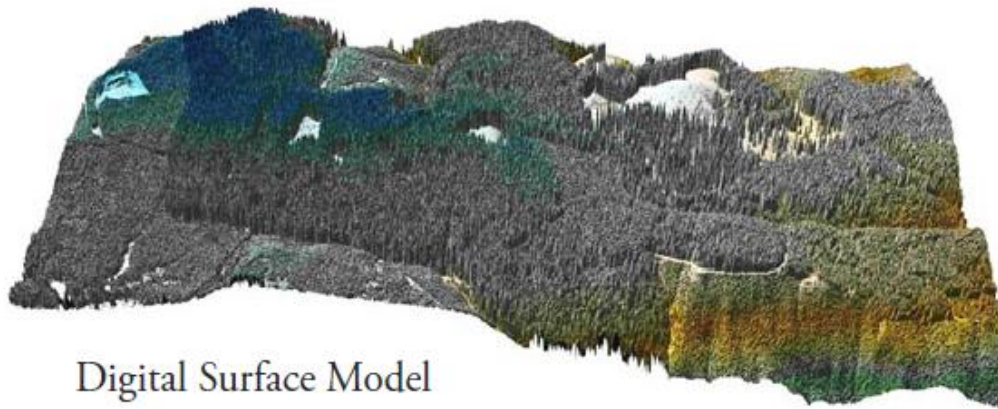


Figura 8 Modelo Digital de Superficies (Matti Maltamo et al., 2014)

Imágenes de intensidades: La intensidad con la que el rebote regresa al sensor es variable según el objeto contra el que chocó. Esta propiedad permite utilizar la intensidad para distinguir los elementos con distintas reflectancias que están sobre la superficie terrestre como carreteras, vegetación, cultivos, etc. Sin embargo, la intensidad con la que se recibe el rebote no es directamente relacionable con el material u objeto sobre el que se produjo ya que varía con las condiciones de vuelo y con el sensor utilizado. Aun así, en muchos casos discrimina distintas situaciones por comparación dentro de la misma imagen. Para ello es necesario normalizar los valores, eliminar los extremos y aplicar una escala de color en blanco y

negro para asignar un nivel de intensidad en esa escala. La imagen resultante se regulariza al rasterizarla al ser constante la distancia entre píxeles.

4.3.4. Formas de onda completa y sistemas de retorno discreto

Una forma de onda es la señal que se devuelve al sensor LiDAR después de haber sido dispersada desde las superficies que el pulso láser intercepta. Los sistemas LiDAR de forma de onda completa registran la señal de vuelta completa dentro de un rango de elevación por encima de un umbral de ruido de la energía de fondo. Un ejemplo de este tipo es el Geoscience Laser Altimeter System de la NASA (GLAS) (Figura 9)). La escena muestra un dosel de dos pisos de abeto Douglas (*Pseudotsuga menziesii*) en una pendiente suave de $4,9^\circ$. En superficies con vegetación en terreno relativamente plano, se produce una forma de onda típicamente bimodal (Rosette et al., 2012).

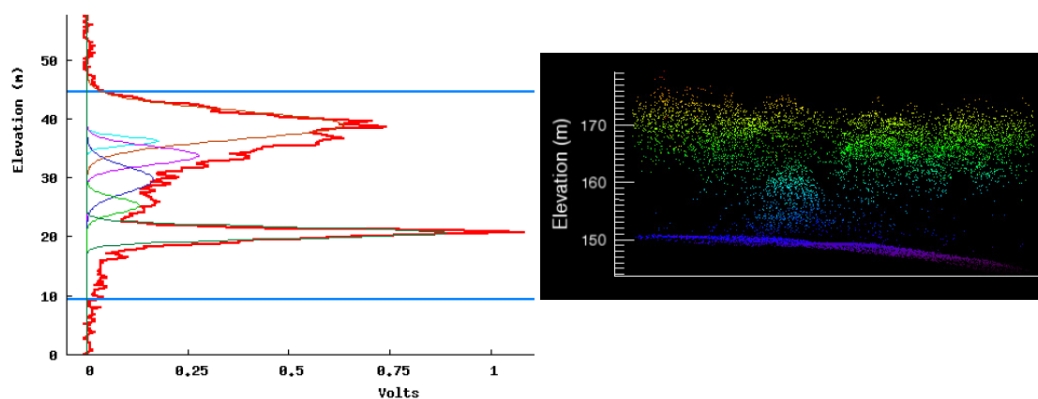


Figura 9 Forma de onda completa

El principio y el final de la señal de forma de onda por encima del umbral de ruido de fondo están representados respectivamente por las líneas azules horizontales superior e inferior (ruido medio + 4.5σ en el caso de GLAS). La amplitud de la forma de onda (eje x) representa tanto el área superficial interceptada en cada elevación como la reflectividad de las superficies a la longitud de onda emitida (1064 nm).

El gradiente en el comienzo de la señal aumenta lentamente inicialmente debido a la superficie relativamente pequeña del follaje y las ramas del dosel superior. A medida que la energía penetra hacia abajo a través del dosel, la amplitud de la forma de onda aumenta a medida que se interceptan más elementos del arbolado, antes de disminuir hacia la base de las copas de los árboles. Se puede observar un segundo pico, que corresponde a un árbol más corto, por encima del pico abrupto y estrecho, que es el devuelto por el suelo. Por debajo de la superficie del suelo, la señal se puede ver que desciende gradualmente. Esto se relaciona tanto con la suave pendiente de este sitio más el efecto de dispersión múltiple del resto de elementos de la escena, que retrasan parte de la señal devuelta al sensor.

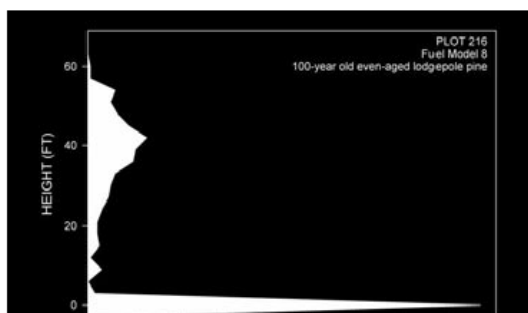


Figura 10 Respuesta de suelo y vegetación sin sustrato arbustivo

Debido a la compleja señal de onda que se produce, a menudo se simplifica utilizando la descomposición gaussiana de la forma de onda (Figura X, izquierda). Representando la forma de onda como la suma de los gaussianos, se suaviza la señal, pero permite un medio de retención e identificación

de los elementos dominantes de la señal para facilitar la interpretación.

Los sistemas LiDAR de huella estrecha pueden producir un muestreo denso de la superficie objetivo. La señal devuelta también está en forma de onda, sin embargo, con sistemas de retorno discretos, sólo se registran los ecos designados dentro de la forma de onda. Éstas pueden ser las primeras y últimas devueltas, o a veces, también una serie de puntos intermedios. La agregación de estos retornos de pulsos de LiDAR múltiples emitidos permite que la escena sea reconstruida como una "nube de puntos" de superficies interceptadas y geolocalizadas. Esto se ve en la Figura X, a la derecha, que ilustra la misma ubicación que se ve dentro de la forma de onda GLAS. La nube de puntos LiDAR de pequeña huella puede interpretarse de forma más intuitiva como un piso superior dominante de altura aproximadamente uniforme y un solo árbol de menor altura en el centro de la escena. Los puntos se colorean con respecto a su elevación.

4.3.5. Patrones de distribución de la huella LiDAR

Los patrones de distribución de las huellas difieren entre los sistemas LiDAR. El perfilado LiDAR implica la localización sistemática de huellas a intervalos a lo largo de la trayectoria del sensor en el suelo. Estos pueden ser contiguos, como el Sistema Portátil de Láser Aerotransportado (Nelson, Parker y Hom, 2003), o colocados a distancias discontinuas a lo largo de la órbita terrestre en el caso del Geoscience Laser Altimeter System (Schutz, Zwally, Shuman, Hancock y DiMarzio, 2005) (Figura X). Esto

generalmente permite el muestreo de áreas extensas, sin embargo, requiere un medio para extrapolar estimaciones de parámetros biofísicos para las áreas donde se adquirieron los datos.



Figura 11 Huella de LiDAR satelital (Matti Maltamo et al., 2014)

El escaneado láser obtenido a partir de una plataforma aerotransportada, se produce perpendicularmente a la dirección de desplazamiento y, en general, produce una densa distribución de pequeñas huellas. La anchura de la hilera y la densidad de huella dependen de la altitud y velocidad de la aeronave más el ángulo de exploración aplicado. Un espejo de exploración dirige pulsos de láser hacia adelante y hacia atrás a través de la línea de vuelo,

causando que los datos sean capturados típicamente en una disposición de diente de sierra.

El ángulo de exploración máximo del nadir para el instrumento se puede personalizar de acuerdo a las necesidades de cada campaña. Los ángulos de exploración más estrechos mejoran las posibilidades de que cada pulso penetre densamente en los doseles de vegetación y de que el sensor reciba un pulso devuelto desde el suelo, ya que hay mayor probabilidad de un camino expedito a través del dosel hasta el suelo. La práctica habitual es crear una superposición de líneas de vuelo similar a las pasadas fotogramétricas de aproximadamente un 60%. Se pueden combinar múltiples líneas de vuelo para proporcionar una cobertura completa del área deseada mediante software especializado. El escaneo láser de huella estrecha es generalmente empleado a escala de estudios forestal, en gran parte debido a los costos.

4.3.6. Adquisición de datos LiDAR

El proceso de adquisición de datos LiDAR requiere de una planificación que dependerá de las necesidades del proyecto. Por ejemplo, el planteamiento variará si el objetivo es generar modelos digitales del terreno o si lo que se busca es extraer información sobre la vegetación.

Cuando el objetivo principal es obtener información útil en la gestión forestal se deberían tener en cuenta los siguientes condicionantes:

Época de realización del vuelo. En las masas forestales de especies caducifolias los vuelos realizados en invierno no tendrán apenas información de los estratos de copas de los árboles. En cambio, la proporción de puntos de suelo será mucho mayor y por tanto los modelos digitales del terreno (MDT) que se generen con estos datos serán mucho más precisos. Esta limitación no existe en masas de frondosas perennifolias y resinosas.

Densidad de pulsos. La utilización de datos LiDAR para obtención de parámetros forestales se puede hacer desde dos orientaciones metodológicas distintas: utilizando variables de masa o utilizando variables de árbol individual.

Esta segunda opción requiere la individualización de los pies dentro de la masa y para ello requiere una mayor densidad de rebotes por metro cuadrado que la que emplea variables de masa. En general se admite que la densidad mínima para individualizar árboles es de 4 rebotes por metro cuadrado (Gatziolis y Andersen, 2008). Si se utilizan las variables de masa, densidades de 1 punto por metro cuadrado o incluso menores son suficientes.

Frecuencia de emisión de pulsos. A igualdad de puntos por metro cuadrado, los datos tomados con equipos de altas frecuencias tendrán un mayor porcentaje de rebotes sobre la superficie que los de frecuencias más bajas. La energía que puede concentrar un pulso emitido por un equipo de alta frecuencia es menor que la que concentra un pulso emitido por ese mismo equipo con una frecuencia más baja

y consecuentemente tendrá menor capacidad de penetración en el dosel. En consecuencia, si se quieren utilizar modelos digitales de la superficie de la masa forestal se deberán usar equipos de más alta frecuencia. Para usar información de la distribución de los puntos en toda la anchura de las copas es ventajoso usar equipos de frecuencias más bajas que además generaran más puntos de suelo para mejorar la precisión de los MDT.

Ajuste de los parámetros de discretización. Consiste en ajustar el valor mínimo de diferencia de energía en los máximos de la onda rebotada para considerarlo como un eco diferenciado. Cuanto más pequeño es este valor, más rebotes por pulso aparecen, pero tienden a concentrarse en el ápice de los árboles y sobre todo en los dominantes (Gatziolis y Andersen, 2008). Sin embargo, un ajuste fino de este valor se asocia con pérdida de precisión en la posición de los puntos de los estratos inferiores de vegetación. Cuando el objetivo es el estudio del estrato de árboles dominantes conviene reducir este valor.

Semiángulo de apertura. Cuanto mayor es el ángulo de divergencia del rayo emitido, mayor es la superficie interceptada por la sección del cono que forma. Esto supone que se mejore la robustez y disminuya la varianza cuando se utilizan métodos de variables de masa (Gatziolis y Andersen, 2008), a cambio se asumirá una menor precisión en la determinación de la posición de cada uno de los puntos. Así pues, cuando se usan métodos de individualización de árboles es preferible reducir el ángulo de apertura. En cualquier caso, el semiángulo de

apertura nunca debe ser mayor de 30° . Con ángulos mayores aparecen problemas de "multipathing" o cambios en la dirección de propagación del rayo que disminuyen la precisión en la localización de los puntos reflejados. El USDA Forest Service recomienda que el semiángulo no sea mayor de 12° en terrenos llanos y 10° en terrenos escabrosos (Gatziolis y Andersen, 2008).

4.3.6.1. Planeamiento del vuelo LiDAR

Los parámetros y condiciones del vuelo para la toma de datos dependerán de los condicionantes mencionados.

Para una correcta planificación del vuelo se tendrán en cuenta fundamentalmente dos aspectos: la superficie a cubrir y la densidad de puntos deseada. Se planificarán las líneas de vuelo considerando, en primer lugar, el ancho de cada pasada y en segundo los solapes entre estas según altura de vuelo y el ángulo de apertura.

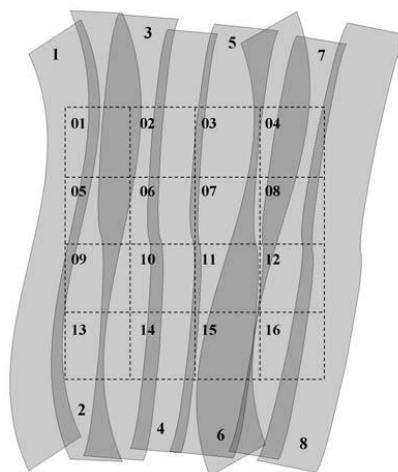


Figura 12 Líneas de vuelo LiDAR aerotrasportado (Matti Maltamo et al., 2014)

La densidad de pulsos se calcula a partir de dos parámetros: la frecuencia de emisión de pulsos y la velocidad de vuelo. A veces será necesario planificar pasadas secundarias para asegurar la densidad mínima pretendida. Otro aspecto importante a tener en cuenta es la climatología de la región de vuelo; se evitarán las épocas de lluvias o de nieblas (Gatziolis y Andersen, 2008).

Para el ajuste de datos post vuelo se seleccionarán puntos de control sobre el terreno coincidentes con las estaciones GPS de apoyo de corrección diferencial.

Una vez diseñado el vuelo, se ubicarán y ajustarán las estaciones GPS de apoyo en tierra y los aparatos (sensor, GPS e IMU) del avión y se realiza el vuelo.

Cada pasada del avión se denomina "misión" para la cual se calibrarán y ajustarán los sensores LiDAR y GPS de la aeronave.

Los datos brutos obtenidos posteriormente serán sometidos a una serie de procesos y filtros previos a su conversión al formato final de usuario. El software empleado para estos procesos es muy específico, siendo normal que sea la misma empresa que ejecutó el vuelo la que se ocupe de la elaboración de los datos para su comercialización.

La preparación de datos debe incluir:

- Ajuste de posición de los puntos que asegure cobertura completa y densidad mínima.

- Homogeneización de la densidad eliminando solapes.
- Filtrado y clasificación de los puntos en las categorías señaladas.

Esta fase de adquisición de datos escapa normalmente al control del usuario final. Sin embargo, es muy necesario conocer estas rutinas ya que marcarán muchos de los procedimientos posteriores y las posibilidades de aplicación de los datos.

4.4. El LiDAR en los estudios forestales

La investigación sobre el uso de datos LiDAR para aplicaciones forestales comenzó a mediados de la década de 1970 en experimentos que usaban simples instrumentos de generación de perfiles. Desde aproximadamente 1990, el acento se ha puesto en el desarrollo de herramientas para realizar inventarios forestales que explotan la capacidad de los instrumentos del Escáner Láser Aerotransportado (ELA) capaces de proporcionar cobertura de aéreas completas.

El LiDAR se ha incorporado como un componente esencial de los inventarios forestales en varios países (Maltamo et al., 2011; Næsset, 2007; Woods et al., 2011). Además, hoy en día son muchos los países que están adquiriendo datos LiDAR para construir modelos de elevación del terreno detallado que luego sirven como un nuevo origen de datos para realizar inventarios y cartografía de la vegetación (Nord-Larsen y Riis-Nielsen, 2010). Hoy en día, pueden obtenerse datos LiDAR para aplicaciones en una variedad de campos de gran número de proveedores y empresas de topografía, cada una con múltiples sistemas.

Por lo tanto, se puede decir que el LiDAR está reconocido como una técnica madura y bien establecida con transcendencia científica y práctica. En los últimos años se han publicado numerosos artículos de revisión científica (por ejemplo, (Hyypä et al., 2008; Koch et al., 2008; Lim, Treitz, Wulder, St-Onge y Flood, 2003; Maltamo et al., 2007; McRoberts, Cohen, Naesset, Stehman

y Tomppo, 2010; Næsset, 2004) y capítulos de libros (Hyypä et al., 2009; Koch et al., 2008; Packalén, Maltamo y Tokola, 2008) que tratan sobre las aplicaciones forestales de esta disciplina.

Los primeros estudios sobre modelado de parámetros forestales mediante datos LiDAR continuaron investigaciones anteriores usando datos de LiDAR de perfilado aéreo que, a su vez, continuaron incluso anteriores investigaciones que empleaban la fotografía aérea. (Nelson, 2013) apunta que la primera mención de posibles aplicaciones forestales fue realizada por (Rempel RC, 1964) los cuales emplearon un sistema de perfiles LiDAR para obtener alturas de planta y árbol. Los primeros experimentos para medir árboles fueron realizados por Solodukhin et al., (1976), que comparó perfiles láser de árboles talados con mediciones de cinta, para posteriormente montar el mismo instrumento en un avión y adquirir perfiles aerotransportados (Solodukhin et al. 1979). Más tarde se llevaron a cabo estudios similares utilizando perfiladores láser en América del norte (Aldred, Bonnor y others, 1985; Maclean y Krabill, 1986; Nelson, Krabill y MacLean, 1984), y algunos estudios desarrollaron métodos para estimar el volumen de bosque y la biomasa a partir de perfiles de cubierta aerotransportada (Maclean y Krabill, 1986; Maclean y Martin, 1984).

(Arp, Griesbach y Burns, 1982) también utiliza un sistema de generación de perfiles, junto con la fotografía aérea, para cartografiar la altura del bosque en Venezuela. Los

primeros usos de los datos LiDAR para inventarios forestales eran extensiones de aplicaciones similares a partir de datos fotogramétricos. (Maclean y Martin, 1984) reportaron R^2 de 0,75 0,87 para un modelo de regresión de la relación entre el volumen de madera y áreas transversales de perfiles de cabida cubierta obtenidas de fotografías aéreas.

Los primeros estudios de LiDAR imitaron este enfoque utilizando los datos de perfiles de sistemas LiDAR adquiridos de un estrecho transecto a lo largo de la línea de vuelo de avión. (Aldred et al., 1985) utilizan datos de Perfil de LiDAR para estimar alturas de copa dentro de los 4,1 m alturas reales y asignar grados a las clases de densidad de cubierta con un 62% de precisión. (Maclean y Krabill, 1986) obtienen R^2 0,9 para un modelo de regresión de la relación entre el volumen comercial bruto y áreas transversales derivadas de perfiles láser. (Nelson et al., 1984) utiliza los datos de un sistema de generación de perfiles para caracterizar perfiles verticales de dosel. (Schreier, Lougheed, Tucker y Leckie, 1985) demostraron la utilidad de los perfiles de datos LiDAR para distinguir entre coníferas, caducifolios y vegetación de bajo crecimiento. (Nelson, Krabill y Tonelli, 1988) elaboraron modelos lineales de las relaciones entre altura, volumen por unidad de área y biomasa por unidad de área como variables de respuesta y métricas de perfil de altura derivados de LiDAR como variables predictoras con R^2 de 0,50 y 0,60.

La importante experiencia obtenida a partir de estos primeros estudios de perfiles LiDAR permitió incorporar modelos de regresión utilizando mediciones LiDAR para predecir atributos forestales de interés con fines de inventario forestal. Los sucesores de los perfiladores LiDAR fueron los sistemas LiDAR de huella estrecha en el rango 0,1-2,0 m y diámetros de huella ancha en el rango de 5,0-10,0 m.

Un paso importante hacia el desarrollo comercial de LiDAR fue la integración de sistemas de posicionamiento Global (GPS) y los sistemas de navegación inercial (INS) en la década de 1990, lo cual facilitó la colocación exacta del escáner y, en consecuencia, de los datos registrados. Utilizando algunos de los sistemas primero comerciales o precomerciales equipados con GPS e INS (Næsset, 2004), varios investigadores reportaron relaciones relevantes entre mediciones de campo de altura y LiDAR, especialmente los datos de altura media. Además, se estimaron valores medios o valores totales de las variables forestales biofísicas como el área basal (Means et al., 2000) y el volumen de los pies (Means et al., 2000; Naesset, 1997) a partir de datos LiDAR de huella estrecha con exactitud similar o mayor que los producidos por los métodos puros basados en muestras de campo.

Los primeros estudios se centraron en las relaciones entre los atributos forestales a nivel de parcela basados en mediciones de campo y las distribuciones de altura basadas en datos LiDAR. Estos enfoques se denominarán basados en área para diferenciarlos de los enfoques de árboles

individuales propuestos unos años más tarde por (Brandtberg, 1999) que se centrarán en la detección de árboles individuales a partir de modelos de superficie construidos usando datos LiDAR. Ambos enfoques de han sido ampliamente probados para inventarios forestales basados en LiDAR. También se han desarrollado variantes y métodos híbridos.

4.4.1. LiDAR Aerotransportado y estimación de parámetros forestales a escalas regional y local

Como ya se ha comentado, el uso de datos de exploración láser aerotransportados en aplicaciones forestales ha alcanzado un interés creciente durante las últimas décadas. En la actualidad, el LiDAR es percibido como un sistema de evaluación rentable y preciso de la estructura vertical y horizontal de las áreas forestales y, por lo tanto, una alternativa válida o complementaria a los métodos actuales de inventario (Wulder et al., 2009). La localización vertical de los puntos se realiza habitualmente con precisiones cercanas a los 15cm, mientras que la exactitud de geolocalización horizontal se encuentra en el rango de los 20-30cm.

Los sistemas de LiDAR aerotransportados proporcionan una cobertura relativamente densa con un tamaño de huella en rango de decenas de centímetros. Sin embargo, a menudo se considera que subestiman la altura del dosel como resultado de la distribución y la densidad de puntos. Esto se debe al cierto grado de penetración de la señal que varía según las especies (Næsset, 2004). En general, esta

subestimación es menos pronunciada para árboles en forma de cono como las piceas o el abeto Douglas (*Pseudotsuga menziesii*) que para árboles de forma esférica como muchas planifolias o incluso el pino silvestre (*P. sylvestris* L.). Las coníferas normalmente crean formas más compactas con menos penetración de energía a través del dosel que las planifolias. Por lo tanto, los retornos de energía tienden a ser producidos a partir de las capas externas del dosel. Sin embargo, el grado de penetración está en última instancia relacionado con una combinación de factores tales como la densidad de muestreo, la divergencia del haz y el ángulo de escaneo (Suárez, Ontiveros, Smith y Snape, 2005).

Los posibles escenarios se ilustran en la Figura X . Con una mayor densidad de pulsos láser, estas dificultades se reducen, pero a costa de mayores gastos operativos y restricciones de duración de vuelo que limitan la cobertura espacial.

- a. El láser golpea la parte superior del dosel
- b. Se ignoran los árboles pequeños cercanos a los vecinos más grandes,
- c. La situación más probable: los retornos del láser no golpean el verdadero ápice del árbol,
- d. Uno de los pulsos es interceptado a una altura inferior debido a la penetración del dosel sugiriendo erróneamente dos copas de los árboles,

- e. Los árboles en un montículo se pueden asignar una mayor altura en ausencia de un buen modelo de la superficie del terreno
- f. En una situación de escasa densidad algunos árboles pueden ser ignorados completamente

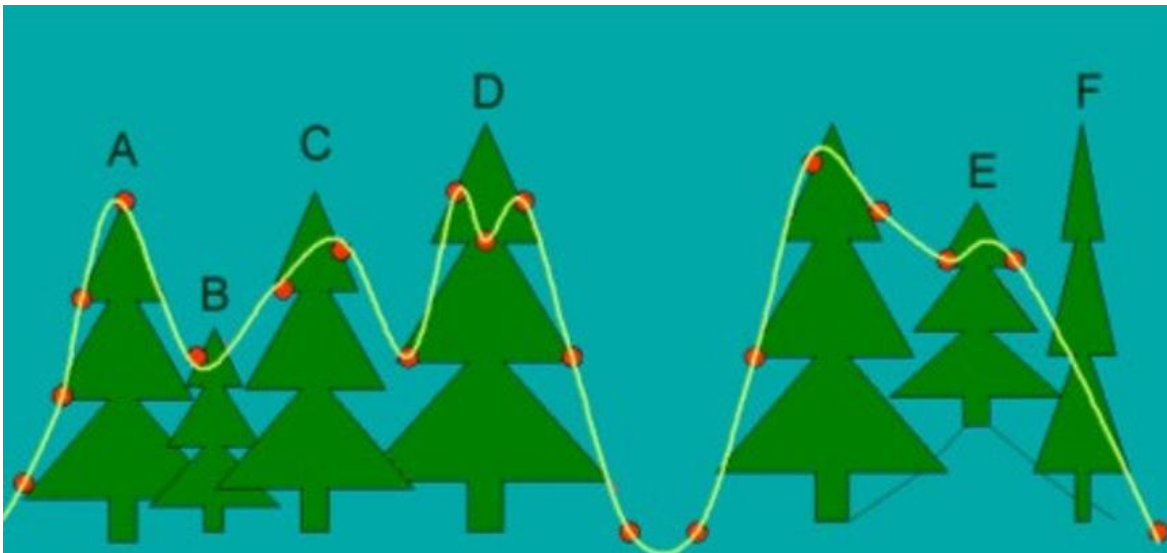


Figura 13 Posibles escenarios de interceptación de la nube de puntos LiDAR del dosel forestal. Fuente: (Suárez et al., 2005)

A pesar de la percepción común, el LiDAR no crea imágenes tomográficas y, por lo tanto, no pueden considerarse representaciones tridimensionales en sentido estricto. Sólo los huecos en la cubierta del dosel y la transmisividad a través de las hojas permiten que la energía del láser sea devuelta desde el suelo. Al igual que con los datos de forma de onda completa, el punto crítico para el cálculo de las métricas de alturas de la vegetación es distinguir entre los puntos devueltos por las superficies del suelo y la que no lo son.

Puesto que la penetración de energía del LiDAR a través de la cubierta de vegetación varía con la estructura del bosque, la densidad y el ángulo de exploración láser, el último retorno de un pulso de LiDAR emitido puede no necesariamente ser devuelto desde la superficie del suelo. Por lo tanto, es necesario un medio para filtrar puntos y diferenciar los que llegan al suelo de aquellos que son interceptados a diferentes alturas dentro del dosel (Kraus y Pfeifer, 1998; Zhang et al., 2003). En (Sithole y Vosselman, 2004) se puede encontrar una completa comparación de diferentes enfoques y una descripción completa de los filtros. Esto permite clasificar los puntos en clases de suelo y clases de vegetación (Figura X).

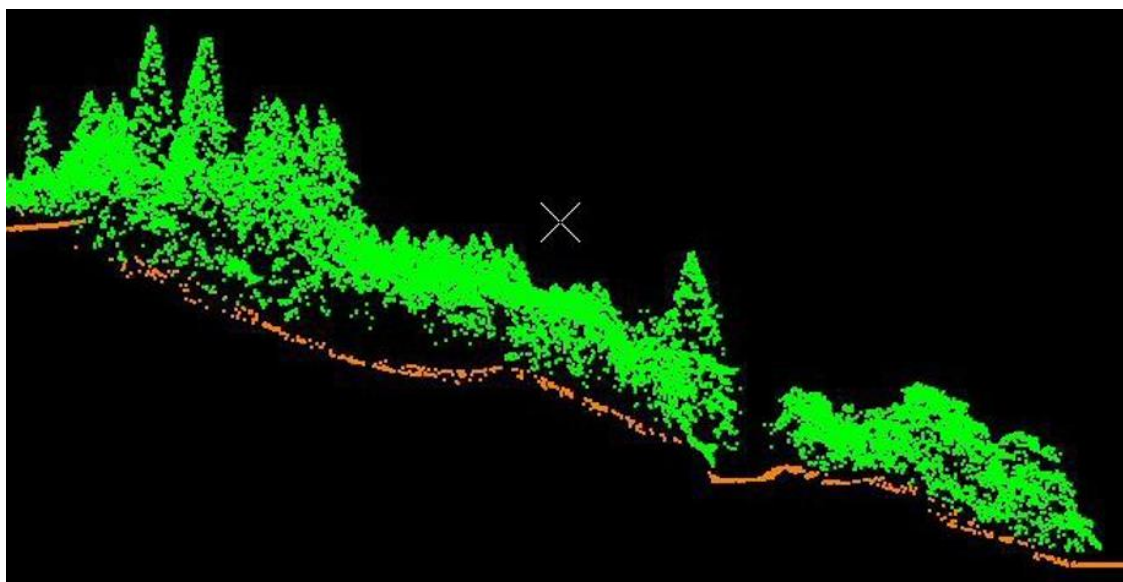


Figura 14 Distribución y estructura de la vegetación mostrada por clasificación de una nube de puntos LiDAR aerotransportado.

Otros algoritmos pueden refinar la clasificación para identificar adicionalmente características como edificios, cables eléctricos, etc.

4.4.2. Enfoques basados en área

Un concepto fundamental subyacente al enfoque basado en el área es que la nube de puntos 3D que consiste en las alturas correspondientes a todos los ecos para una superficie dada, por ejemplo, una parcela de muestra, contiene información que puede usarse para caracterizar la superficie del suelo y la distribución vertical del material biológico en capas de vegetación por encima de esa área.

Para la mayoría de las aplicaciones, la distribución de alturas verticales para un área particular se ignora la posición horizontal de ecos individuales dentro del área en cuestión, pero mantiene las alturas para todos los ecos. La distribución entera de los datos LiDAR, o la porción de la misma que representa la vegetación, puede usarse para calcular métricas discretas que están relacionadas con las propiedades de la distribución del material biológico. Las métricas incluyen parámetros de la distribución de altura LiDAR tales como la media o percentiles y parámetros relacionados con la densidad de la copa, tales como las tasas de penetración. La distribución de alturas LiDAR también puede ser estimada mediante funciones de distribución continua.

(Magnussen y Boudewyn, 1998) demostraron que la proporción de alturas de retornos LiDAR de huella estrecha en o por encima de una altura de referencia dada correspondía bien con la proporción de área foliar por encima de esta altura. Aunque el área foliar no es generalmente una variable de

inventario de interés, una contribución derivada de este estudio fue la demostración de la utilidad de los percentiles de altura LiDAR como variables predictoras. Este enfoque innovador se ha demostrado posteriormente útil para la predicción de una amplia gama de atributos forestales. Varias extensiones de este concepto son especialmente relevantes.

- En primer lugar, las densidades de pulsos LiDAR para las unidades espaciales utilizadas para la calibración del modelo deben ser suficientemente grandes para producir estimaciones fiables de los parámetros de las distribuciones de altura de los retornos utilizadas como variables predictoras del modelo.
- En segundo lugar, las densidades de retornos deben ser suficientemente grandes para estimar estos parámetros para ambas unidades espaciales utilizadas para calibrar modelos y unidades espaciales utilizadas para la estimación de superficie. Esta última característica permite la calibración de modelos utilizando datos de campo y LiDAR para parcelas forestales.

Como señalan (Magnussen y Boudewyn, 1998), el número de pulsos por unidad espacial debe producir estimaciones fiables de las métricas LiDAR utilizadas para predecir los atributos del bosque. En este sentido, las densidades de pulso mínimas rara vez han sido inferiores a 0,1 pulsos / m² (Holmgren y Jonsson, 2004; Naesset, 1997), y las áreas de parcela mínima rara vez han sido menores de 200 m² (Anderson y Breidenbach, 2007; Breidenbach, Gläser y

Schmidt, 2008; Gobakken y Næsset, 2008; Næsset, 2002). Aunque se han publicado varios informes sobre los efectos de diferentes densidades de pulso en las estimaciones de volumen y biomasa basadas en LiDAR, los resultados no son directamente comparables debido a que no se usaron áreas comunes de parcela.

De manera similar, los informes de los efectos de tamaños de parcela variables no son directamente comparables porque no se usaron densidades de pulso comunes. Sin embargo, cuando los resultados de estos estudios comparativos se expresan en términos de pulsos por parcela, (Breidenbach et al., 2008; Gobakken y Næsset, 2008; Maltamo, Malinen, Packalén, Suvanto y Kangas, 2006; Thomas, Treitz, McCaughey y Morrison, 2006) concluyeron que la reducción de la densidad de pulso a 100-225 pulsos por parcela no tuvo efectos adversos sobre la calidad del ajuste de los modelos de volumen o biomasa cuando se usaron deciles de altura como variables predictoras. (Holmgren y Jonsson, 2004; Strunk, Temesgen, Andersen, Flewelling y Madsen, 2014) informaron de que la reducción del número de pulsos por parcela a 50 o menos no tuvo efectos adversos graves sobre los ajustes o estimaciones del modelo. (Thomas et al., 2006) compararon no sólo calidades de ajuste sino también distribuciones de alturas de retornos para densidades de pulso de 200 y 2000 pulsos por parcela, e informaron que las distribuciones eran esencialmente las mismas.

(Næsset, 1997) comparó las medias de las alturas de 36 rodales a partir de una muestra de 15 parcelas por rodal

con tres medias basados en métricas LiDAR: media aritmética de alturas LiDAR, altura media ponderada LiDAR y media altimétrica del rodal a partir de la mayor altura LiDAR de celdas de 15x15m. Esta última medida resultó ser la que menos se desviaba de las medias basados en las mediciones de los árboles. (Naesset, 1997) estimó el volumen de rodal a partir de múltiples modelos lineales y no lineales utilizando tres métricas LiDAR como variables independientes: altura media del rodal calculada como la media de las alturas máximas de retorno para celdas de pequeño tamaño, altura media de todos los ecos LiDAR y densidad de cubierta del dosel, calculado como la proporción de ecos de las copas. Los valores de R^2 oscilaron entre 0,45 y 0,89, dependiendo del sitio.

La contribución fundamental de estos dos estudios fue la introducción de la idea de mosaico espacial compuesto por una cuadrícula de celdas y la predicción de la variable de respuesta para cada celda de la cuadrícula.

La intuición de (Magnussen y Boudewyn, 1998) con respecto a la necesidad de un número suficiente de pulsos para caracterizar las distribuciones verticales de alturas de ecos LiDAR, en combinación con el enfoque de celdas de cuadrícula propuesto por (Næsset, 1997; Naesset, 1997), llevó a Næsset y Bjercknes a seleccionar las celdas de cuadrícula con las mismas áreas que las parcelas utilizadas para calibrar los modelos de predicción. (Næsset, 2004) proporcionó una justificación parcial para seleccionar unidades espaciales de áreas iguales, así como cierta cautela en el uso de métricas como la altura máxima que

dependen del tamaño de la parcela. Sin embargo, (Zhao, Popescu y Nelson, 2009b) propusieron modelos usando métricas LiDAR basadas en distribuciones de altura de eco que pueden ser calibradas usando unidades espaciales de un tamaño y aplicadas a unidades espaciales de diferentes tamaños.

(Magnussen, Naesset y Gobakken, 2013) propusieron el uso de momentos de las distribuciones de alturas del dosel LiDAR para predecir la distribución a nivel de parcela de atributos forestales tales como el diámetro del árbol, el área basal y el volumen. A pesar de estos avances, sigue siendo algo incierto la cuestión de si los modelos construidos utilizando parcelas con áreas particulares y densidades de pulso pueden aplicarse sin efectos adversos a las celdas de cuadrícula con diferentes áreas y densidades de pulso de una manera fácilmente implementable y eficiente.

Son numerosas las investigaciones sobre procedimientos para caracterizar las relaciones entre parámetros forestales y las métricas LiDAR. Algunos ejemplos: (Means et al., 2000) usaron modelos lineales y datos LiDAR de huella estrecha para predecir altura, área basal y volumen. (Andersen, Strunk, Temesgen, Atwood y Winterberger, 2014; Hudak, Crookston, Evans, Hall y Falkowski, 2008; Latifi, Nothdurft y Koch, 2010; Packalén y Maltamo, 2006). (Breidenbach et al., 2008; Maltamo, 2004; Vauhkonen, Mehtätalo y Packalén, 2011) investigaron modelos de efectos mixtos lineales con intercepciones al azar de nivel

de soporte, y (Breidenbach y Kublin, 2009; Junntila, Maltamo y Kauranne, 2008) investigaron métodos bayesianos.

(Asner et al., 2012; Breidenbach et al., 2008; Jensen, Humes, Conner, Williams y DeGroot, 2006; Lefsky, Cohen y Spies, 2001; Næsset, 2007) investigaron el uso de modelos comunes para las relaciones entre parámetros forestales y métricas LiDAR para diferentes condiciones forestales. Para variables respuesta como el volumen, la biomasa y el carbono, las perspectivas de uso de modelos comunes resultaron positivas bajo la condición de que los sistemas de sensores y los parámetros de adquisición fueran estables en las áreas de estudio (Næsset, 2009). (Næsset y Gobakken, 2008) apuntaron que la calidad del ajuste de los modelos a los datos de la biomasa aérea mejora significativamente si se incluyen variables que relacionaban proporciones de píceas y planifolios.

4.4.3. Enfoques de análisis a nivel de árbol individual

Para los inventarios a nivel de árbol, los datos LiDAR se usan típicamente para detectar copas de árboles individuales y para predecir atributos de interés usando conjuntos de modelos alométricos. Algoritmos y técnicas similares a los desarrollados para las imágenes aéreas se han empleado con este tipo de datos (Hyypä et al., 2008). Los inventarios a nivel de árbol consisten en una secuencia de pasos que incluye la detección de árboles, la extracción de características y la estimación de los atributos de los árboles. Para estas aplicaciones, las contribuciones de datos LiDAR incluyen la estimación de las alturas del dosel

vegetal, estimación de propiedades físicas tales como diferencias de altura entre las copas de árboles potenciales durante la detección de árboles, la estimación de alturas individuales de los árboles y tamaños de copas utilizados como entrada para modelos alométricos (Persson, Holmgren y Söderman, 2002) y la resolución de problemas relacionados con las variaciones de geometría y radiometría de imágenes espectrales (por ejemplo, (Mäkinen, Korpela, Tokola y Kangas, 2006)).

El enfoque de "árbol individual" se basa en la localización, delineación de copa, medida de la altura y extracción de algunos parámetros adicionales para cada árbol presente en la zona de estudio. Las variables de masa para parcelas, rodales, cantones o unidades superficiales mayores, pueden ser deducidas de las obtenidas de los pies que las componen.

Este enfoque trabaja de forma muy similar a la de los inventarios forestales pie a pie o por muestreo en parcelas de radio fijo o variable, pero con la ventaja de reducir el trabajo de campo a la necesaria verificación de la fiabilidad de los resultados obtenidos. Por ello, en la actualidad es la forma más extendida de trabajar cuando se pretende obtener datos para la gestión diaria de los montes, mientras que el enfoque de áreas o superficies se utiliza en estudios más generales como cálculo de biomasa o caracterización de superficies homogéneas en el monte. Además, la individualización de árboles permite obtener resultados con el mismo nivel de precisión para cualquier superficie estudiada.

La detección de árboles individuales generalmente se basa en modelos raster de altura de dosel (Canopy Height Model, CHM) interpolados a partir de datos de altura LiDAR (por ejemplo, (Persson et al., 2002; Popescu, Wynne y Nelson, 2003) aunque también hay ejemplos de utilización de técnicas basadas en nubes de puntos, especialmente para su uso en métodos de segmentación. Una consideración importante es que normalmente no se pueden detectar todos los árboles. El grado de detección de árboles individuales se ve afectado por el algoritmo de detección y su parametrización (Kaartinen et al., 2012). Además, el éxito de la detección está fuertemente afectado por las condiciones de los bosques, en particular la densidad de las parcelas y el patrón espacial, lo que sugiere problemas relacionados tanto con copas de árboles entrelazados como con árboles debajo del dosel dominante. Los métodos para resolver estos problemas incluyen la selección de un nivel apropiado de filtrado a priori (Ene, L., Næsset, E. y Gobakken, T., 2012; Heinzl, Weinacker y Koch, 2011), el uso de datos de onda completa (Reitberger, Schnörr, Krzystek y Stilla, 2009) y el de algoritmos 3D mejorados (Lahivaara et al., 2014; Tang, Dong y Buckles, 2013).

Las tareas de estimación de este enfoque requieren que la información relevante para estimar los atributos de los árboles de interés debe extraerse de los datos de LiDAR. La estimación a nivel de árbol combina mediciones directas, propiedades específicas de especies que se pueden predecir a partir de los datos LiDAR y alometría de árboles en forma de relaciones dimensionales entre partes de plantas.

Además de las estimaciones de localizaciones, el componente de detección y delineación de árboles típicamente produce alturas, dimensiones de corona y distribuciones de altura e intensidad para segmentos de árbol individuales (por ejemplo, Holmgren y Persson 2004). Las estimaciones de estas últimas propiedades son particularmente útiles para el reconocimiento de especies.

El uso de modelos alométricos para estimar atributos como el diámetro normal (DN) plantea problemas, ya que las relaciones entre las dimensiones de los árboles medibles y el DN están lejos de ser deterministas, sino más bien afectadas por factores como la densidad de árboles y el manejo silvícola (Korpela et al., 2007; Maltamo et al., 2007). Se han investigado métodos de regresión geoméricamente ponderados como un medio para superar al menos algunos de estos problemas (Salas, Ene, Gregoire, Næsset y Gobakken, 2010).

Múltiples estudios a nivel de árbol se han centrado en la detección de árboles individuales, la extracción de características y los pasos necesarios para la estimación a partir de este procedimiento. Sin embargo, pocos estudios han aportado estimaciones a nivel de parcela completa obtenidas por este procedimiento (Korpela et al., 2007; Peuhkurinen, Mehtätalo y Maltamo, 2011).

La exactitud de las estimaciones agregadas depende de múltiples factores que incluyen tanto la detección del propio árbol como los errores de estimación a nivel ese nivel. Los árboles no detectados y los errores en las

predicciones del modelo alométrico degradan la exactitud de las estimaciones a nivel de parcela y restringen seriamente la aplicación del modelo de árbol individual en los bosques semi-naturales (Korpela et al., 2007). Se están desarrollando enfoques que tratan de eludir los efectos de estas fuentes de error, como la detección semiespecular de árboles (Breidenbach, Næsset, Lien, Gobakken y Solberg, 2010). El enfoque de árbol individual puede ser, sin embargo, más interesante para aplicaciones tales como la planificación de la adquisición de madera para la cual no se requieren necesariamente estimadores no sesgados (por ejemplo, (Peuhkurinen, Maltamo, Malinen, Pitkänen y Packalén, 2007)).

Por último, debemos señalar las propiedades complementarias de los enfoques de árboles individuales y los enfoques de área. Además de la detección de árboles reales, los datos sobre árboles individuales pueden ser generados mediante la predicción de distribuciones de diámetro con el enfoque basado en el área, mientras que la combinación de estos enfoques se ha encontrado especialmente beneficioso para reducir los errores de detección de árboles (Breidenbach et al., 2010; Lindberg, Holmgren, Olofsson, Wallerman y Olsson, 2010; Maltamo, 2004) propusieron generar datos de referencia de campo para el enfoque basado en el área mediante enfoques modificados de detección de árboles. (Vauhkonen et al., 2011) utilizaron una predicción de diámetro normal basada en métodos de área para orientar los análisis detallados

de los niveles de árboles para rodales suficientemente maduros para la tala.

4.5. El LiDAR en el estudio de la biomasa forestal

En este apartado se mostrarán ejemplos de estudios e investigaciones llevadas a cabo mediante el empleo de la tecnología LiDAR para el estudio de la biomasa forestal, tanto a nivel de los componentes del árbol, como del árbol completo. Así mismo se analizarán los métodos y resultados de los dos grandes enfoques en el estudio de la biomasa forestal desde plataformas aerotransportadas: los métodos de masas y los métodos de árbol individual.

4.5.1. Estimación de la biomasa. Definiciones y Términos

Se ha mencionado arriba que la biomasa de un árbol puede dividirse en fracciones o componentes. Las descripciones de los componentes típicos de la biomasa se dan en la Tabla 6, y algunos se ilustran en la Figura 15. Algunos componentes - como el follaje - se definen directamente, pero otros necesitan más especificaciones para describirse sin ambigüedad. La biomasa se cuantifica típicamente en peso seco, que es el peso total excluyendo el peso del agua contenida en la materia en cuestión. A menudo se encuentran cifras de biomasa dadas en una base por área, típicamente en unidades de t/ha, es decir, tonelada métrica por hectárea.

Tabla 6 Componentes de biomasa del árbol

Componente	Observación
Biomasa total sobre el suelo	Debe definirse la altura del tocón.
Biomasa del tronco	Puede dividirse en biomasa de corteza y de tronco o puede referirse a biomasa de tronco comercializable
Biomasa de ramas	Puede dividirse en ramas vivas y muertas
Biomasa de follaje / acículas	
Biomasa de tocón	
Biomasa de copas	Típicamente la suma de la biomasa de rama y follaje
Biomasa total subterránea	La inclusión del tocón debe ser definida, también mínima espesor de las raíces
Biomasa de raíz	
Total biomasa arbórea	

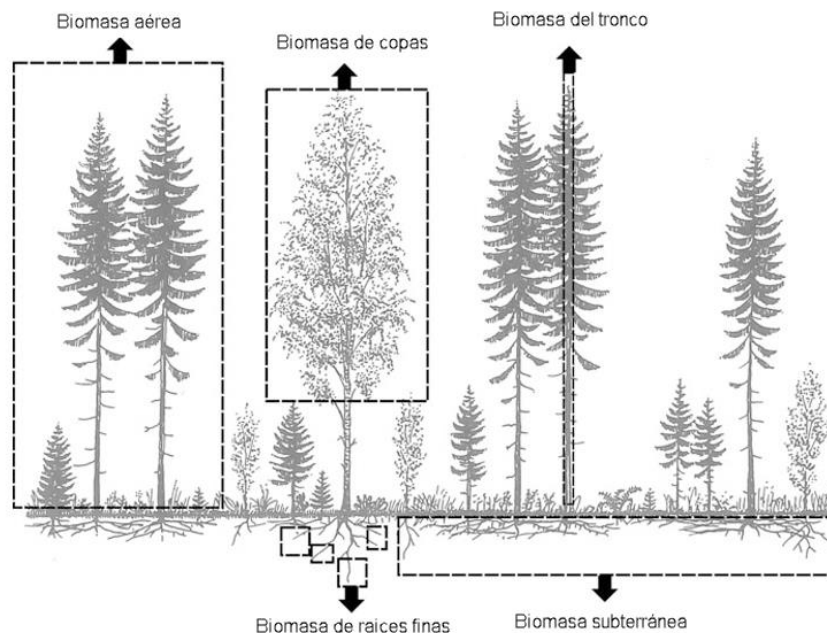


Figura 15 Componentes de biomasa de árbol (Matti Maltamo, Erik Næsset y Jari Vauhkonen, 2014)

El contenido de carbono de la madera y la corteza de los árboles es aproximadamente el 50% de la biomasa seca, por lo tanto, las estimaciones de carbono se pueden obtener multiplicando las estimaciones de biomasa seca por 0,5. (Montero González et al., 2006) fijan este valor en 49,7 para el *P. radiata*). Como tal, la mayoría de los esfuerzos de estimación de carbono almacenados por los bosques se centran en la estimación de la biomasa seca.

Para obtener información sobre estas fracciones de biomasa a partir de los datos LiDAR se debe encontrar una relación directa o indirecta entre los datos de teledetección y el componente de biomasa en cuestión. Lo que aquí se entiende por relación directa se puede ilustrar con un ejemplo: considérese un caso en el que la altura de un árbol se reduce cortándola a mitad del tronco y retirando la parte superior. Esto afectará claramente a cómo los pulsos láser serán reflejados por este árbol, así que podemos decir que hay una relación directa entre la altura del árbol y los datos LiDAR. Si caváramos un túnel y elimináramos la mitad de las raíces de un árbol, no tendría ningún efecto sobre la forma en que los pulsos láser se reflejan de este árbol. Esto significa que no hay una relación directa entre la biomasa de las raíces y los datos LiDAR. Afortunadamente, sí existe una relación indirecta entre la biomasa de las raíces y los datos LiDAR, y esto se discutirá a continuación.

Si consideramos como componente terrestre la biomasa de tronco, que, a diferencia de las raíces, sí puede ser impactado por pulsos de láser, es razonable creer que

existe igualmente algún tipo de relación indirecta al estimar la biomasa de tronco. Después de todo, la mayor parte de la información inherente a los datos LiDAR no estará directamente relacionada con el tronco, ya que son pocos los pulsos que se reflejarán en él. Por lo tanto, en estos casos también debemos buscar una relación indirecta, que casi seguramente implicará el concepto de alometría, es decir, la relación sistemática entre las propiedades biofísicas -como el tamaño o las formas de ciertas partes- encontradas en muchos organismos vivos.

La existencia de estas relaciones alométricas es un factor determinante en el marco metodológico que constituye la base de las estimaciones forestales basadas en datos LiDAR y es especialmente visible en el proceso de estimación de los componentes de biomasa. Las relaciones alométricas - que ciertas propiedades de los árboles covaríen de una manera sistemática -, permiten el desarrollo de modelos alométricos, también conocido como ecuaciones alométricas.

Un ejemplo típico de un modelo alométrico con el que la mayoría de los forestalistas están familiarizados, es un modelo de Diámetro Normal (DN) / altura-volumen de árbol único, es decir, una ecuación de volumen. Al analizar un gran número de árboles muestreados con diámetro medido con exactitud a la altura del pecho (DN), la altura del árbol y el volumen del tronco pueden derivarse de un modelo predictivo expresado como una función que toma el DN y la altura como entrada y produce el volumen del tronco como la salida. Un propósito típico de este modelo es utilizar propiedades que son fáciles de medir, como DN y altura en

este ejemplo, para obtener una predicción de una propiedad que requiere mucho tiempo o es difícil de medir directamente, como el volumen del tronco. Estos principios de modelado también pueden usarse para otras propiedades de un árbol, y en particular la que aquí nos interesa, los modelos de biomasa alométrica.

La razón por la cual la alometría es importante al estimar la biomasa con datos LiDAR es que sólo hay ciertas propiedades -en particular la altura del árbol- que pueden estar directamente relacionadas con la información de los datos LiDAR. Un fragmento crucial de información que se necesita en estimaciones de biomasa basadas en el enfoque de área es la altura sobre el suelo de los ecos láser. Así que cuando el volumen del tronco se deriva utilizando modelos areales, se basa en el hecho de que el volumen está sistemáticamente relacionado con las alturas de los árboles.

4.5.2. Modelos alométricos de biomasa

A menudo se necesitan mediciones de campo de la biomasa de un árbol tipo como dato de referencia en la estimación de la biomasa a diversas escalas espaciales. Esto sirve igualmente para los métodos basados en la teledetección, por lo que incluso en la elaboración de cartografía a escala regional, los datos de referencia pueden consistir en mediciones de campo de la biomasa en árboles individuales. La biomasa de estos árboles se utiliza - junto con los datos de teledetección - para construir modelos predictivos.

Dado que el contenido de humedad de los árboles varía, se recurre habitualmente al peso seco para cuantificar la biomasa forestal, como hemos mencionado anteriormente. Sin embargo, la obtención de mediciones precisas del peso seco de los componentes de biomasa de un árbol implica un proceso que requiere mucho tiempo y trabajo, que incluye el muestreo, el secado y el pesaje destructivos. En consecuencia, el uso de modelos alométricos existentes - con frecuencia denominados ecuaciones de biomasa- es común cuando se obtienen datos de biomasa forestal de un solo árbol a partir de mediciones de campo. Estos modelos alométricos son generalmente modelos específicos de la especie con diámetro normal y a veces también, la altura del árbol u otras características como variables explicativas. Los modelos de biomasa alométrica se derivan generalmente de un conjunto de árboles muestreados de manera destructiva.

Existen estudios que han recogido y revisado un gran número de modelos de biomasa alométrica a escala regional y nacional, como (Zianis, Muukkonen, Mäkipää, Mencuccini y others, 2005) para Europa, (Jenkins, Chojnacky, Heath y Birdsey, 2003). Además de los modelos de biomasa total sobre el suelo, también existen en ciertas regiones modelos para la biomasa de componentes tales como ramas, corteza, raíces y follaje para ciertas especies.

En España son varios los estudios que se han llevado a cabo, algunos de los cuales se han mencionado arriba: (Canga et al., 2013; Montero González et al., 2006) Pérez-Cruzado, etcétera.

Los errores asociados con las predicciones de algunos componentes de biomasa de estos modelos alométricos -como la biomasa de ramas o raíces- suelen ser mayores que las predicciones de otros componentes como la madera del tronco. Significa que, por ejemplo, la biomasa de ramas varía más que la biomasa de la madera del tronco para un diámetro y una altura dados, lo que para muchas especies del árbol es intuitivamente razonable. Esto implica, sin embargo, que la biomasa de la madera de tronco en muchos casos puede predecirse con mayor precisión que, por ejemplo, la biomasa de ramas, utilizando estos modelos alométricos.

Se ha demostrado que la elección de modelos alométricos afecta las predicciones de un modelo de biomasa basado en datos LiDAR. (Zhao, Guo y Kelly, 2012) concluyen que el uso de modelos regionales de biomasa para derivar los datos de referencia en la mayoría de los casos es mejor que el uso de modelos nacionales, siempre y cuando los modelos regionales sean desarrollados con un conjunto representativo de datos. Dado que la recolección de datos de referencia de campo para modelos de biomasa alométrica es costosa y requiere mucho tiempo, algunos modelos regionales se basan en un número limitado de árboles muestreados, lo que podría afectar la aplicabilidad incluso a nivel regional. Por lo tanto, es aconsejable considerar a fondo la elección de modelos alométricos, cuando sea posible.

4.5.3. Métodos de relación de componentes y factores de expansión de la biomasa

Existe otro enfoque para las estimaciones de biomasa, puesto en marcha en el Programa de Inventario y Análisis Forestal del Servicio Forestal del Departamento de Agricultura de los Estados Unidos (USDA), que proporcionó estimaciones de biomasa y carbono a nivel nacional compatibles con las estimaciones de volumen a nivel de árbol. Se describe en (Heath, Hansen, Smith, Miles y Smith, 2009). El método se denomina *método de la relación de componentes* (*Component Ratio Method*, CRM) y sus pasos se pueden resumir de la siguiente manera: (1) se convierte el volumen de madera de tronco a biomasa usando un conjunto de densidades específicas de madera; (2) se calcula la biomasa de la corteza del tronco usando una compilación de muestras de corteza y densidades específicas de corteza; (3) se calcula la biomasa de ápices y ramas como una proporción de la biomasa de tallo basada en las proporciones de los componentes de Jenkins et al. (2003); (4) se calcula la biomasa del tocón a partir de las ecuaciones de (Raile y Station, 1982); y (5) se suman las partes para obtener la biomasa total sobre el suelo. La biomasa de las raíces se calcula como una proporción de la biomasa del tallo basada en las proporciones de los componentes descritos en (Jenkins et al., 2003).

Las ventajas de utilizar este enfoque son: (1) la existencia de una gran experiencia en el campo de la investigación y de la operativa de inventarios forestales en la extracción de estimaciones de volumen de árboles

individuales; (2) la vinculación de la biomasa al volumen y luego la conversión de la biomasa en carbono mediante el factor de conversión 0,5, proporciona no sólo estimaciones consistentes de volumen y biomasa, sino también estimaciones de carbono que coinciden con las predicciones de volumen y biomasa.

Basándose en la misma lógica de la relación entre la biomasa y el carbono, existe una relación entre la biomasa y el volumen. (Somogyi et al., 2008) recopilaron una base de datos para calcular la biomasa y el carbono para bosques de la región eurasiática a partir de variables proxy derivadas de los inventarios forestales, como el volumen de los árboles. Esta base de datos contiene varios tipos de expansión, conversión y factores combinados, por varias especies de árboles o grupos de especies, para calcular la biomasa de árboles individuales. Este enfoque se basa en la siguiente ecuación que involucra factores de expansión de biomasa (BEF):

$$\text{Biomasa (t)} = V \times D \times \text{ExpF} \times R$$

Donde V es el volumen (m³), D es la densidad de la madera (t m⁻³), ExpF es el factor de expansión de la biomasa (adimensional), y R es la relación raíz-brote (adimensional, la relación entre la raíz y la biomasa total).

La detección remota LiDAR proporciona una estimación directa de la elevación de las superficies interceptadas. En el contexto de la vegetación, si se pueden distinguir las señales del suelo y la vegetación, se pueden calcular

las alturas relativas sobre el suelo de las copas de los bosques. Dado que se necesita un diámetro de tallo y una estructura de copa adecuados para soportar las dimensiones de los árboles, la altura de la vegetación está estrechamente relacionada con el volumen y por lo tanto con la biomasa.

4.5.4. LiDAR satelitales y estimaciones de la biomasa a escala global

El Geoscience Laser Altimeter System (GLAS) de la NASA a bordo del Satélite ICESat es actualmente el único sistema LiDAR por satélite que ha proporcionado una cobertura de muestreo casi global durante un período prolongado de tiempo. Por lo tanto, ofrece un conjunto de datos único de perfiles verticales de la superficie de la Tierra.

ICESat fue lanzado en enero de 2003 y la misión continuó hasta que el láser final cesó de trabajar en el otoño de 2009. Durante este período, el láser fue empleado durante aproximadamente un mes al año, en primavera y en otoño, y además durante el verano anterior a la Vvda de la misión.

Se trata de un sistema LiDAR de onda completa, que operaba a una altitud de 600 km, viajando a 26.000 km/h emitiendo pulsos láser de 1064nm a 40Hz. Esto proporcionó un muestreo de la superficie de la Tierra con intervalos de centro de huella a 172 m de distancia. El diámetro y la excentricidad de la huella han variado considerablemente entre las campañas láser desde un eje principal de $148,6 \pm 9,8$ m hasta $51,2 \pm 1,7$ m (Figura X). La información completa sobre el sensor está disponible en Abshire et al., 2005;

Brenner et al., 2003; NSIDC, 2010a; Schutz et al., 2005; Schutz, 2002. Los datos y herramientas para procesarlos están disponibles gratuitamente en NSIDC, 2010b.

Características del GLAS El patrón de muestreo sistemático y la representación del perfil de vegetación en la forma de señal de LiDAR de onda completa permite mapear la distribución espacial de parámetros de vegetación para grandes áreas. La cobertura estacional permite repetir mediciones cercanas a una frecuencia que no sería factible utilizando métodos de muestreo convencionales.

Sin embargo, el sistema fue diseñado principalmente para aplicaciones criosféricas y por lo tanto la configuración no se considera óptima para el análisis de vegetación. En terreno inclinado, con vegetación, la vegetación y las superficies del suelo pueden ocurrir a las mismas elevaciones. Esto hace que las señales de tierra y vegetación se combinen dentro de la forma de onda y, cuando no es posible distinguir un pico de tierra, hace que la señal devuelta de la vegetación sea razonablemente identificada. Además, la nubosidad densa provoca que la ausencia de señales de retorno válidas cause huecos de muestreo. Para las regiones con una alta cobertura de nubes como las zonas tropicales, esto puede servir para empeorar la ya escasa densidad de muestreo cerca del ecuador, producida por la órbita polar del satélite.

Dentro de cada huella, asumiendo que la parte superior del dosel es el comienzo de la señal de la forma de onda, la exactitud con que la señal devuelta por la vegetación puede

ser identificada depende de la capacidad de identificar una superficie del terreno representativa en la dicha forma de onda. Los métodos que se han empleado para lograr esto han incluido el uso de un MDT independiente para incorporar la inclinación del terreno dentro de las huellas LiDAR (Lefsky et al., 2005; Rosette, North y Suárez, 2008) o utilizar la descomposición gaussiana de la forma de onda para localizar un pico correspondiente al suelo (SUN, RANSON, Kimes, BLAIR y KOVACS, 2008).

Por lo tanto, la altura de la vegetación puede estimarse como la diferencia de elevación entre el inicio de la señal de forma de onda y la superficie de tierra identificada estimada dentro de la forma de onda. Los estudios anteriores informan que el RMSE es de 3+ metros.

Aplicaciones en la estimación de la biomasa a escala global. Por lo general, los índices de forma de onda de la Altura de la Mediana de la Energía (*Height of Median Energy*, HOME) o percentiles de altura relativa por encima del suelo (*Relative Height Percentiles Above Ground*, RHI) se calculan utilizando la distribución de la energía acumulada dentro de esta región de la forma de onda retornada por la vegetación. Lefsky et al., 2007, ideó un método alternativo de la estimación de un parámetro de altura de la vegetación que representa la rugosidad del terreno y de la vegetación mediante la forma de onda bordes anterior y posterior en lugar de aislar la señal devuelta de la vegetación.

Se han utilizado técnicas similares de fusión de datos para determinar la distribución de la biomasa por varios autores y abarcando una amplia gama de tipos de vegetación y se han aplicado desde escalas regionales a continentales (Fatoyinbo y Simard, 2013). El último grupo de investigación del Centro de Investigación Woods Hole, dirigido por Josef Kellndorfer, está trabajando para estimar la biomasa de los bosques tropicales a nivel mundial (Baccini et al., 2009). Además, la investigación utilizando GLAS está en progreso como parte de la NASA Carbon Monitoring System initiative (NASA, 2017) para determinar la distribución de biomasa dentro de los Estados Unidos (así como para producir mapas de biomasa de mayor resolución a nivel de condado utilizando LiDAR de datos aéreos).

Los productos globales de superficies de vegetación derivados de GLAS combinados con datos de sensores ópticos (Lefsky, 2010) o combinando datos GLAS y radar abren múltiples posibilidades para mejorar la comprensión de procesos globales y permitirán una mejor la evaluación de la biomasa a escala global.

4.5.5. Métricas y modelos de predicción de biomasa forestal derivadas de datos LiDAR

Existen varios tipos de técnicas de regresión en uso para predecir los componentes de la biomasa a partir de datos LiDAR. En general, se utilizan los mismos tipos de modelos estadísticos para estimar los componentes de la biomasa que los utilizados para la estimación del volumen del

tronco y otras características del bosque utilizando datos LiDAR.

Los modelos de regresión lineal son los más comunes y suelen implicar técnicas de reducción de en el caso de que las variables derivadas LiDAR sean numerosas o altamente correlacionadas. También se han utilizado con éxito modelos de regresión no paramétricos basados en técnicas de aprendizaje automático tales como *random forest machines* y *support vector machines* para predecir componentes de biomasa forestal.

Tal y como se ha mencionado, es necesario encontrar una relación entre los datos LiDAR y el componente de biomasa objetivo para desarrollar un modelo predictivo. Esta relación puede ser directa o indirecta, pero distinguir entre los dos tipos de relación en la práctica no resulta necesaria, ni siempre es posible. Son muchos los autores que opinan que en el contexto del desarrollo de modelos predictivos y la búsqueda de variables predictoras adecuadas, basta con encontrar algún tipo de relación, (Hauglin, Dibdiakova, Gobakken y Næsset, 2013).

En algunos procedimientos de modelado -como la regresión escalonada- los parámetros reales derivados de datos LiDAR que se usan como variables predictoras en el modelo se determinarán a través del proceso de ajuste del modelo y, típicamente, se compone de un subconjunto de un amplio rango de variables extraídas antes de la adaptación del modelo.

Un enfoque diferente sería seleccionar a priori una o algunas características derivadas LiDAR que se cree que será capaz de predecir la relación entre los datos LiDAR y la propiedad biofísica deseada. Un ejemplo de esto podría ser usar la altura máxima sobre el suelo de los ecos asignados a un solo árbol para modelar la altura del árbol o el DN. Otro ejemplo podría ser el uso de características geométricas derivadas de los ecos asignados a la copa de un solo árbol como predictores para propiedades como la biomasa de la copa o del follaje. Más concretamente, esto podría significar utilizar el volumen de un casco convexo tridimensional o el área de una proyección bidimensional de los ecos láser asignados a una corona de árbol como variables predictoras, como lo hizo (Hauglin et al., 2013) y (Kankare, Vastaranta et al., 2013).

4.5.5.1. Estimación de componentes de biomasa mediante datos LiDAR

La mayoría de estos estudios siguen la metodología basada en área descrita arriba, es decir, se predice la biomasa a nivel de parcela, de masa o a nivel regional.

El enfoque metodológico típico para la estimación de componentes de biomasa a través de LiDAR es similar al Método de Áreas descrito en el que los volúmenes de tronco se predicen para unidades de área.

El enfoque se puede resumir de la siguiente manera: Los datos LiDAR se asignan espacialmente primero a unidades areales, típicamente celdas raster o parcelas muestrales,

se analizan los pulsos de retorno para cada unidad para posteriormente extraer un conjunto de parámetros estadísticos. Generalmente los parámetros se obtienen a partir de los cálculos de las distribuciones de los datos de altura de los ecos láser y suelen incluir medidas como momentos y estadísticas de orden de las distribuciones.

Estos parámetros LiDAR son posteriormente relacionados a valores de referencia tomados en medidas de campo mediante modelos de regresión. Cuando se calcula la biomasa, estos valores de referencia se derivan habitualmente de modelos alométricos existentes.

Dado que la biomasa se obtiene utilizando modelos alométricos, las mediciones reales de campo realizadas en cada parcela son típicamente registros de parámetros de árbol como el Diámetro Normal, especie y a menudo alturas de árboles. Estas mediciones reales representan las entradas a los modelos alométricos, que a su vez producirán la biomasa como salida. Esta biomasa modelada se agregará para todos los árboles en una parcela y se utilizará como valor de referencia.

El supuesto subyacente aquí es que la cantidad de biomasa del componente seleccionado está relacionada con los datos LiDAR. Esto suele ser el caso cuando los modelos alométricos que emplean el DN como uno de la principal entrada, ya que el DN está indirectamente relacionado con los datos LiDAR. Como se ha descrito anteriormente, los errores asociados a la biomasa predicha a partir de modelos alométricos varían según el componente de biomasa, lo que

significa que, por ejemplo, la biomasa subterránea en la mayoría de los casos será modelizada con un error mayor que la biomasa de tronco. Debido a la aleatoriedad de estos errores, tenderán a disminuir al producir agregaciones a nivel de parcela.

El siguiente paso consiste en desarrollar modelos de regresión predictiva con las mediciones de campo de nivel de parcela - en este caso biomasa - y métricas derivadas de datos LiDAR de las mismas parcelas, siguiendo el método basado en área descrito arriba.

Las ecuaciones de regresión se emplearán en el siguiente paso para predecir la biomasa en nuevas unidades de área, típicamente celdas raster. Estas pueden posteriormente ser agregadas para obtener cifras de biomasa para las unidades espaciales deseadas, por ejemplo, rodales forestales.

La mayoría de las veces la modelización se realiza para la biomasa total sobre el suelo, pero algunos estudios también aportan estimaciones separadas para componentes de biomasa, como follaje, ramas o incluso la biomasa subterránea (Hauglin, Gobakken, Lien, Bollandsås y Næsset, 2012; Lim y Treitz, 2004; Næsset y Gobakken, 2008). Esto se logra simplemente aplicando un modelo alométrico apropiado al calcular la biomasa de referencia en las parcelas de muestra.

En lugar de utilizar variables del modelo derivadas de la distribución de altura de los ecos LiDAR, algunos investigadores han propuesto métodos en los que no sólo se tienen en cuenta las alturas de los ecos láser, sino

también la distribución planar de los ecos. Esto se ha realizado habitualmente en estudios que identifican árboles individuales y donde se extrae información relacionada con estos árboles identificados. Los estudios descritos por (Bortolot y Wynne, 2005; Gleason y Im, 2012) son ejemplos de este último enfoque, pero las predicciones son en estos casos aún se hace a nivel de la parcela.

Los pasos para la metodología basada en área, en lo que se refiere a la obtención de la biomasa total o la biomasa de componentes, se resumen en la Fig.

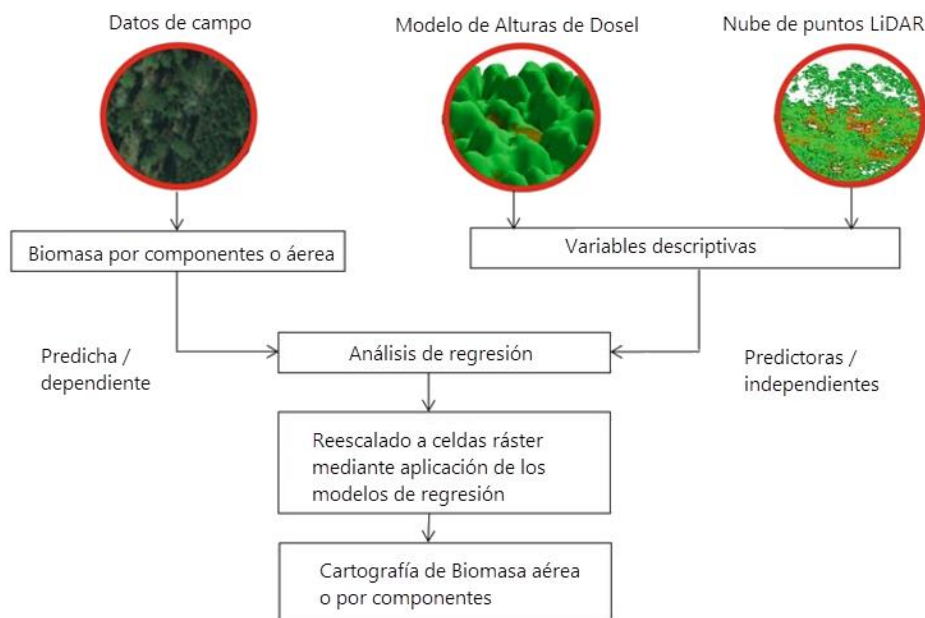


Figura 16 Resumen de los pasos de la metodología de masa/área (Matti Maltamo et al., 2014)

Una desventaja potencial del enfoque basado en área es que los modelos de regresión utilizados para vincular las

métricas derivadas LiDAR para trazar la biomasa tienen una portabilidad limitada a otras especies u otras áreas de estudio, o pueden cambiar en el tiempo para la misma área a medida que la estructura tridimensional del bosque sufre cambios. Más aún, los modelos pueden depender de los instrumentos utilizados para recolectar datos porque los diferentes sistemas LiDAR aerotransportados responden de manera diferente al mismo bosque (Næsset y Gobakken, 2008) (Næsset y Gobakken 2008).

Las métricas de datos LiDAR que pueden resultar significativas en la explicación de la biomasa a nivel de parcela y los modelos de regresión referenciados en la literatura a menudo carecen de puntos en común, p.ej. las métricas que incluyen la altura media, máxima y media de la copa, la altura media cuadrada de la copa o las alturas cuantil. Estas métricas se han utilizado solas o combinadas en modelos lineales o no lineales, con o sin transformaciones, habitualmente logarítmicas.

En el contexto de la exploración láser aerotransportada de los parámetros estructurales de los bosques, la mayoría de los modelos de estimación descritos en la literatura existente generalmente no sólo son específicos del lugar o de la especie, sino que también son dependientes de la escala, debido a que los modelos se ajustan utilizando datos recogidos en tamaños de parcela determinados y se debe aplicar a una escala o un tamaño de píxel (celda ráster) apropiado al tamaño de parcela utilizado en el modelo de ajuste (Næsset, 2002).

Derivar la biomasa para componentes individuales del árbol se vuelve aún más difícil para el enfoque basado en área, dado que las ecuaciones alométricas para componentes de la biomasa han sido desarrolladas a nivel de árbol individual para las especies de árboles más comunes.

4.5.5.2. Estimación de componentes de biomasa a partir del enfoque de Copa de Árbol Individual (CAI)

El otro enfoque alternativo para estimar los atributos forestales a partir de los datos LiDAR es denominado de Copa de Árbol Individual (CAI) o de Individualización de Copas. En (Sorin C Popescu, 2007) se describe cómo se pueden estimar los componentes de la biomasa utilizando datos LiDAR a nivel de árbol. En el enfoque CAI descrito por (Sorin C Popescu, 2007) se calcula el DN a partir de los datos LiDAR y luego se puede usar este DN derivado con un modelo alométrico para obtener una estimación de biomasa para el árbol en cuestión. Un enfoque similar se utilizó anteriormente por (Hyypä, Kelle, Lehikoinen y Inkinen, 2001) para derivar volúmenes de tronco de árboles individuales.

Aunque el DN no se mide directamente mediante datos LiDAR, se pueden obtener buenas estimaciones de DN utilizando la información que se capta mediante los datos LiDAR, en particular las métricas relacionadas con la altura y el tamaño de la corona. La relación alométrica entre altura y DN resulta robusta para muchas especies, y se utiliza en numerosos modelos alométricos.

En el estudio particular descrito por (Sorin C Popescu, 2007) se utilizó la altura y el diámetro de la copa derivados de LiDAR para estimar el DN para los pinos de Loblolly. Cuando los datos LiDAR se utiliza para identificar y mapear árboles individuales, las dimensiones y la ubicación de los árboles individuales se derivan con alta precisión, por lo tanto, pueden generarse mapas de la biomasa de árbol individual aplicándolo a la extensión de la cobertura aérea de los datos LiDAR. Este enfoque fue utilizado por varios estudios, como (Sorin C Popescu, 2007; Zhao, Popescu, Meng, Pang y Agca, 2011). La biomasa del tallo se asignó al píxel en el mapa de biomasa correspondiente a la ubicación del árbol y la biomasa del follaje se distribuyó uniformemente sobre los píxeles cubiertos por la copa. Este mapa de biomasa se utilizó para validar estimaciones de biomasa hechas desde el espacio, como las derivadas del Geoscience Laser Altimeter System (GLAS) a bordo del Ice Cloud and land Elevation Satellite (ICESat) (Popescu, Zhao, Neuenschwander y Lin, 2011).

(Zhao et al., 2009a). El mapa se utilizó como dato de referencia para sintetizar los conjuntos de datos de entrenamiento y de prueba a múltiples escalas para validar dos modelos de biomasa invariante en escala, un modelo funcional lineal y un modelo no lineal equivalente que utiliza distribuciones de altura de copas derivadas LiDAR y funciones de cuartil de altura de copa como predictores, respectivamente.

La estimación de componentes de biomasa mediante el enfoque CAI presenta varias ventajas sobre el enfoque MBA, sobre todo que, una vez que los árboles se cartografían con los datos LiDAR, el método CAI puede aplicarse para derivar componentes de biomasa para modelos alométricos a nivel de árbol. En otras palabras, este método aprovecha los modelos alométricos para biomasa de masa existentes mientras que los datos LiDAR sólo se emplean para el mapeo y la medición precisa de los árboles. En la biometría forestal tradicional, las relaciones alométricas se derivan a partir de mediciones de campo a nivel de árbol individual y no a nivel de parcela o área. Con la MBA, en sí misma, la relación entre las métricas de datos LiDAR y la biomasa por unidad de área no tiene otro significado práctico que no sea para extrapolar los valores de biomasa a la extensión de los datos LiDAR.

Al igual que con el MBA, también los componentes de la biomasa - como las raíces - que no tienen relación directa con los datos LiDAR se pueden estimar usando el enfoque CAI. Sin embargo, hay algunos componentes de la biomasa, en particular ramas y follaje, donde los datos LiDAR pueden contener información más directa. En muchos casos, la mayoría de los ecos láser que interactúan con los árboles individuales se reflejan en las copas de los árboles, por lo tanto, en los elementos como el follaje y las ramas. Esto sugiere que los datos LiDAR pueden contener información directa sobre la superficie reflectante, consecuentemente sobre la biomasa de las ramas y del follaje.

(Hauglin et al., 2013) demostraron en un estudio con piceas de Noruega que los datos LiDAR en realidad pueden contener más información con respecto a la biomasa de rama que las mediciones de campo del DN. Un enfoque alternativo podría ser, por tanto, estimar la biomasa de estos componentes directamente de los datos LiDAR. Aunque es atractivo utilizar los datos de forma remota de manera tan directa, hay algunos desafíos que deben resolverse para poder utilizar plenamente esta información.

En el cartografiado de la biomasa forestal debe asegurarse que los datos utilizados para desarrollar el modelo corresponden a los recursos forestales que van a ser estimados. En esquemas operativos como los utilizados en los inventarios forestales esto se logra utilizando una muestra estadística del área objetivo como datos del modelo. Esto significa que se debe recolectar un conjunto de datos para cada área inventariada. (Hauglin et al., 2013; Kankare, Rätty et al., 2013; Kankare, Vastaranta et al., 2013; Rätty et al., 2011) investigaron el uso de los datos obtenidos mediante muestreos destructivos para el modelado directo de la biomasa de ramas a partir de datos LiDAR.

En la mayoría de los escenarios operativos esto suponía costes prohibitivamente altos, y por lo tanto no resultaban una opción. Sin embargo, apuntaban que el desarrollo del método podría resolver este problema utilizando datos obtenidos por otros medios, como los obtenidos usando la *Exploración Láser Terrestre* (ELT).

La cartografía de los componentes de biomasa o de biomasa a nivel de árbol individual, siguiendo el enfoque general, tiene la ventaja de utilizar ecuaciones alométricas específicas de especies ampliamente investigadas y probadas en la literatura existente o la práctica forestal, ya sean ecuaciones regionales o generalizadas.

Otra ventaja es el hecho de que el enfoque de los CAI trae efectos de escala mínimos al inventario de la biomasa forestal. Los mapas de la biomasa a escalas por encima de los niveles de árbol surgirán integrando los resultados de los árboles hasta la escala deseada. Siempre y cuando los algoritmos para el procesamiento de los datos LiDAR y la identificación de los árboles individuales se realicen con la suficiente precisión, se podría esperar una gran precisión en la derivación espacial de biomasa siguiendo este enfoque. Sin embargo, como se puede observar en la literatura, la segmentación CAI y los algoritmos de medición varían en su rendimiento de identificar todos los árboles correctamente y poder derivar así sus mediciones.

Uno de los principales problemas es la ineficiencia de muchos algoritmos de segmentación de árboles para identificar el sotobosque y los árboles suprimidos, o para delinear los árboles bajo condiciones de cubierta muy densa. Los errores sistemáticos que surgen de esta delimitación errónea de las copas motivaron el desarrollo de enfoques semi-CAI, como alternativa al enfoque general de los CAI. En algunos casos el enfoque de CAI puede ser eficaz porque los principales contribuyentes a la biomasa aérea son los árboles dominantes (Zavitkovski, 1976) que

podrían identificarse más fácilmente mediante algoritmos de delineación de árboles a partir de datos LiDAR.

4.5.5.3. Estimación de la biomasa a nivel de masa

El inventario a nivel de masa del LiDAR aerotransportado sigue el método originalmente diseñado por Næsset, 1997a, b. Como se ha mencionado, estos dos estudios son particularmente relevantes debido a que ambos han fomentado los trabajos basados en la noción de que los datos LiDAR pueden ser utilizados en inventarios forestales a gran escala, siempre que los datos georeferenciados de las parcelas de campo puedan ser utilizados en una primera fase para desarrollar relaciones empíricas entre métricas LiDAR tales como los percentiles de altura relativa sobre el suelo y los principales parámetros de manejo forestal. Estas relaciones se utilizan para estimar, en una segunda fase, los parámetros del bosque para todas las parcelas de testeo en el área de estudio, conocido como procedimiento de dos etapas para el inventario de masas forestales (Næsset, 2002).

De esta manera, este enfoque proporciona una estimación útil de parámetros clave tales como la altura máxima, la cubierta del dosel, la densidad de árboles, el área basal y el volumen. Las relaciones establecidas permiten que parámetros forestales como la biomasa se deduzcan directamente del uso de métricas de LiDAR y que esta evaluación se implemente en áreas forestales grandes donde haya cobertura LiDAR.

Este método tiene, por supuesto, algunas limitaciones. Depende en gran medida de la recopilación de abundantes datos sobre el terreno para formar relaciones empíricas entre datos de campo y LiDAR que a menudo no son fácilmente transferibles a otras áreas de estudio. Pueden existir tipos de relaciones diferentes para especies morfológicamente similares (piceas de Noruega, picea de Sitka o alerce) y adicionalmente, la estructura del rodal puede resultar un factor determinante. En este enfoque no se pueden parametrizar los efectos derivados de la distribución espacial de los calveros, su tamaño y la distribución espacial de los pies (ya sea perfectamente alineados, creciendo en rodal natural o aclarados en diferentes intensidades) sobre la interceptación vertical de los impactos láser, lo que puede ocasionar la aparición de valores anómalos en las métricas LiDAR.

Sin embargo, el valor real de este método se ha demostrado en estudios regionales, donde la combinación de LiDAR y mediciones de campo es capaz de optimizar los muestreos tradicionales, particularmente como parte de los inventarios forestales de grandes superficies (Hollaus, Wagner, Schadauer, Maier y Gabler, 2009).

En España (Cuasante, 2009) aplicando esta metodología en montes de coníferas en la provincia de Burgos, obtuvieron modelos de regresión para las variables de masa: densidad de árboles, área basimétrica, volumen de madera, diámetro medio cuadrático y altura dominante para cinco estratos diferentes con R^2 entre 0,37 y 0,65, 0,65 y 0,83, 0,65 y

0,84, 0,22 y 0,77, 0,60 y 0,94 y RMSE entre 25% y 39%, 18%, 26%, 16% y 31%, 10% y 21%, 5% y 10% respectivamente.

También en España (Condés y Riaño, 2005) aplican esta metodología para el cálculo de la biomasa foliar de una masa de pino silvestre obteniendo una ecuación basada en la altura del punto más alto de cada celda con R^2 de 0,82 y RMSE menor del 10%

(García, 2011) aplica esta metodología para la realización de un inventario forestal obteniendo ecuaciones de volumen total, biomasa total aérea, área basimétrica y pies por hectárea con R^2 de 0,90, 0,89, 0,89 y 0,80 respectivamente.

CAPÍTULO 3

EL ÁREA DE ESTUDIO

5. El área de estudio: delimitación y presentación general

La elección de Bizkaia como marco de este estudio de carácter fundamentalmente metodológico es debido fundamentalmente a dos motivos: en primer lugar, por tratarse de un territorio de gran potencial energético desde el punto de vista de la biomasa forestal. Según el Banco de Datos del Mapa Forestal datos del último IFN el 58% del territorio lo ocupa el bosque -46% bosque de plantación- ("Usos del suelo. Banco de datos del Mapa Forestal," 2006). En segundo lugar, siendo uno de los objetivos principales del estudio el estimar la biomasa a partir de datos LiDAR Aerotransportados, poder disponer de datos de este tipo con la suficiente calidad se antoja imprescindible. El Departamento de Medio Ambiente, Planificación Territorial, Agricultura y Pesca del Gobierno Vasco realizó un vuelo LiDAR en 2008 que cubría las provincias de Álava y de Bizkaia y ofrecía las suficientes garantías de calidad para poder ser empleado en este ensayo metodológico. Además, la Infraestructura de Datos Espaciales *Geoeskadi* ofrece acceso libre a dichos datos, por lo que se consideró una buena oportunidad para llevar a cabo esta tesis.

La Provincia de Bizkaia tiene una superficie de 2.236 km², que representan el 30,66% de la Comunidad Autónoma Vasca y se encuentra situada al norte de la península ibérica, limita al oeste con la Comunidad Autónoma de Cantabria, al

sur con la provincia de Burgos y el territorio histórico de Álava, al este con Gipuzkoa y al norte con el mar Cantábrico.

Presenta un relieve montañoso especialmente en el sur y en el este de la provincia, aunque se dan, en general, formas suaves y poco abruptas. La red hidrográfica es moderadamente densa y está relativamente jerarquizada. Salvo algunos sectores meridionales, la mayor parte del territorio se articula dentro de la vertiente cantábrica.

Desde el punto de vista geológico, se inserta dentro de la macrounidad de los Pirineos y, más concretamente, forma parte de los Montes Vascos o Cordillera Vasco-cantábrica, los cuales están conformados a su vez por las morfoestructuras del Anticlinorio de Bizkaia, el Sinclinorio de Bizkaia y la Franja Alóctona de Lekeitio-Tolosa o Anticlinal Norte de Bizkaia, las cuales presentan un trazado general noroeste-sureste. (Ruiz Urrestarazu y Galdos Urrutia, 2008).

Desde el punto de vista fisográfico, los Montes Vascos se caracterizan por presentar relieves de cabecera típicos de una zona de montaña media, con alturas entre los 1000 y 1500 metros, con escasa masividad de las estructuras (< 500 km²), con relieves discontinuos y marcados por la proximidad al litoral. Destacan las morfoestructuras calizas que constituyen los relieves dominantes en cabecera y en los interfluvios, con presencia abundante de elementos típicos del modelado kárstico, tanto a nivel del exokarst como del endokarst.

El resultado de esta configuración morfoestructural son una serie de cuencas hidrográficas alargadas, drenadas por cursos de perfil longitudinal marcadamente fluvio-torrencial (con un 2-3 % de pendiente media, aunque en las cabeceras se rebasa el 50), pequeñas (< 1000 km², excepto la del Nerbion-Ibaizabal), con un trazado dendrítico, alimentación pluvial y discordantes con la estructura.

Los valles fluviales presentan perfiles transversales en V, con relleno exiguo de fondo de valle. Todos los cursos de agua importantes se integran en el litoral mediante estuarios de tipo ría, parcialmente colmatados por sedimentos Holocenos. (González y Serrano, 1996).

Bizkaia presenta un clima oceánico litoral (litoral fresco o de fachada occidental), abierto a las influencias marítimas. Este tipo climático se caracteriza por su marcada homogeneidad térmica -temperaturas moderadas en invierno y suaves en verano- y por sus elevadas precipitaciones.

Los valores pluviométricos son superiores a los 1.400-1.500 mm. anuales, pudiendo alcanzarse incluso cifras por encima de los 2.000 mm. Estas se incrementan gradualmente de oeste-este a este y de norte a sur, aunque es el relieve el elemento fundamental que explica su distribución. Los valores más bajos, -1.000-1.200 mm-, se producen en la ría de Gernika y en el valle del Nerbion-Ibaizabal (1.115 mm en Basauri, 1.022,5 mm en Punta Galea), incrementándose a medida que aumenta la altitud y a medida que nos desplazamos hacia el oeste.

Además de elevadas precipitaciones, esta variedad climática oceánica se caracteriza por presentar una buena distribución de las precipitaciones a lo largo del año, con un máximo otoño-invernal muy nítido y un máximo secundario en primavera (normalmente en abril o en marzo). Diciembre suele ser el mes más húmedo (en algunos observatorios es noviembre), mientras que el más seco es, en todos los casos, julio. (Ruiz Urrestarazu, 1982; Ruiz Urrestarazu y Galdos Urrutia, 1989).

Los días de tormenta, que suponen alrededor de 13 días de media al año, siendo más frecuentes en entre los meses de abril y octubre son especialmente peligrosos ya que descargan importantes volúmenes de agua en cortos lapsos de tiempo, lo que unido a las características de la red hidrográfica y la estrechez de los valles cantábricos hacen que el riesgo de inundaciones sea elevado.

Desde el punto de vista térmico se caracteriza por sus moderadas temperaturas, hasta tal punto que no puede hablarse de período frío, mientras que las de verano resultan suaves. La elevada nubosidad mitiga la pérdida de calor por irradiación en invierno y evita el calentamiento excesivo en verano, acompañado por el papel atemperante del mar que suaviza los rigores del invierno y aminora los excesos del verano, consecuencia de todo ello la oscilación térmica anual resulta muy baja: 10,9 en Sondika. La temperatura media anual oscila entre los 14,4° C. de Punta Galea, y los 12,7° C de Carranza; ningún mes baja de los 5° C. Las únicas excepciones se dan en el interior del

territorio o en zonas con alturas por encima de los 700-800 metros. (Uriarte, 1996)

El riesgo de heladas es bajo: 13,1 días de media anual, con enero (3,3 días), el mes que presenta un mayor riesgo. Entre los meses de mayo y septiembre son fenómenos excepcionales, especialmente en la costa.

Respecto al régimen de vientos, la máxima frecuencia corresponde a los vientos de componente noroeste (24,2%), seguidos por los de componente S (18,7%) y norte (15,57%). También son importantes los vientos del oeste (8,5%), mientras que el resto de las direcciones apenas tienen incidencia. El porcentaje de calmas supone el 21% de las observaciones, y el de vientos flojos representa el 62,7%, lo que permite afirmar que predominan los tipos de tiempo poco ventosos. Sin embargo, hay que destacar la importancia de los episodios de vientos muy fuertes (fuerza 5 o más), e incluso de los muy fuertes y atemporalados, que, aunque representan tan solo el 4,8% de las observaciones, a menudo traen consigo afecciones a infraestructuras y a los sistemas forestales. (Uriarte, 1996)

Principales formaciones vegetales

El bosque potencial de Bizkaia es el bosque caducifolio, compuesto sobre todo por especies frondosas como el haya (*Fagus sylvatica*) a partir de los 600 metros, y el roble pedunculado (*Quercus robur*), el roble albar (*Q. petraea*) y el roble negro (*Q. pyrenaica*) en el piso inmediatamente inferior. Otras especies que acompañan a menudo son los castaños (*Castanea sativa*), fresnos (*Fraxinus excelsior*),

abedules (*Betula celtiberica*), arces (*Acer campestre*), olmos (*Ulmus minor*), tejos (*Taxus baccata*), y acebos (*Ilex aquifolium*) entre otros. Sin embargo, la fuerte presión antrópica (presión demográfica, plantaciones de coníferas y mirtáceas como el eucalipto, la especialización ganadera, la ocupación urbana industrial y viaria, etcétera.) ha propiciado una progresiva degradación de este tipo de bosque.

Las manchas forestales supervivientes suelen presentar un aspecto degradado, a menudo mostrando las etapas de sustitución, observándose secuencias regresivas de orlas espinosas, landas acidófilas de helechos (*Pteridium aquilinum*), brezos y brecinas (*Erica vagans* y *E. cinerea*) y argomas (*Ulex europaeus* y *U. gallii*) al final de la serie. (Meaza, G., 1997)

La repoblación con especies alóctonas como el pino insigne ha contribuido a menguar la extensión de este bosque potencial debido en gran parte a su rápido crecimiento y su adaptabilidad al terreno y al clima de buena parte del territorio vizcaíno, lo que le ha convertido en el indiscutible protagonista del paisaje forestal de los interfluvios de la provincia. La campiña está ocupada por prados de siega y diente, así como cultivos de forrajeras como la alfalfa, el trébol, etcétera. para la alimentación del ganado, coexistiendo con otros cultivos como el de las hortalizas, maíz, y árboles frutales como el manzano. (Aseguinolaza et al., 1989; Meaza, 1996)

5.1. El pino radiata en la CAE y en Bizkaia

Los pinares de *P. radiata* son la formación más importante en superficie del País Vasco, ocupan 125.000 hectáreas que suponen el 32% de la superficie arbolada autonómica y, a nivel nacional, suponen el 49% de la superficie de esta formación arbolada. (Ministerio de Agricultura, Alimentación y Medio Ambiente, 2013) Su importancia también queda reflejada en las cifras de existencias: estos pinares aportan el 23% de los pies mayores de la comunidad y el 44% del volumen con corteza. Es por tanto la especie que más aparece representada entre las parcelas medidas en el IFN4, (507 parcelas, 30% del total).

Los bosques de pino radiata, repoblaciones de carácter productor, se presentan en altitudes inferiores a los 600 metros, por todo el territorio de Gipuzkoa y Bizkaia, y en el extremo noroccidental de Araba/Álava (coincidiendo con los terrenos de menor cota de esta provincia).

Según los últimos Inventarios Forestales, la superficie ocupada por el pino radiata (Figura 17). Sin embargo, en cuanto a existencias maderables nos encontramos en máximos, según el último Inventario Forestal de 2011 que cuantifica unas existencias maderables totales de 30 millones de m³ de pino radiata y un promedio de 216 m³/ha. En el primer Inventario Forestal de 1972 se midieron 15 millones de m³ y un promedio de 94 m³/ha.

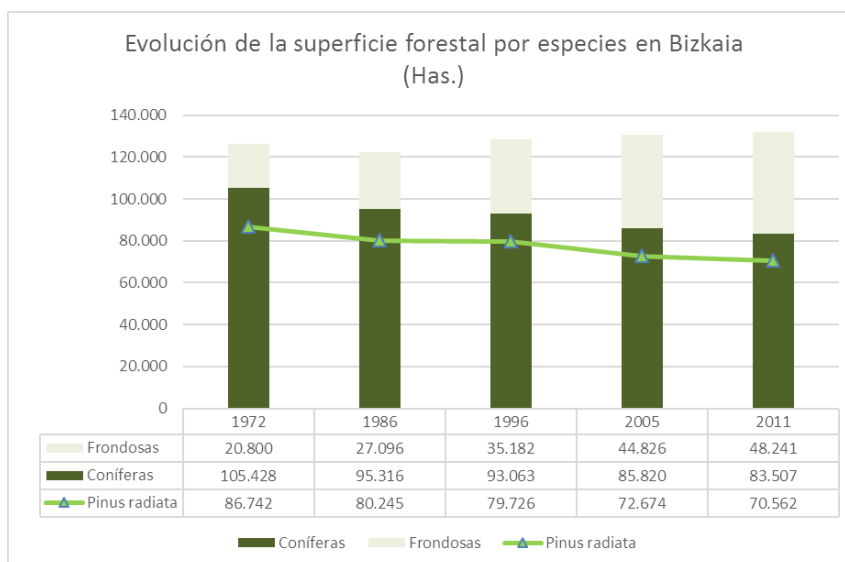


Figura 17 Evolución de la superficie forestal en Bizkaia Fuente: Inventarios Forestales ⁴

La superficie total ocupada por coníferas en los montes vascos ha pasado de suponer en 1972 más del 62% del total de la superficie arbolada a ser menos del 50% en 2011. Sin embargo, más del 60% del total de las existencias maderables se sigue concentrando en masas de pino (tabla 7).

⁴ Los conceptos "Total coníferas" y "Total frondosas" incluyen una superficie correspondiente al llamado "Bosque mixto" (coníferas + frondosas) en los primeros mapas (1972 y 1986). En estos primeros mapas no se diferenciaban algunas especies forestales que entonces ocupaban escasa superficie. Aparecían englobadas en el grupo de "Otras coníferas" o de "Otras frondosas". Tampoco existía el concepto de "Bosque mixto atlántico", que empezó a emplearse en 1996 para englobar a mezclas de diversas frondosas de regeneración natural. Algunas de estas superficies, publicadas en el País Vasco, pueden no coincidir exactamente con las publicadas por el Ministerio, ya que se basan en el método de agrupación por estratos. En el último Inventario Forestal de la CAPV, con trabajo de campo realizado en 2011, se incluyen los resultados del mapa forestal 2010.

Esto coincide con un importante bache en la demanda de madera. Se calcula que el crecimiento anual de estos pinares alcanza los 1,8 millones de m³, frente a una demanda actual de las industrias comprendida entre 1 y 1,5 millones de m³ cada año (Cantero, 2014).

La madera de pino radiata, que hasta hace pocos años era la única especie que aportaba volúmenes importantes de aprovechamiento forestal, coexiste con la oferta procedente de otras masas forestales que van aumentando su superficie año a año. Por ejemplo, los eucaliptos (15.000 ha), el abeto Douglas (6.500 ha) o el pino laricio (13.700 ha). En ciertas zonas costeras y en montes de suelos pobres o arenosos el pino radiata ha ido dejando sitio a los eucaliptos. Cuando la altitud se torna factor limitante, en cotas superior a los 500 m, suele plantarse abeto Douglas en suelos profundos y fértiles o pino laricio de Córcega en suelos más pobres o con mayor riesgo de viento o nieve.


Los principales problemas para movilizar esta madera son de tipo socio-económico, relacionados básicamente con el minifundismo⁵, el coste de la mano de obra, el abandono de la gestión forestal por edad o por cansancio o las restricciones legales para la corta y la saca. En ocasiones, también existen limitaciones físicas en muchos montes vascos, como la ausencia de una red viaria en

⁵ Sólo el 15% de la superficie de pino radiata se ubica en montes públicos frente a los aproximadamente de 20.000 propietarios privados, lo que da una superficie media de 5,6 ha de pinares por propietario; en la mayor parte de las ocasiones, además, repartidas en varias parcelas no colindantes.

condiciones, fuertes pendientes medias, riesgos erosivos o climatología adversa.

En este contexto, la obtención de bioenergía es una de las posibles salidas para la madera, la leña u otro tipo de materia vegetal de los montes vascos. Muchas entidades locales vascas son propietarias de montes y podrían ser autónomas en materia energética con unas pequeñas inversiones iniciales.

Tabla 7 Variación de superficies y densidades de *P. radiata*

Superficie (ha)			
	IFN3	IFN4	Variación (%)
Pinares con F.c.c. entre 70 y 100%	83.067,23	84.849,54	2,15
Pinares con F.c.c. < 70%	11.292,46	16.722,39	48,08
Replantaciones jóvenes de pinares	34.701,16	23.641,88	-31,87
 Total pinares de <i>Pinus radiata</i>	129.060,85	125.213,81	-2,98

Densidad de la formación (existencias por hectárea)		
	IFN3	IFN4
Pies mayores (pies/ha)	455,76	451,95
VCC (m ³ /ha)	184,99	221,13
Pies menores (pies/ha)	361,27	328,92

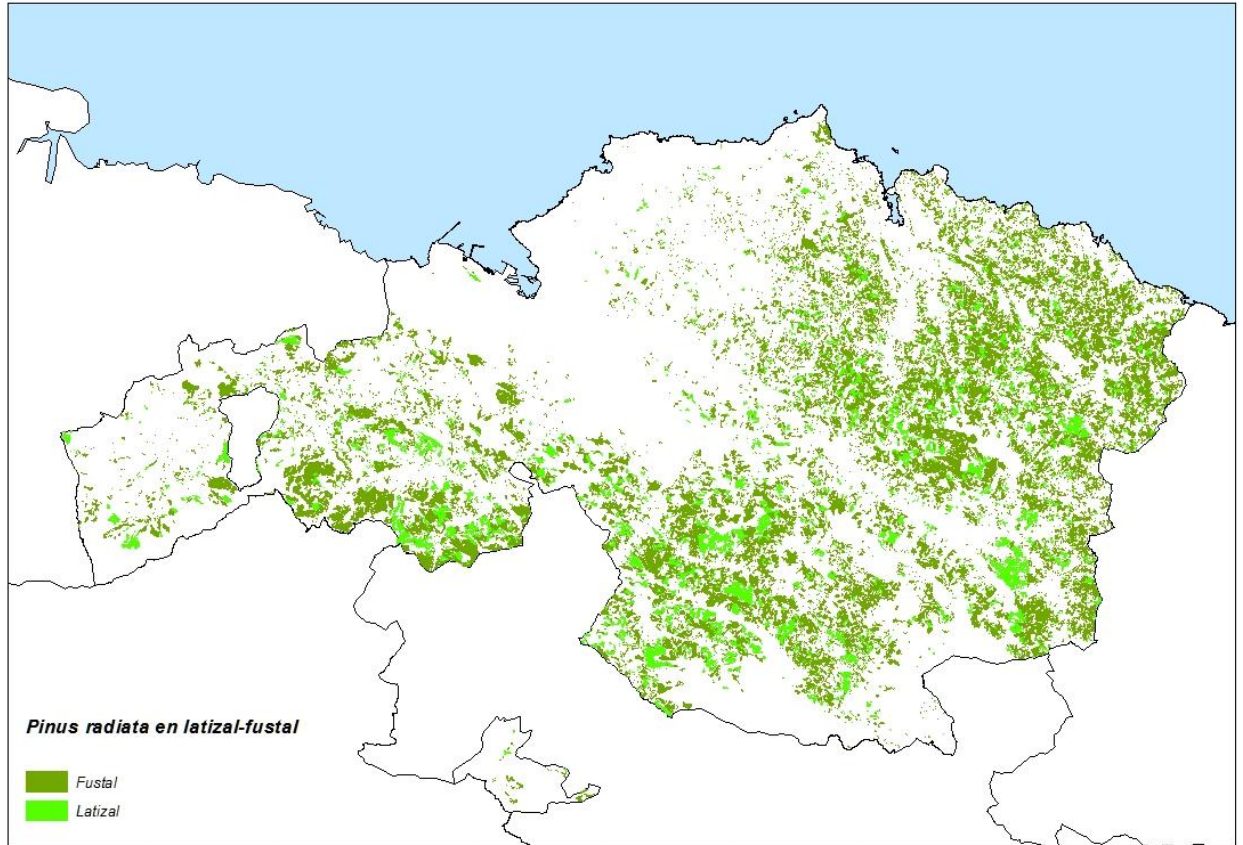


Figura 18 Distribución del *Pinus radiata* en formaciones de Latizal y Fustal en Bizkaia (Elaboración propia a partir del IFN4)

CAPITULO 2

DESARROLLO DE MODELOS Y ANÁLISIS DE RESULTADOS DE ESTIMACIÓN

6. OBTENCIÓN DE LAS VARIABLES DE LOS MODELOS DE ESTIMACIÓN

En este apartado se explican los procedimientos seguidos para la obtención de las variables que formarán parte de los procesos estadísticos conducentes a obtener los modelos de predicción de la biomasa aérea de *P. radiata* en masas de latizal/fustal en Bizkaia.

6.1. Tratamiento de los datos de campo de la biomasa forestal

Los datos de campo a partir de los cuales se generará la variable independiente de los modelos de predicción, es decir la biomasa, provienen de una única fuente: el cuarto Inventario Forestal Nacional. Como se ha comentado, la estimación de la biomasa no se realiza de manera directa en la fase de campo, sino que hará falta poner en relación algunas de las variables dasométricas obtenidas en las parcelas muestrales con la biomasa por unidad de superficie mediante una serie de ecuaciones denominadas ecuaciones alométricas (Canga et al., 2013; Montero González et al., 2006; Pérez-Cruzado, 2008).

6.1.1. El Inventario Nacional Forestal

Diseño del muestreo

Una vez elegidos los parámetros que van a ser objeto de investigación, todos los del inventario anterior más un grupo de nuevos, se seleccionan un conjunto de parcelas de muestreo que, al tratarse de un inventario continuo, se preparan para que sean permanentes y puedan visitarse sucesivas veces.

El diseño de muestreo se basa en la estratificación de las superficies forestales en función de sus características dasométricas, estructurales y botánicas o, en algunos casos, de condicionamientos geográficos o institucionales, pero tomando siempre la provincia como unidad para los cálculos.

Cada estrato se forma por agrupación de superficies forestales (llamadas teselas, manchas, polígonos o recintos) de características similares, cuyo perímetro esté delimitado a la escala de trabajo (normalmente 1:25.000).

Al ser un inventario forestal continuo, la muestra del IFN4 viene condicionada por las del Segundo y Tercer Inventario Forestal Nacional (IFN2 e IFN3). Como norma general se repiten todas las parcelas usadas en el proceso de datos del inventario anterior, añadiéndose las correspondientes a los nuevos terrenos arbolados y excluyéndose, para los cálculos de existencias que no para los de comparación, las de aquéllos que han pasado a no

forestal arbolado en los últimos 10 años (Ministerio de Medio Ambiente, 2014).

Además, se selecciona otra muestra de parcelas en las zonas de monte desarbolado distribuidas de la misma manera que sus homólogas arboladas (puntos de cruce de las líneas UTM kilométricas). Esta muestra se apea total, parcialmente o no se apea según disponibilidades presupuestarias, al ser consideradas de menor interés que las arboladas. No se consideran los pocos árboles encontrados en estas parcelas pues, además de no tener apenas importancia en el total provincial, las cifras serían muy imprecisas por el elevado error de muestreo que conllevan las anteriores circunstancias. Por este motivo, la información que se presenta es más escasa que la de la zona arbolada y hay que utilizarla con prudencia al ser mayor el error de muestreo y, quizá, pudiera tener algún ligero sesgo.

La información suministrada se separa, dentro de los tres primeros ámbitos, en la correspondiente a las zonas de uso forestal monte arbolado de las de uso forestal monte desarbolado. De esta segunda parte se han seleccionado las tablas con la información considerada como más interesante.

Una vez seleccionados los puntos de muestreo UTM, que se convertirán en parcelas, el paso siguiente consiste en transferirlos a las más modernas y adecuadas fotografías aéreas que sea posible conseguir. Dichas fotografías pasan al jefe de los trabajos de campo que, después de la pertinente planificación, las reparte a los jefes de

equipo. Estos, ayudados por los planos generales de la zona, por las informaciones disponibles del inventario anterior y por las fotos, se desplazan en un vehículo todoterreno hasta las cercanías del punto objetivo. Después, ya a pie y mediante la observación estereoscópica de los pares de contactos, y los croquis de acceso del IFN3 cuando se trate de parcelas de repetición, recorren el correspondiente terreno identificado hasta llegar a la parcela.

El emplazamiento del punto centro de la parcela, determinado mediante una decisión aleatoria sin que influyan en ella la topografía, vegetación, gustos del operador, etcétera., se formaliza clavando un rejón metálico tubular, que queda oculto al marcharse el personal. En las parcelas de repetición este rejón ya está en el terreno y se localiza con la ayuda adicional del rumbo y la distancia de los árboles medidos en el IFN3, que se identifican ahora, y de un detector de metales. Si por alguna razón es imposible encontrar el rejón de estas parcelas se actúa como si fueran nuevas (Ministerio de Medio Ambiente, 2014; Tracsatec).

En algunas comunidades autónomas se están utilizando ortofotos recientes en vez de fotografías aéreas. Este método tiene la ventaja de la disminución de costes -se pueden utilizar vuelos modernos existentes de cualquier escala para hacer ortofotos- y de la mayor precisión de los puntos UTM ubicados en los soportes gráficos, y el inconveniente de la pérdida de la visión estereoscópica,

paliada en parte con los fotogramas antiguos del IFN3 que se llevan también al campo.


Luego de ubicarse el jefe de equipo en el centro de la parcela, comienza la medición de los parámetros seleccionados para este inventario siguiendo las normas escritas en un manual de instrucciones de los trabajos de campo creado al efecto. El tamaño de la muestra (superficie a muestrear) es variable en función del diámetro de los individuos (parcelas de radio múltiple).

Existen doce bloques de información en el registro de cada parcela:

- El primero trata de la identificación del punto (provincia, hoja, fotograma, número, clase, etcétera.).
- El segundo de su clasificación según el fotointerpretador y según el operador de campo con respecto a una serie de parámetros (nivel de usos del suelo, nivel morfoestructural, especie, ocupación, estado, etcétera.).
- El tercero de la estructura de edades y del origen de la vegetación leñosa.
- El cuarto, el más importante de todos, estudia la dendrometría de los árboles empezando por su localización en coordenadas polares y su numeración, diferente si son árboles nuevos que, si son los mismos que se midieron en el IFN3, siguiendo por la especie botánica, dos diámetros normales y la altura total, y acabando por la calidad, forma de cubicación, estado

fitosanitario y parámetros especiales. Esto se completa con datos sobre regeneración, matorral leñoso y especies forestales arbóreas presentes.

- El quinto analiza el suelo a través de su rocosidad, su textura, su contenido en materia orgánica, su pH y su tipo.
- El sexto examina los riesgos bien de erosión, bien de incendios.
- El séptimo, selvicultura, registra los tipos de cortas de regeneración y los tratamientos de mejora del vuelo y del suelo.
- El octavo mide y dibuja la fisiografía de la parcela.
- El noveno describe el itinerario de acceso y referencia mediante texto y gráficos.
- El décimo recoge las observaciones que sobre el apeo quiera hacer el operador.
- El undécimo anota los datos de control (jefe de equipo, hora de comienzo, etcétera.).
- El duodécimo explica cómo se han hecho las fotografías en la parcela que sirven para caracterizar el paisaje y visualizar la estructura de la masa.



CUARTO INVENTARIO FORESTAL NACIONAL
ESTADILLO DE DATOS DE CAMPO
FORESTAL ARBOLADO

N.º ESTADILLO
 CLASE
 TIPO

1. IDENTIFICACIÓN

PROVINCIA HOJA VUELO PASADA FOTO COORDENADAS:
 AÑO VUELO PASADA FOTO X Y

2. CLASIFICACIÓN

NIVEL DE USOS DEL SUELO Nivel1 NIVEL MORFO-ESTRUCTURAL Nivel2 Nivel3 FRACCIÓN DE CABIDA CUBIERTA TOTAL DE LA VEGETACIÓN FRACCIÓN DE CABIDA CUBIERTA DE LA VEGETACIÓN ARBÓREA DISTRIBUCIÓN ESPACIAL COMPOSICIÓN ESPECÍFICA

ESPECIE OCUPACIÓN ESTADO ESPECIE OCUPACIÓN ESTADO ESPECIE OCUPACIÓN ESTADO

3. ESTRUCTURA DE EDADES Y ORIGEN

ESPECIE FORMA PRINCIPAL DE MASA EDAD FUENTE DE LA INFORMACIÓN FIABILIDAD ORIGEN DE LA MASA TRATAMIENTO DE LA MASA

Figura 19 Estadillo de levantamiento de parcelas del IFN4

5. SUELO

POCOSIDAD

TEXTURA

CONTENIDO EN MATERIA ORGÁNICA

REACCIÓN DEL SUELO pH

TIPO DE SUELO

6. RIESGOS

EROSIÓN

MANIFESTACIONES EROSIVAS

INCENDIOS

MODELO DE COMBUSTIBLE

ESPESOR DE LA CAPA MUERTA, CÉSPED, MUSGO Y LÍQUENES

PRESENCIA DE REGENERACIÓN

EFFECTIVIDAD DE LA REGENERACIÓN

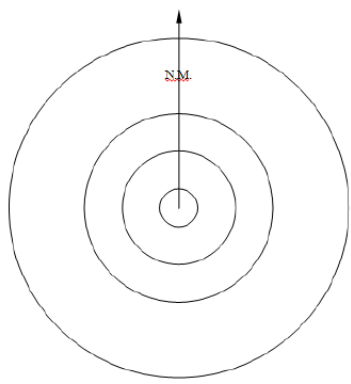
7. SELVICULTURA

EXISTENCIA DE CORTAS

TRATAMIENTO DE MEJORA DEL VUELO

TRATAMIENTO DE MEJORA DEL SUELO

8. FISIOGRAFÍA DE LA PARCELA



CONTINUACIÓN DE LAS OBSERVACIONES

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

9. ITINERARIO DE REFERENCIA		
<u>REFERENCIA</u>		<u>CROQUIS</u>
DESCRIPCIÓN		
.....		
.....		
.....		
LOCALIZACIÓN <input type="checkbox"/>	ACCESO <input type="checkbox"/>	LEVANTAMIENTO <input type="checkbox"/>
10. OBSERVACIONES		11. DATOS DE CONTROL
.....		JEFE DE EQUIPO
.....		FECHA
.....		HORA DE COMIENZO
.....		HORA DE FINALIZACIÓN
.....		
12. FOTOGRAFÍAS DE LA PARCELA		
ARCHIVO: <input type="checkbox"/>	FOTO 1: <input type="checkbox"/>	RUMBO: <input type="checkbox"/>
FOTO 2: <input type="checkbox"/>	RUMBO: <input type="checkbox"/>	DISTANCIA AL CENTRO DE LA PARCELA <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
		FOTO 3: <input type="checkbox"/>

Figura 20 Estadillo de levantamiento de parcelas del IFN4

La recogida de esta gran cantidad de datos requiere personal especializado, aparatos de medición y registro complejos y un manual de instrucciones extenso y minucioso. Las mensuras y las explicaciones, menos algunos dibujos complicados, se registran en un ordenador portátil, y un software hecho específicamente para este proyecto. (Ministerio de Medio Ambiente, 2014).

6.1.2. El cuarto inventario forestal nacional en el País Vasco

El Inventario Forestal CAPV 2011 es una estadística oficial, integrada en el Plan Vasco de Estadística 2009-2012. Su realización es responsabilidad del Órgano

Estadístico del Departamento de Medio Ambiente, Planificación Territorial, Agricultura y Pesca del Gobierno Vasco, y cuenta con la colaboración financiera y operativa de las Diputaciones Forales de Álava, Bizkaia y Gipuzkoa. Este Inventario, se ha desarrollado en el marco metodológico del Inventario Forestal Nacional IFN4, y se ha realizado en colaboración con el Ministerio de Agricultura, Alimentación y Medio Ambiente, responsable del IFN4, que ha aportado soporte técnico y financiero para, el levantamiento de las parcelas de campo y su procesado, la valoración económica, el estudio de biodiversidad y la evolución del paisaje forestal.

Previo a la ejecución del inventario se realizó durante los años 2010 y 2011 el mapa forestal de la CAPV. Este documento fue la base para la determinación de los estratos inventariables y la colocación de las parcelas de inventario.

El mapa forestal se elaboró por fotointerpretación sobre pantalla utilizando las ortofotos en color del vuelo del año 2009 con tamaño de pixel de 0,25 metros. El tamaño mínimo de tesela cartografiable fue de 0,20 hectáreas, de forma que una tesela pudiera albergar una parcela de inventario de 25 metros de radio.

Las fases que se siguieron en el trabajo de fotointerpretación fueron:

Recogida de la información: las Diputaciones Forales enviaron datos catastrales de las parcelas en las que se habían registrado peticiones de corta de arbolado o ayudas

por reforestación. También se contaba con los mapas de rodales actualizados de montes públicos. De esta forma, se contaba con una referencia geográfica de la ubicación de cortas y plantaciones forestales en los últimos años.

Transformaciones temporales: dado el plazo de 5 años entre el anterior mapa y el nuevo, en el caso de las plantaciones forestales se supuso que evolucionaron con la edad; así, los arbolados en estado de masa repoblado en 2005 han pasado 5 años después al estado monte bravo, los arbolados en estado de masa monte bravo en 2005 han pasado 5 años después al estado latizal y los arbolados en estado de masa latizal en 2005 han pasado 5 años después al estado fustal. Las zonas que en 2005 aparecían como Talas y no han sido plantadas ni regeneradas en estos 5 años, dejan de considerarse Talas para pasar a ser Matorral o Herbazal en cada caso.

Teselado: tomando como base el anterior mapa forestal 2004-2005, se fotointerpretaron directamente en pantalla los cambios detectados en la vegetación sobre ortofoto en color con pixel de 0,25 metros.

Codificación: las teselas, visualizadas en pantalla, fueron codificadas siguiendo las normas del mapa forestal MFE25, incluido en el IFN4. Con vistas a que los fotointerpretadores tuvieran mejores referencias del tipo de vegetación de cada tesela, se contó con la información aportada por anteriores ortofotos anuales, cartografía de vegetación Eunis 2005 y las propias parcelas de campo del Inventario forestal 2005 (incluido en el IFN3)

Corrección de errores: una vez codificadas las teselas de las distintas hojas, se comprobó que no existieran teselas con campos vacíos o con valores erróneos y que no hubiera bordes defectuosos en las teselas, concentrados sobre todo en los límites entre hojas.

Se siguió el límite oficial de la CAPV y de cada Territorio Histórico aportado por el Ministerio de Medio Ambiente (MMA), aunque se amplió esta superficie cartografiada del País Vasco en una distancia de 1 km para evitar vacíos de información en los bordes.

En los Territorios Históricos de Bizkaia y Gipuzkoa, las Diputaciones Forales aportaron cartografía digitalizada de los rodales de diversos montes públicos, a escala aproximada de trabajo 1/5.000. En el caso de Álava, donde no se dispone de rodalización de los montes públicos, la guardería forestal realizó el trabajo de corrección y actualización del mapa de vegetación.

Para la ejecución del inventario, cada parcela de campo es asignada a un estrato según la información del mapa forestal. Los estratos se definen en función de la especie forestal, densidad del arbolado y estado de masa o rango de edad. Se busca en la formación de estratos la homogeneidad de las teselas que lo componen en cuanto a la variable principal a estimar en el inventario: volumen de madera.

Tabla 8 Estratos forestales del IFN4

Estrato	DEFINICIÓN			
	Formación forestal dominante	Ocupación %	Estado de masa	FCC%
01	<i>P. sylvestris</i>	>=70	Fustal. Latizal	70-100
02	<i>P. sylvestris</i>	>=70	Fustal. Latizal	5-69
03	<i>P. sylvestris</i> y <i>P. nigra</i> de plantación	>=70	Fustal. Latizal	5-100
04	<i>P. nigra</i> y <i>P. pinaster</i>	>=70	Fustal. Latizal	5-100
05	<i>P. nigra</i> y <i>P. nigra</i> con <i>Q. faginea</i> o <i>Q. Ilex</i>	>-70;30<=Esp.<70	M. bravo-Repoblado	5-100
06	<i>P. radiata</i>	>=70	Fustal. Latizal	20-100
07	<i>P. radiata</i>	>=70	M. bravo-Repoblado	5-100
08	<i>Q. robur</i> y <i>Q. robur</i> con <i>Fagus sylvatica</i>	>=70;30<=Esp.<70	Fustal. Latizal	5-100
09	<i>Quercus pyrenaica</i>	>=70	Fustal. Latizal	5-100
10	<i>Quercus faginea</i>	>=70	Fustal. Latizal	70-100
11	<i>Q. faginea</i> con <i>P. sylvestris</i> o con g.	30<=Esp.<70	Fustal. Latizal	70-100
12	<i>Quercus faginea</i>	>70	Fustal. Latizal	5-69
13	<i>Quercus ilex</i>	>=70	Fustal. Latizal	70-100
14	<i>Quercus ilex</i>	>=70	M. bravo-Repoblado	5-69
15	<i>Q.faginea</i> at <i>Q. ilex</i> y <i>Q. pyrenaica</i>	>-70;30<=Esp.<70	Fustal. Latizal	5-100
16	<i>Fagus sylvatica</i>	>=70	Fustal. Latizal	70-100
17	<i>Fagus sylvatica</i>	>=70	M. bravo-Repoblado	5-100
18	Arboles de ribera y frondosas región atlántica	>=70;30<=Esp.<70	Todos	5-100

Los estratos se materializan de forma que tengan una superficie mínima que asegure una suficiente representación de parcelas de campo que evite que los errores de muestreo sean muy elevados, la búsqueda de esta superficie mínima lleva en ocasiones a agrupar en un estrato teselas de diferentes especies. Se definieron 18 estratos en Álava, 12 en Bizkaia y 14 en Gipuzkoa.

Las parcelas de inventario se situaron en los vértices de la malla kilométrica de coordenadas UTM en el sistema de referencia ED50, cuando éstos se encontraban dentro de una tesela de las definidas en el Mapa Forestal. En esta edición del inventario y por motivos presupuestarios, sólo se midió una de cada dos parcelas que cumplían las

condiciones anteriores por lo que sólo se cuenta con la mitad de la información de campo que en ediciones anteriores del IFN. En esta edición se midieron 1.685 parcelas.

Para el replanteo de parcelas y medida de variables se siguió, con algunas modificaciones, la metodología desarrollada para el IFN2 (ICONA, 1990).

Es muy importante tener en cuenta que, según esta metodología de medición, se replantean en el campo parcelas de radio variable de forma que un árbol entrará o no en la parcela según su diámetro normal. Los árboles con diámetro inferior a 7,5 centímetros se consideran pies menores y no se tienen en cuenta para el cálculo de las variables derivadas. Los pies cuyo diámetro se encuentra entre 7,5 y 12,5 centímetros sólo se cuentan si se encuentran a menos de 5 metros del centro de la parcela. Los pies con diámetro entre 12,5 y 22,5 centímetros se cuentan si su distancia al centro de la parcela es menor de 10 metros. Aquellos árboles cuyo diámetro normal esté entre 22,5 y 42,5 cm se tendrán en cuenta si su distancia al centro de la parcela es menor de 15 metros. Los arboles de más de 42,5 centímetros de diámetro se contarán todos dentro de una parcela de 25 metros de radio.

El peso de cada árbol en el cálculo de variables derivadas depende pues de su diámetro según el denominado factor de expansión que se define como la relación que existe entre la superficie de referencia (en general una hectárea) y la

superficie de la subparcela cuyo diámetro corresponde a las condiciones expuestas más arriba.

El uso de estos factores no ofrece problemas si se pretende estimar existencias a una escala adecuada, pero es problemático si se pretenden estimar distribuciones diamétricas de parcelas reales (Bravo, del Río, Montero, Ordoñez y Cañellas, 2002).

Una vez obtenidos los datos de campo, la densidad en número de pies por hectárea de cada parcela se obtiene sumando la densidad correspondiente a cada pie muestreado. El área basimétrica por hectárea de cada parcela se obtiene sumando la sección normal correspondiente a cada árbol muestreado teniendo en cuenta que, en función de su diámetro, cada árbol muestreado representa a un número determinado de pies por hectárea según su factor de expansión.

Se aplican las ecuaciones de cubicación a cada árbol muestreado en las distintas parcelas de campo para obtener el volumen y el crecimiento anual de cada árbol y, por extensión, de cada parcela multiplicando el volumen de cada árbol por el factor de expansión que le corresponde.

Para elevar los datos a nivel de provincia, se toman todas las parcelas de un mismo estrato y se calculan con todas ellas los valores medios por hectárea de los distintos parámetros, estos valores constituyen los parámetros dasométricos. Estos valores medios se multiplican por la superficie del estrato y se obtienen los valores absolutos del estrato. La agregación de los valores absolutos de todos los estratos proporciona los datos de número de pies,

volúmenes y crecimientos desglosados por especies en cada provincia.

Para este estudio se utilizarán los datos de *P. radiata* (pino insigne) obtenidos en el IFN4 de la CAPV. Sobre esta especie se desarrolla todo el trabajo posterior.

6.1.3. Los datos de campo del IFN4 en Bizkaia

Los datos de campo de este estudio corresponden a los tomados en Bizkaia entre el 17 de enero y el 15 de junio de 2011 para el Cuarto Inventario Forestal Nacional (IFN4). En total 581 parcelas de inventario fueron replanteadas y medidas en Bizkaia.

Hay que tener en cuenta, sin embargo, que el rápido crecimiento de esta especie provoca que las diferencias entre los valores de las variables dasométricas obtenidas con los datos de campo procedentes del IFN4 de 2011 y los obtenidos con la aplicación del procedimiento desarrollado para los datos LiDAR obtenidos en 2008 complican su uso. A esto hay que sumarle el hecho de que el pino radiata está sometido a un régimen de cortas muy intenso por lo que en los tres años de desfase muchas de las parcelas medidas han sufrido reducciones considerables en el número de pies que contienen. Aunque la variación en superficie respecto al IFN3 es mínima (tabla 8), la división de cabidas en función de la F.c.c. y la evolución de las existencias indican menor superficie de repoblaciones jóvenes en favor de masas más maduras, con mayor densidad de volumen.

Para solventar este problema se ha recurrido al contraste mediante fotointerpretación de las ortofotos de ambos años (2008 y 2011) y se han descartado aquellas parcelas que evidenciaban una pérdida considerable de individuos.

De las parcelas medidas en Bizkaia se separaron, en primera instancia, las 182 que tenían como especie principal el pino radiata, cuyo estado de desarrollo de la masa se correspondía con el de:

1. **Latizal:** fase que se encuentra entre el momento en el que las ramas se pueden encontrar a lo largo de todo el fuste y el momento en el que los pies tienen unos 20 cm de diámetro normal.
2. **Fustal:** Se caracteriza esta clase de edad porque sus pies tienen diámetros normales superiores a 20 cm llamándose bajo o joven fustal a los pies con diámetros normales comprendidos entre 20 y 35 cm, fustal medio a aquellos que los tienen entre 35 y 50 cm y fustal viejo o alto fustal, cuando los diámetros normales son superiores a 50 cm. en las que, además, según el IFN4 la composición específica, en función de las especies presentes en la parcela apeada, era:
 1. **Homogénea o puras,** es decir a masas monoespecíficas con una única especie arbórea o a masas masas heterogéneas o mezcladas pie a pie. La normativa española precisa que una masa es monoespecífica o pura cuando al menos el 90% de los pies pertenecen a la misma especie.

2. Heterogénea o mezclada pie a pie. Masas de diferentes especies que se juntan o bien se entremezclan por golpes o grupos, siempre que tengan una altura similar).

De las 182 parcelas del IFN4 se eliminaron aquellas cuya altura media no superaba los 4 m para evitar el efecto distorsionador que puede tener el matorral sobre los datos finales, además de tener en cuenta el crecimiento en altura de los árboles en los tres años de desfase entre la toma de datos LiDAR y los datos de campo.

Como se ha señalado, los datos LiDAR usados en el proceso se tomaron tres años antes que las medidas en campo. Tras comparar un a uno los árboles inventariados en el IFN3 y el IFN4 se ha observado que el crecimiento en altura de los pinos en estos años ha sido de media 0,33 metros por año. (González 2006). Por esto, para este estudio, se consideraron y contabilizaron para los cálculos de la biomasa a nivel de parcela sólo los pies cuya altura fuera mayor de 6 metros. Se evitaba así que los árboles incluidos invalidaran la premisa de altura mínima de trabajo de 4 m.

Esto ha sido posible gracias a que, en ambos inventarios, durante la toma de datos de campo, se midió el rumbo en grados centesimales con aproximación de ± 1 grado de cada árbol desde el centro de cada parcela. También se midió la distancia en metros a cada árbol desde el centro de la parcela con aproximación de $\pm 0,1$ metros. De esta forma se obtuvieron las coordenadas polares de cada árbol respecto al centro de la parcela y se pudieron identificar aquellos

pies apeados en ambos inventarios. Así aparecen recogidos en el estadillo levantado para cada parcela (Fig. X)

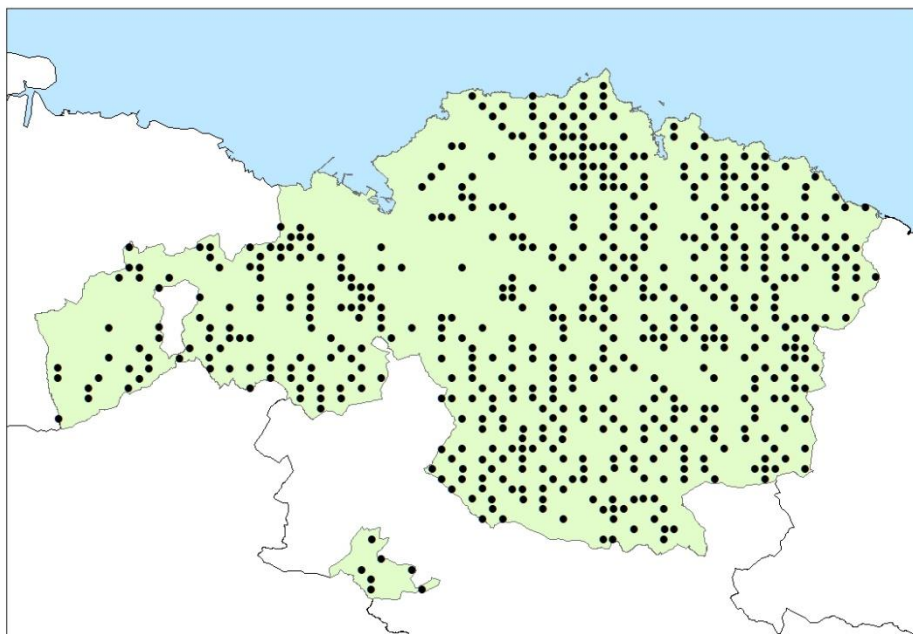


Figura 21 Parcelas con *P. radiata* del IFN4

Para cada árbol se calcularon con los datos anteriores sus coordenadas UTM en ETRS89. Estas coordenadas se corrigieron posteriormente para compensar los errores propios de la medición con GPS y los que conlleva la aproximación que se hace en la reproyección y cambio de sistema de referencia.

Una vez eliminados los pies anteriores, se seleccionaron únicamente aquellas parcelas en las que el *P. radiata* representaba más del 80% de los pies medidos o en las que las coníferas representan más del 90% y el pino radiata más del 60% de los pies medidos.

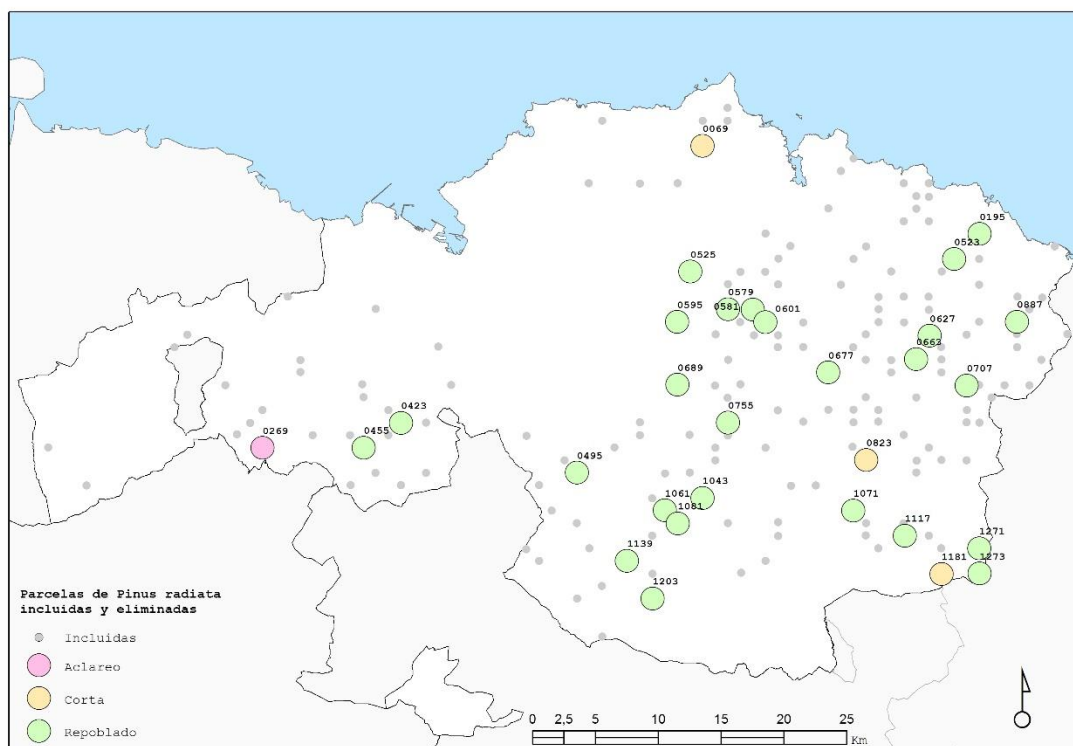


Figura 22 Parcelas incluidas y eliminadas

Para evitar los errores que durante los tres años de desfase entre ambas fuentes de datos pudieran haber ocasionado la dinámica de cortas y repoblaciones, se realizó un análisis visual para detectar y eliminar aquellas parcelas en las que se habían realizado claras o cortas previas al apeo del IFN4. También se analizaron y eliminaron las parcelas que habían sido repobladas en fechas inmediatamente anteriores a la adquisición de los datos LiDAR y que, por lo tanto, presentaban pies con alturas inferiores a los 4m, aunque en el estadillo de campo, consecuencia del crecimiento en altura, presentarían alturas superiores al umbral preestablecido.

Con estos criterios fueron seleccionadas 152 parcelas. En ninguna de ellas se realizaron claras o cortas en los tres años que median entre la toma de datos LiDAR y los del IFN4.

En cada una de las parcelas seleccionadas se realizaron las siguientes operaciones:

- a. Ordenación ascendente de árboles por diámetro normal (Dn).
- b. Separación de cuatro grupos de árboles:
 - a. $10,0 \text{ cm} \leq \text{Dn} < 12,5 \text{ cm}$.
 - b. $12,5 \text{ cm} \leq \text{Dn} < 22,5 \text{ cm}$.
 - c. $22,5 \text{ cm} \leq \text{Dn} < 42,5 \text{ cm}$.
 - d. $42,5 \text{ cm} \leq \text{Dn}$
- c. Asignación a cada árbol del factor de expansión según su Dn. Para ello se considera la hectárea como unidad superficial de referencia. El factor es el que corresponde al cociente entre la superficie de una hectárea y la superficie de la subparcela a la que pertenece el árbol según su Dn. Los árboles del grupo 1 de la relación anterior se miden en una parcela de 5 m de radio, los del 2 en una parcela de 10m de radio, los del 3 en una parcela de radio 15 m y los del grupo 4 en una parcela de 25m de radio. En resumen, los factores de expansión son los de la tabla 9:

Tabla 9 Factores de expansión según Dn.

Dn. del árbol. (cm)	Radio parcela (m)	Superficie parcela (m²)	Factor de expansión
10,0 - 12,5	5	78,54	127,324
12,5 - 22,5	10	314,16	31,831
22,5 - 42,5	15	706,86	14,147
Mayor de 42,5	25	1.963,45	5,093

6.1.4. Cálculo de la biomasa de las parcelas del IFN4

El cálculo de la biomasa se ha realizado utilizando las ecuaciones alométricas a partir de variables dasométricas explicadas en el capítulo de la metodología, concretamente se han aplicado las siguientes ecuaciones:

1.- (Canga et al., 2013) donde se midieron los diámetros y las alturas de cada árbol por métodos destructivos, obteniendo un modelo por regresión lineal múltiple capaz de explicar el 94% de la variabilidad con un RMSE de 139 kg:

$$W_{br} = 0,009892(d^2h)^{1,023} - 0,00434d^2h + 61,57 - 6,978d + 0,3463d^2 \quad 6$$

2.- (Pérez-Cruzado, 2008) con un modelo de cálculo de biomasa a nivel de árbol realizado integrando distintas fracciones de biomasa del árbol (tabla 10).

Tabla 10 Ecuaciones alométricas de *P. radiata* de Pérez-Cruzado

Fracción	Ecuación ⁷	R2 Ajust.	RMSE
Madera	$P = 0,0123 \cdot d^{1,6042} \cdot h^{1,4131}$	0,96	53,56
Corteza	$P = 0,0036 \cdot d^{2,6564}$	0,92	11,14
Ramas gruesas	$P = 1,937699 + 0,001065 \cdot d^2 \cdot h$	0,66	19,95
Ramas finas	$P = 0,0363 \cdot d^{2,6091} \cdot h^{-0,9417}$	0,81	6,02
Ramillos	$P = 0,0078 \cdot d^{1,9606}$	0,69	1,49
Acículas	$P = 0,0423 \cdot d^{1,7141}$	0,79	6,97

3.- (Montero González et al., 2006) que igualmente parte de un modelo de integración de los componentes de biomasa del árbol completo (tabla 11).

⁶ Los pesos de cada fracción se expresan en kg de peso seco, el diámetro normal (d) se expresa en cm y la altura total (h) en metros.

⁷ Igual que el anterior

Tabla 11 Ecuaciones alométricas de *P. radiata* de Montero Gonzalez

Biomasa	Parámetros		R ² adj	SEE
	a	b		
<i>P. radiata</i>				
BT	-2,61093	2,48739	0,977	0,193270
BF	-3,02878	2,56358	0,976	0,200075
BR ₇	-10,5693	3,64861	0,710	0,525330
BR ₂₋₇	-4,12515	2,1173	0,746	0,615400
BR ₂	-3,53532	1,75877	0,669	0,616072
BA	-5,03445	2,05803	0,739	0,609518
Br	-2,78485	2,14449	0,939	0,309544

BT = biomasa aérea total del árbol.

BF = biomasa del fuste.

BR₇ = biomasa de las ramas de diámetro mayor de 7 cm.

BR₂₋₇ = biomasa de las ramas de diámetro entre 2 y 7 cm.

BR₂ = biomasa de las ramas de diámetro menor de 2 cm.

BA = biomasa de las acículas.

BH = biomasa de las hojas.

Una vez calculada la biomasa de cada árbol, se ha procedido a calcular la biomasa total por hectárea referida a cada parcela acumulando la biomasa de cada árbol multiplicada por el factor de expansión que le corresponde. A modo de ejemplo se muestra su obtención para el caso de la parcela 0138 (tabla 12).

Tabla 12 Ejemplo de cálculo de la biomasa en una parcela del IFN4

Parc.	nArbol	Rumbo	Dist.	Esp.	H	Dm	FE	Biomasa Canga	Biomasa Pérez-Cruzado	Biomasa Montero
0138	1	16	9,60	028	21,60	239,5	14,147	2,71	2,94	2,86
	3	110	13,50	028	22,80	292,5	14,147	4,40	4,43	4,70
	4	111	21,30	028	28,40	453	5,093	4,81	4,39	5,02
	5	111	3,50	028	23,30	376,5	14,147	7,94	7,10	8,80
	6	116	7,90	028	23,80	385,5	14,147	8,47	7,58	9,33
	7	140	12,30	028	26,70	335,5	14,147	6,56	6,69	6,60
	8	152	10,60	028	30,80	372	14,147	8,98	9,41	8,54
	9	182	9,80	028	22,10	362,5	14,147	7,09	6,29	8,01
	10	203	9,10	028	26,40	354	14,147	7,37	7,26	7,55
	11	208	15,10	028	23,50	307,5	14,147	5,02	4,99	5,32
	12	222	8,20	028	21,00	262,5	14,147	3,29	3,35	3,59
	13	263	11,40	028	20,90	302	14,147	4,53	4,28	5,08
	14	293	1,00	028	23,70	351,5	14,147	6,85	6,39	7,42
	15	295	7,00	028	23,70	327,5	14,147	5,82	5,63	6,22
	16	298	10,80	028	24,10	412,5	14,147	9,94	8,67	11,04
	17	327	13,00	028	21,40	277	14,147	3,76	3,76	4,10
	18	331	9,90	028	19,90	176	31,831	2,99	3,55	2,99
	19	352	11,30	028	21,00	239,5	14,147	2,67	2,85	2,86
	20	352	8,70	028	12,50	132	31,831	1,47	1,36	1,46
	TOTAL BIOMASA PARCELA								104,67	100,92

- *Parc.*: código de la parcela.
- *nArbol*: Número de árbol medido dentro de cada parcela.
- *Rumbo*: Rumbo de la línea que une el centro de la parcela con cada árbol en grados centesimales.
- *Dist.*: Distancia del árbol al centro de la parcela en metros.
- *Esp.*: Código especie del IFN4.
- *H*: Altura del árbol en metros.
- *Dm*: Media aritmética de los dos diámetros normales medidos en milímetros.
- *FE*: Factor de expansión.
- *Biomasa* (Canga / Pérez-Cruzado / Montero): Biomasa en t/h.

Las coordenadas UTM del centro de cada parcela se determinaron en campo con navegador GPS en el sistema de referencia ED50. El error cometido en este tipo de medidas sin corrección en postproceso posterior puede ser de varios metros, lo cual podía afectar gravemente a los datos del estudio. Para evaluar el posible impacto de esta fuente de error se ha realizado una comprobación que viene explicada en el punto siguiente.

Las coordenadas del centro de las parcelas seleccionadas se almacenaron en una tabla de formato excel junto con los datos relativos a pies mayores recogidos en el estadillo y se cargaron en el Sistema de Información Geográfica ArcGIS como capa de eventos. A la capa de centros de parcela se le aplicó una reproyección de código EPSG 1632. De esta forma se obtuvo la capa de centros de parcela en el sistema de referencia ETRS89. Todos los cálculos posteriores están referidos a este sistema.

En la tabla 13 y la figura 23 se muestran los valores de biomasa calculados para las 151 parcelas incluidas en el estudio a partir de las tres ecuaciones alométricas empleadas (Canga et al., 2013; Montero González et al., 2006; Pérez-Cruzado, 2008):

Tabla 13 Biomasa de las parcelas seleccionadas para el estudio

PARCELA	BIOMASA MONTERO (t/ha)	BIOMASA CANGA (t/ha)	BIOMASA PÉREZ-CRUZADO (t/ha)
0007	257,95	242,40	247,05
0025	200,31	190,07	174,92
0045	498,86	460,87	401,91
0051	213,73	210,07	208,77
0053	331,96	303,26	278,45
0088	54,61	45,44	41,35
0099	168,15	152,17	145,62
0103	163,76	160,28	161,05
0106	279,41	268,61	242,72
0121	255,84	241,32	245,48
0123	138,69	125,28	114,84
0138	111,47	104,69	100,92
0139	152,30	160,59	176,54
0149	280,93	284,39	278,24
0155	150,39	151,51	157,83
0173	81,61	76,75	73,77
0175	167,48	170,37	179,56
0179	205,20	197,31	181,41
0213	361,26	327,98	297,76
0217	184,16	161,17	152,14
0231	136,32	124,72	113,76
0243	280,83	252,72	215,02
0249	231,65	238,04	234,56
0257	247,97	245,85	233,14
0261	193,52	193,88	179,80

0269	144,37	126,13	103,89
0307	146,14	125,70	117,90
0325	141,39	106,97	75,15
0353	207,58	195,15	196,14
0361	209,60	179,81	143,39
0385	164,63	140,74	132,06
0389	155,62	143,62	142,54
0397	128,30	117,95	121,68
0409	124,05	108,03	94,70
0425	217,44	211,90	188,71
0435	101,48	100,93	100,54
0437	125,44	122,49	114,34
0439	151,23	155,01	157,54
0443	118,29	107,13	103,00
0465	77,07	64,01	56,29
0479	349,06	317,98	265,45
0487	194,46	188,94	173,78
0491	49,66	40,22	37,05
0501	80,09	70,47	63,71
0503	330,47	339,87	328,90
0515	239,07	221,72	196,79
0517	277,41	292,81	304,58
0527	385,54	377,88	363,88
0529	189,11	170,03	148,43
0537	125,72	119,99	124,13
0539	173,43	178,50	180,67
0543	194,35	191,08	186,51
0545	151,84	135,33	124,86
0547	278,75	273,05	247,99
0567	275,92	279,82	281,81
0569	134,92	130,59	131,50
0571	266,72	254,48	230,29
0587	174,17	172,63	179,03

Estimación de la biomasa potencial disponible de *Pinus radiata*
para usos energéticos en Bizkaia a partir de datos LiDAR
aerotransportados. Propuesta metodológica

0591	210,37	193,68	183,08
0599	280,88	256,77	218,33
0603	216,22	190,12	168,17
0607	120,70	100,92	94,82
0609	316,14	303,87	274,62
0615	288,06	278,29	254,21
0617	350,13	331,79	280,83
0619	167,65	161,29	149,96
0625	205,96	214,30	229,67
0637	95,33	86,49	87,48
0641	109,96	104,42	104,72
0645	196,20	185,51	168,11
0652	210,37	195,37	171,19
0659	253,83	254,15	242,51
0661	143,79	142,61	148,47
0665	80,52	75,69	80,73
0681	195,47	178,50	174,94
0685	209,23	205,34	196,65
0692	386,39	379,75	352,98
0693	353,86	352,86	337,75
0711	371,39	341,88	329,81
0719	74,07	64,97	58,48
0723	36,42	31,23	28,38
0735	222,26	208,59	192,05
0739	128,25	111,56	96,93
0741	163,36	153,54	141,32
0743	234,32	214,81	185,37
0749	351,73	313,72	266,75
0751	75,09	65,10	62,69
0757	147,90	127,64	107,50
0761	173,23	156,15	144,64
0763	238,78	228,79	201,01
0767	289,92	265,05	256,07
0769	271,55	260,97	240,94
0773	283,84	274,06	255,97
0775	333,70	329,22	298,38
0783	76,20	64,28	60,35
0795	62,24	51,45	47,65
0799	89,69	80,37	74,20
0803	198,55	210,58	210,93
0805	174,80	158,92	158,04
0815	199,47	210,24	213,12
0825	243,25	225,64	199,13
0827	343,07	353,03	333,40
0829	168,15	166,74	161,57
0833	211,99	189,87	170,73
0835	169,89	157,99	159,51

0843	385,89	354,63	302,16
0851	149,57	141,20	134,95
0853	144,16	128,42	125,42
0865	103,32	94,46	90,15
0875	118,13	104,66	101,02
0877	86,45	78,13	73,77
0883	51,30	46,71	44,12
0885	129,31	128,47	130,96
0891	148,60	129,82	117,41
0893	210,50	197,41	183,13
0897	144,76	129,18	124,94
0899	300,38	298,05	279,94
0905	175,48	158,48	145,18
0907	92,28	79,21	77,34
0909	216,48	197,49	180,78
0911	247,08	242,09	219,61
0913	297,17	303,93	296,61
0921	185,88	157,43	135,51
0925	162,55	149,14	151,25
0929	195,60	184,71	158,82
0949	160,52	151,48	154,66
0957	285,79	238,83	190,34
0965	176,64	165,63	147,08
0973	123,17	99,77	88,17
0989	100,56	90,91	91,83
0993	260,06	254,39	251,00
1007	145,08	114,79	89,53
1029	236,13	208,75	178,23
1031	160,79	153,60	147,06
1039	229,42	220,43	228,26
1083	274,85	278,40	257,92
1087	306,67	269,60	227,22
1093	249,83	255,25	248,40
1095	183,88	160,59	149,91
1101	87,36	77,94	72,02
1109	198,64	185,49	177,14
1115	168,40	172,04	181,94
1119	150,00	138,70	126,85
1133	161,98	162,48	154,04
1137	243,44	214,34	182,18
1149	97,89	89,77	77,80
1165	182,20	161,65	148,71
1171	363,02	353,64	315,98
2063	85,26	63,54	54,35
2105	142,01	124,30	120,19
2129	242,80	239,99	246,80

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

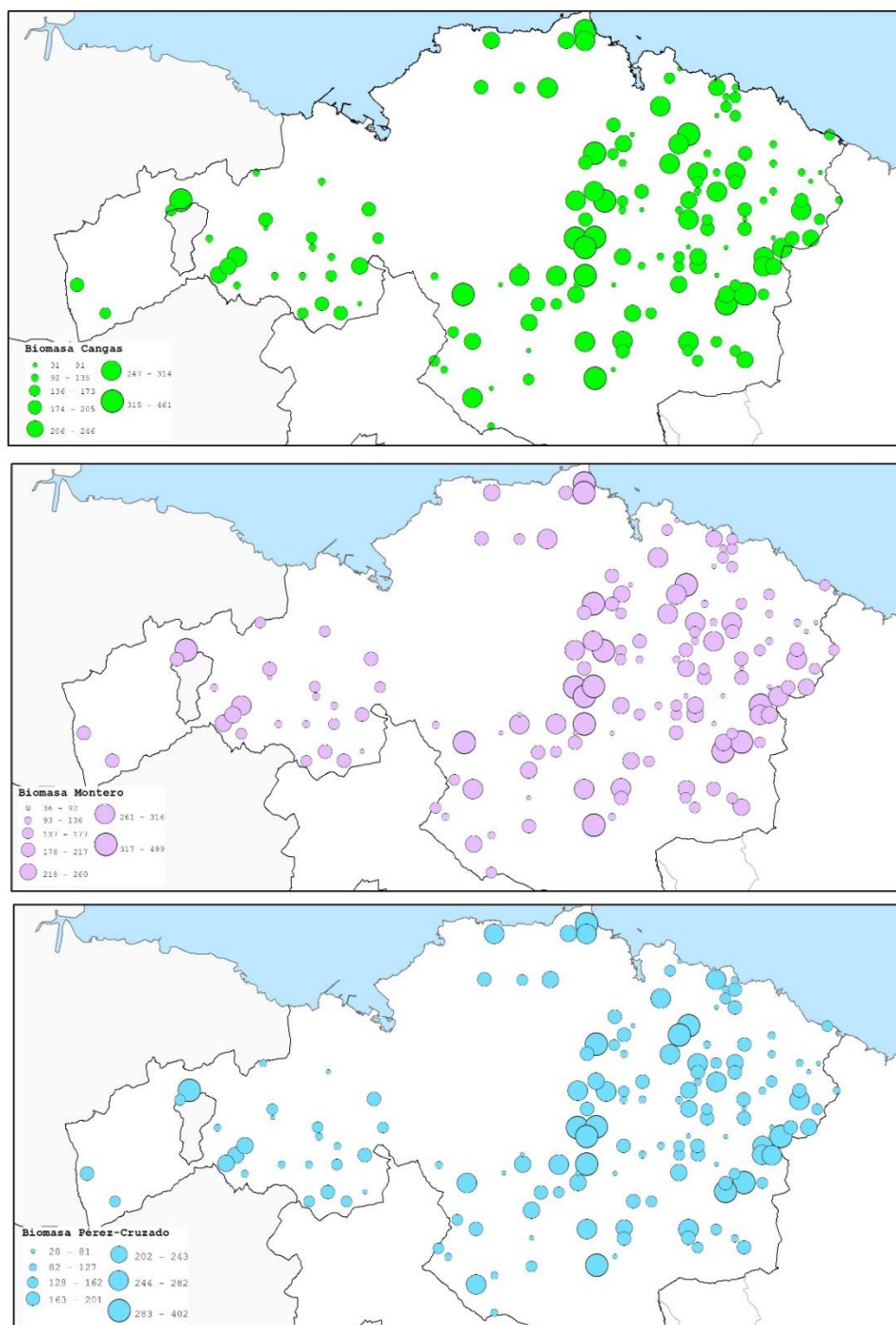


Figura 23 Cantidad de biomasa estimada por parcela a partir de las distintas ecuaciones alométricas (Canga et al., 2013; Montero González et al., 2006; Pérez-Cruzado, 2008)

6.1.5. **Comprobación del posicionamiento de las parcelas de campo**

Más arriba se ha expuesto el problema en el posicionamiento de las parcelas del IFN4. Para intentar solucionar los problemas que esto podría acarrear al estudio se plantearon dos estrategias posibles:

- Calcular las coordenadas precisas en el sistema de referencia actual al rejón de centro de todas las parcelas seleccionadas mediante trabajo de campo, usando sistemas de posicionamiento de alta precisión submétricos.
- Asumir la incertidumbre siempre y cuando se demostrara que, desde el punto de vista de masa forestal, el entorno de las parcelas se mantuviese estadísticamente homogéneo.

La primera opción se descartó, por causas tanto económicas como prácticas, pues no se tiene seguridad de poder encontrar los rejones en campo.

Para contrastar la hipótesis de homogeneidad en el entorno de cada parcela se ha diseñado un procedimiento mediante el cual se han generado una serie de parcelas "virtuales" en torno a cada parcela obteniéndose métricas LiDAR para cada una de ellas para posteriormente analizar si resultan significativamente semejantes, aceptando, en caso de que fuera así, que la falta de exactitud no debería tener especial relevancia en los resultados.

Para evaluar esta homogeneidad, se han obtenido la métrica correspondiente al percentil 95 de altura LiDAR de 8 nuevas parcelas "virtuales" generadas en el entorno de cada parcela original, trasladando su centro 10 metros hacia las 8 direcciones azimutales principales (N, NE, E, SE, S, SO, O, NO).

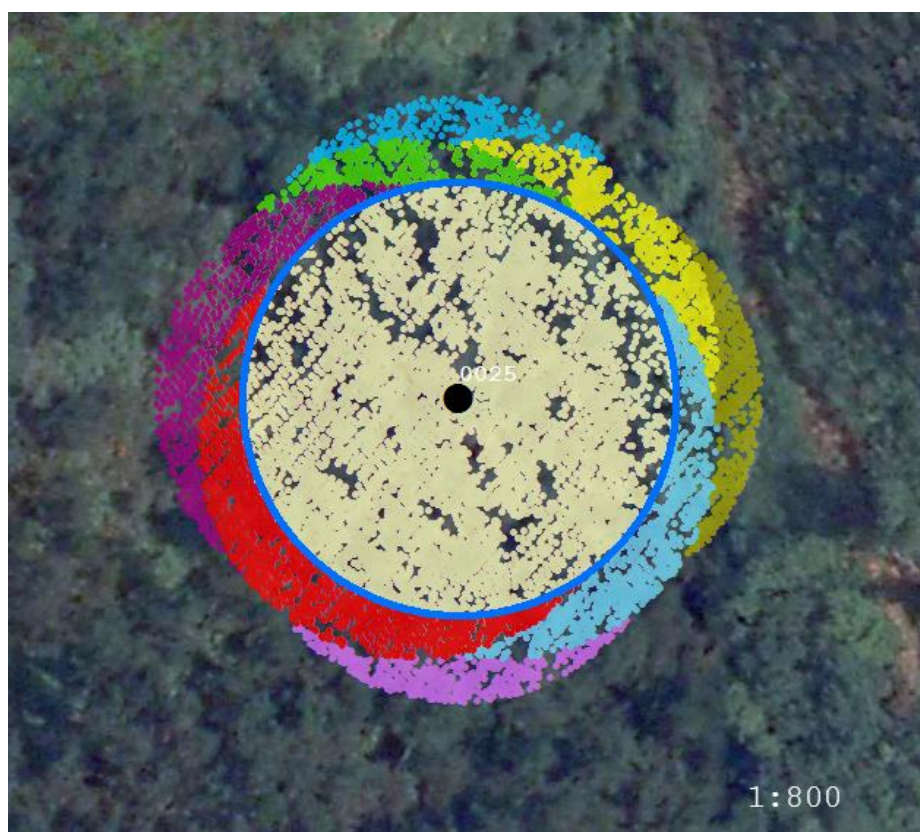


Figura 24 "Parcelas virtuales en las 8 direcciones"

En primer lugar, se ha comprobado que, tal y como se puede ver en la tabla 14, el percentil 95 de la altura LiDAR se ajusta a una distribución normal para todos los conjuntos de parcelas, tanto en el caso de las desplazadas como en las que componen el conjunto de parcelas con coordenadas originales.

Tabla 14 Pruebas de normalidad de la variable Percentil 95 de altura en las parcelas desplazadas.

P95	Shapiro-Wilk	Anderson-Darling	Lilliefors	Jarque-Bera
ORIG	0,048	0,070	0,269	0,180
E	0,053	0,074	0,145	0,174
N	0,037	0,076	0,152	0,130
NE	0,057	0,079	0,251	0,152
NO	0,053	0,062	0,192	0,157
O	0,040	0,052	0,128	0,151
S	0,206	0,193	0,191	0,348
SE	0,058	0,085	0,117	0,238
SO	0,089	0,062	0,077	0,251

Esto nos permite afirmar que en torno a la media más/menos 2 desviaciones estándar deberían encontrarse el 96% de los valores. Los gráficos X al XX muestran la diferencia entre el valor medio del percentil 95 de altura de la parcela original y el de cada una de las parcelas desplazadas, con los límites de más/menos 2 desviaciones estándar. Como se puede observar la mayoría de las parcelas se encuentran entre estos intervalos, lo cual nos permite asumir que el error inducido por posibles errores en la georreferenciación de las parcelas debería tener una incidencia mínima en la calibración de los modelos de predictivos.

Una de las posibles consecuencias de la existencia de estos errores de posicionamiento puede resultar en la aparición de casos atípicos en los modelos de regresión, los cuales, en caso de aparecer, deberán ser analizados de manera individual y evaluar su eventual retirada del modelo.

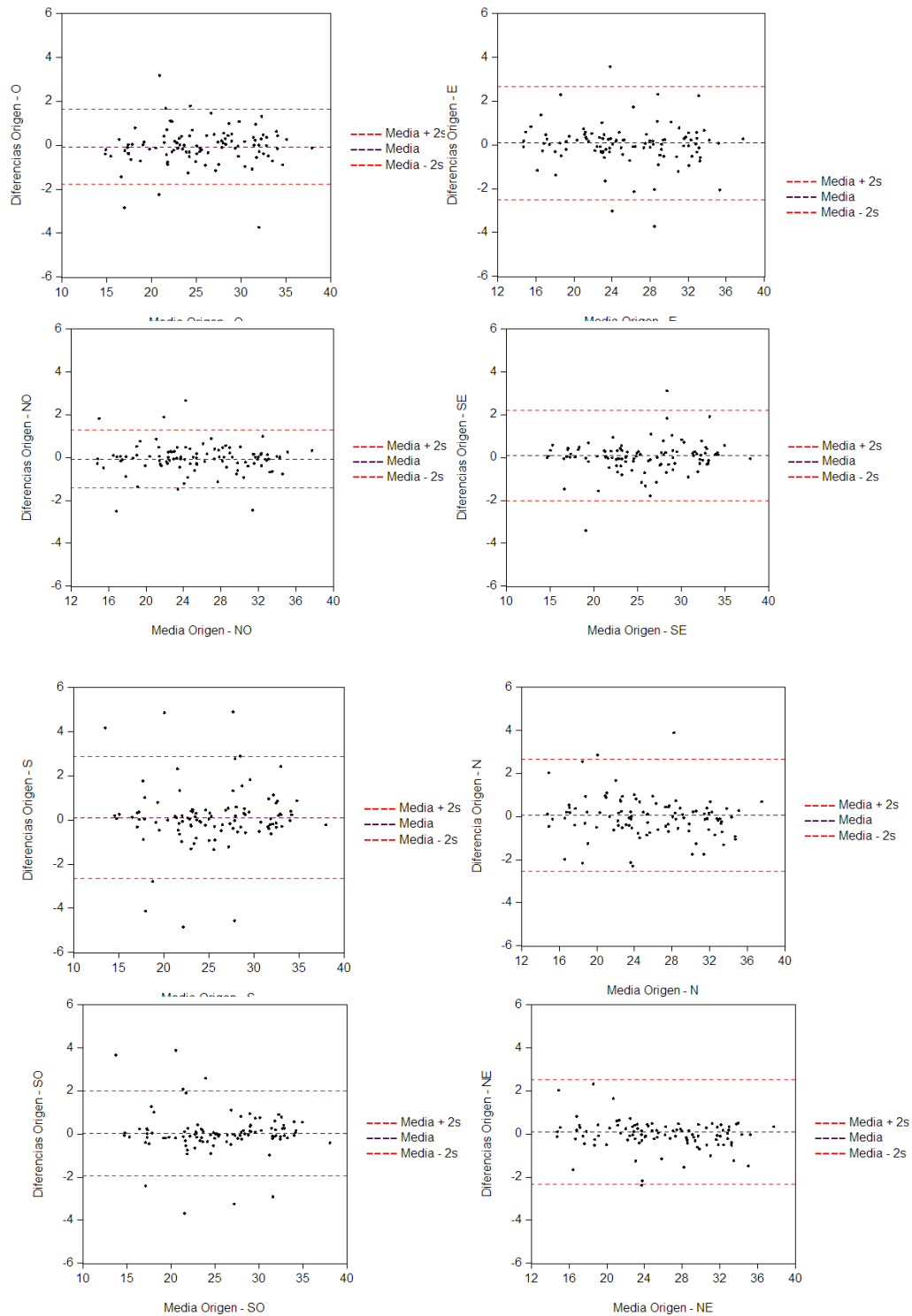


Figura 25 Diferencia entre el valor medio del percentil 95 de altura de la parcela original y el de cada una de las parcelas desplazadas

6.2. Tratamiento de los datos LiDAR

El vuelo del que proceden los datos LiDAR de Bizkaia se realizó entre los días 18 de junio y 10 de julio de 2008.

Los componentes del sistema LiDAR fueron los siguientes:

- Sistema Lite Mapper 5600 con Laser Scanner RIEGL LMS-Q560.
- Sistema de Navegación: CCNS-4 de IGI
- GPS: NovAtel OEM 4-G2 (Internal) 4-G2 L1/L2 2Hz
- INS: AeroControl IID. Frecuencia de registro 256 Hz.

Parámetros de configuración:

Tabla 15 Parámetros de configuración del sensor LiDAR

CARACTERISTICA	VALOR
Scan angle	60°
Pulse Repetition Rate (PRR)	120.000 Hz
Eff. Measurement rate	80.000 Hz
Beam divergence	0,5 mrad
Facets	4
Ground speed	110 Km/h

Densidad de puntos: Densidad media planificada de 2,5 puntos por m².

Altitud de vuelo sobre el terreno: 900 m. con incrementos de variación de 600 m.

Distancia entre líneas de vuelo: 450 m.

Recubrimiento transversal: 60%.

Corredor de línea de barrido en el terreno: 1.039 m.

Ángulo de barrido: 60 ° (ÁNGULO COMPLETO).

Velocidad de la aeronave: 56,6 m/s

Gap Límite: 60 m. Es el máximo hueco sin puntos permitido.

Estaciones de referencia: Como estaciones de referencia se emplearon las integradas dentro de la red GPS de Euskadi y dos estaciones GPS adicionales TRIMBLE 5700 centradas en el área del proyecto con una frecuencia de registro de 2 Hz.

El conjunto de puntos obtenidos se dividió según una cuadrícula de 1Km de lado y se creó un archivo independiente para cada celda.

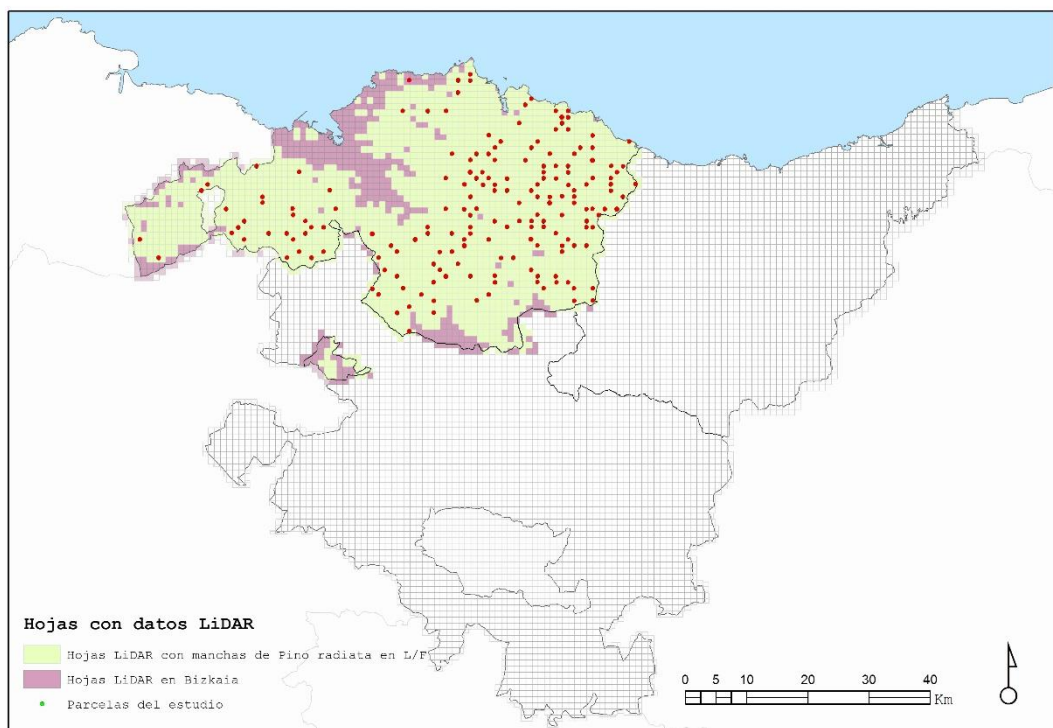


Figura 26 Hojas LiDAR cubriendo la CAE, Bizkaia, la superficie ocupada por el *P. radiata* en Latizal/Fustal y parcelas de campo seleccionadas: (elaboración propia)

El proceso previo de los puntos obtenidos se realizó por la empresa SIGRID S.L. utilizando para ello los programas

de TERRASOLID TerraMatch, V-8 y TerraScan, V-8. Se obtuvo una nube irregular de puntos con una densidad media de 3,18 puntos por metro cuadrado. Tras el proceso, Los puntos quedaron clasificados en alguna de las categorías recogidas en la tabla 5.

Con los puntos clasificados como suelo se confeccionó el Modelo Digital del Terreno (MDT) de la CAPV con una definición de 1m X 1m de lado de celda.

Tanto los archivos. las como los MDT tienen carácter público y pueden ser descargados gratuitamente en la página de la Infraestructura de datos espaciales (IDE) de Euskadi ftp://ftp.geo.euskadi.net/LiDAR/LIDAR_2008_ETRS89/.

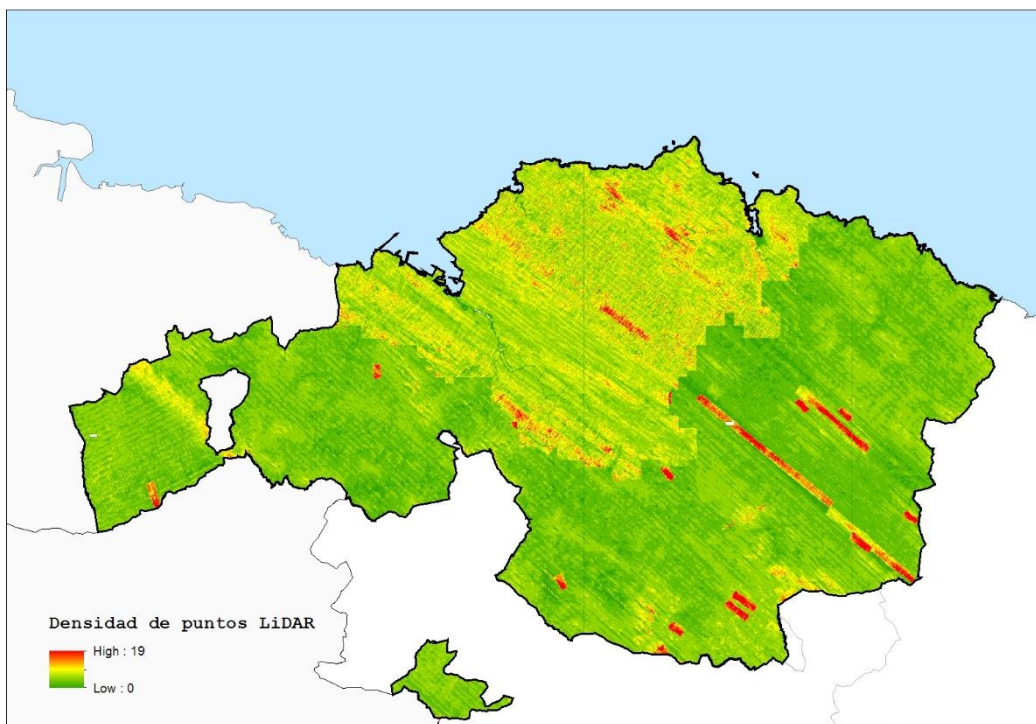


Figura 27 Densidad de puntos LiDAR: (elaboración propia)

Rango de alturas de trabajo: Los archivos de datos generados en los pasos anteriores contienen registros con alturas aberrantes que pueden llegar a cientos de metros. Esto ocurre porque los archivos las originales tenían fallos de depuración y aparecían algunos puntos con altitudes erróneas. Por otra parte, este estudio se centra en la vegetación arbórea de las masas de pino insigne. Se considera que por debajo de 4 metros sobre el suelo la interferencia del matorral puede restar validez a los resultados obtenidos. Como consecuencia de lo anterior, se estableció un intervalo de alturas sobre el MDT de límite inferior 4 metros y de límite superior 50 metros ya que el árbol más alto medido en el IFN4 de Bizkaia tiene una altura de 48,4 metros.

6.2.1. Obtención y procesamiento de las métricas

Los datos brutos que se obtienen del sensor contienen sólo las coordenadas de los puntos donde se produjeron reflexiones. No hay ninguna información explícita que diga cuando un punto pertenece o no al suelo y cuando un punto pertenece o no a la vegetación. Con el debido tratamiento se separó el conjunto de puntos que conforman el suelo dentro de la nube de puntos total. La separación de los puntos del suelo del resto de puntos está basada en el algoritmo de (Kraus y Pfeifer, 1998) y está implementada en el programa FUSION (McGaughey, 2016). En base a estos puntos se obtienen los Modelos Digitales del Terreno. Éstos se emplean para transformar la altitud de cualquier punto en altura sobre el suelo.

Conociendo los datos de altura sobre el terreno, y sabiendo qué hay sobre él, conoceremos los datos de alturas de dichos objetos. En este caso los objetos sobre el terreno son árboles, por tanto, obtenemos información relativa a la altura de la vegetación. Como se ha comentado, las métricas relacionadas con la altura han resultado ser variables con muy buena relación con la biomasa, por lo que es de esperar que se obtengan buenos modelos predictores.

Para la obtención de las variables candidatas se ha empleado el programa informático FUSION, el cual permite mediante el comando *CloudMetrics* calcular una variedad de parámetros estadísticos que describen el conjunto de datos LiDAR.

Para el cálculo de las métricas FUSION puede emplear tanto las alturas de puntos como los valores de intensidad. Estos representan la magnitud del pulso de retorno, el cual representa, en parte, la reflectividad del objeto alcanzado por el pulso láser, el cual variará dependiendo de la composición del objeto interceptado. El inconveniente principal de los datos de intensidad es que los valores varían tanto en cada elemento, en función de su posición en el vuelo, en un mismo vuelo como entre los valores observados en vuelos diferentes, y por lo que son valores relativos de escasa utilidad en este estudio. Por este motivo se decidió no emplearlos.

CloudMetrics produce un registro de salida para cada archivo de datos procesado, por lo cual se han tenido que

procesar mediante archivos de procesamiento por lotes. En nuestro estudio cada parcela supone un fichero “.LAS” de nube de puntos LiDAR. La salida final se agrega a un único archivo de salida formateado como ASCII de valores separados por comas (CSV), que posteriormente será procesado en MS-Excel y MS-Access para dejarlo listo para los posteriores análisis estadísticos.

Las métricas que proporciona el comando CloudMetrics son las siguientes (tablas 16 y 17):

Tabla 16 Métricas LiDAR relacionadas con la altura y la estructura vertical del dosel extraídas a partir de los datos LiDAR con el programa FUSION

MÉTRICAS LiDAR	
Total number of returns	Número total de retornos
Count of returns by return number (support for up to 9 discrete returns)	Número de retornos por número de retorno (hasta 9 retornos discretos)
Minimum	Mínimo
Maximum	Máximo
Mean	Media
Median (output as 50th percentile)	Mediana (salida como percentil 50)
Mode	Moda
Standard deviation	Desviación estándar
Variance	Varianza
Coefficient of variation	Coficiente de variación
Interquartile distance	Distancia intercuartílica
Skewness	Asimetría
Kurtosis	Kurtosis
AAD (Average Absolute Deviation)	AAD (Desviación Absoluta Media)
MADMedian (Median of the absolute deviations from the overall median)	MADMedian (Mediana de las desviaciones absolutas de la mediana general)
MADMode (Median of the absolute deviations from the overall mode)	MADMode (Mediana de las desviaciones absolutas de la moda general)
L-moments (L1, L2, L3, L4)	L-momentos (L1, L2, L3, L4)
L-moment skewness	L-momento de asimetría
L-moment kurtosis	L-momento de curtosis
Percentile values (1st, 5th, 10th, 20th, 25th, 30th, 40th, 50th, 60th, 70th, 75th, 80th, 90th, 95th, 99th percentiles)	Valores de los percentiles (1, 5, 10, 20, 25, 30, 40, 50, 60, 70, 75, 80, 90, 95, 99)
Canopy relief ratio ((mean-min)/(max-min))	Relación de relieve del dosel ((media-min)/(máx-min))
Generalized means for the 2nd and 3rd power (Elev quadratic mean and Elev cubic mean)	Media generalizada para la segunda y tercera potencia (Elevación media cuadrática y Elevación media cúbica)

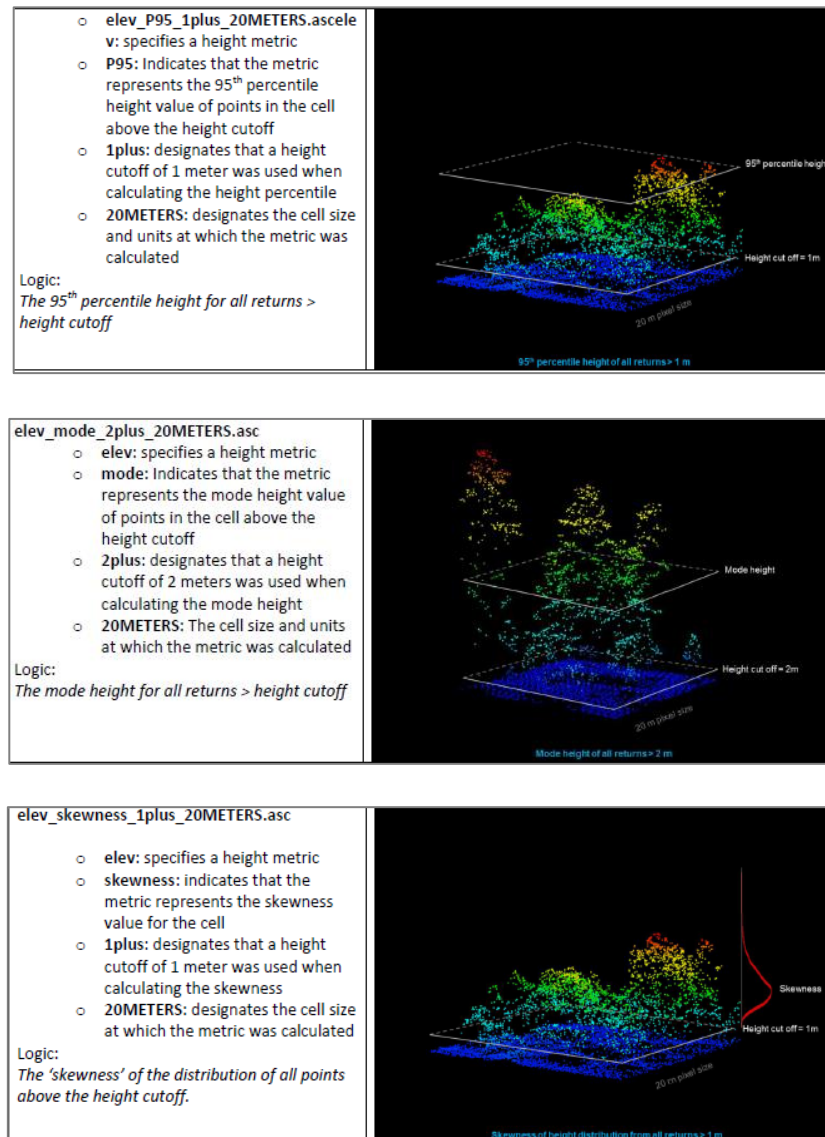


Figura 28 Métricas LiDAR explicadas en representaciones 3D (McGaughey, 2016)

Además de las métricas anteriores, CloudMetrics también calcula varios valores en términos porcentuales de retornos por encima de una altura de corte (4m en nuestro caso), cuando se utiliza el modificador / above: #:

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Tabla 17 Métricas LiDAR relacionadas con la densidad de cubierta del dosel extraídas a partir de los datos LiDAR con el programa FUSION

MÉTRICAS LiDAR	
Percentage of first returns above a specified height (canopy cover estimate)	Porcentaje de primeros retornos por encima de una altura especificada (estimación del dosel de copas)
Percentage of first returns above the mean height/elevation	Porcentaje de primeros retornos por encima de la altura media
Percentage of first returns above the mode height/elevation	Porcentaje de primeros retornos por encima de la moda de la altura
Percentage of all returns above a specified height	Porcentaje de todos los retornos por encima de una altura especificada
Percentage of all returns above the mean height/elevation	Porcentaje de todos los retornos por encima de la altura media
Percentage of all returns above the mode height/elevation	Porcentaje de todos los retornos por encima de la moda de la altura
Number of returns above a specified height / total first returns * 100	Número de retornos por encima de una altura especificada / total de primeros retornos * 100
Number of returns above the mean height / total first returns * 100	Número de devoluciones por encima de la altura media / total de primeros retornos * 100
Number of returns above the mode height / total first returns * 100	Número de devoluciones por encima de la altura del modo / total de primeros retornos * 100

Los valores de grado de cobertura calculados en CloudMetrics cuando se utiliza el modificador / above: # se calculan de la siguiente manera:

$(\# \text{ de primeros retornos} > \text{ altura de corte}) / (\text{número total de primeros retornos})$

La Figura 29 ilustra el concepto de estimación de la cubierta de dosel usando los primeros retornos LIDAR:

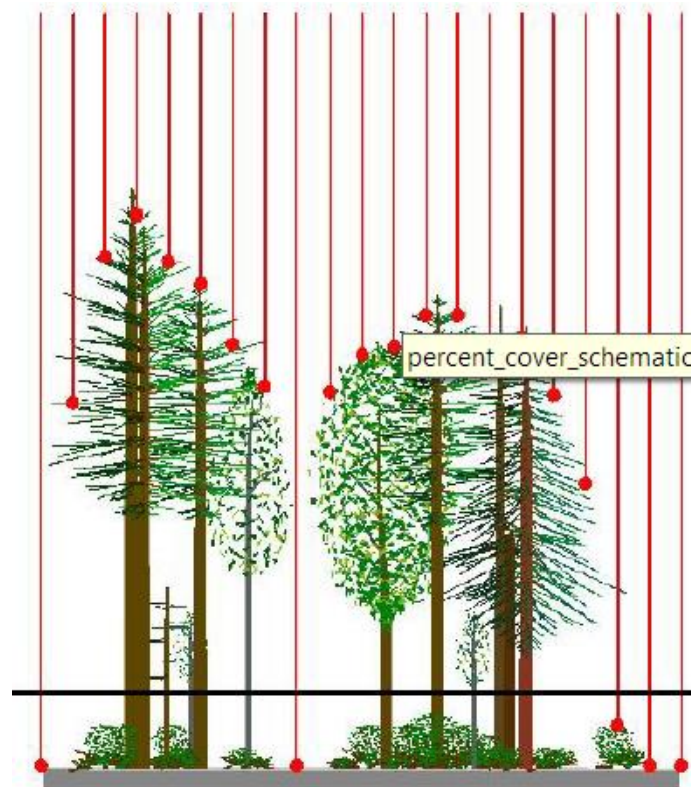


Figura 29 Esquema de cálculo del porcentaje de cobertura con una altura de corte (McGaughey, 2016)

En la Figura 29, el dosel superior se define como cualquier vegetación mayor que la altura de corte (3 metros en este ejemplo) sobre el suelo. De los 21 pulsos LiDAR que entran en el dosel, se registran 16 retornos por encima del umbral de 3 metros. La estimación de la cobertura del dosel se calcula como $16/21$ o 76 por ciento.

Los valores de los percentiles se calculan usando el siguiente método:

$$(n-1)p = i + f \left\{ \begin{array}{l} i \text{ es la parte entera de } (n-1)p \\ f \text{ es la parte fraccional de } (n-1)p \end{array} \right\}$$

$$\text{donde } \left\{ \begin{array}{l} n \text{ es el número de observaciones} \\ p \text{ es el valor percentil dividido por } 100 \end{array} \right\}$$

Si $f=0$ entonces el Valor de Percentil = x_{i+1}

Si $f > 0$ entonces el Valor de Percentil = $x_{i+1} + f(x_{i+2} - x_{i+1})$

Las observaciones se ordenan de menor a mayor

Los momentos se usan típicamente para describir las características de una distribución. Los momentos más comunes son el primero (media), segundo (varianza), tercero (asimetría) y cuarto (curtosis). La varianza proporciona una indicación de la variabilidad en las muestras, la asimetría proporciona indicaciones de cuan asimétrica es la distribución, y la curtosis da una indicación de cuan apuntada es la distribución. La asimetría se calcula utilizando la siguiente ecuación:

$$\text{Skewness} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3}{(N-1)s^3}$$

la curtosis:

$$\text{Kurtosis} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4}{(N-1)s^4}$$

La Desviación Absoluta Media (AAD):

$$\text{AAD} = \frac{\sum_{i=1}^N (Y_i - \bar{Y})}{N}$$

La Desviación Absoluta Mediana desde la Mediana (MADMedian) :

$$MAD_{median} = \text{median}_i(|X_i - \text{median}_j(X_j)|)$$

Desviación Absoluta Mediana desde la Moda (MADMode) :

$$MAD_{mode} = \text{median}_i(|X_i - \text{mode}_j(X_j)|)$$

Las medias generalizadas para la 2ª (Elev quadratic mean) y 3ª (Elev cubic mean) potencias (con $p=2$ y $p=3$ respectivamente) :

$$M_p = \left(\frac{1}{n} \sum_{i=1}^n x_i^p \right)^{\frac{1}{p}}$$

La moda se calcula dividiendo el rango de datos (elevación o intensidad) en 64 clases y contando el número de retornos en cada clase. La moda es la clase con el mayor número de valores. En los casos en que varias clases tienen el mismo número de devoluciones, la moda será la clase con el valor más bajo. Los valores de moda no son necesariamente únicos ya que varios valores pueden ocurrir con igual frecuencia. Además, la precisión de la estimación de modo varía dependiendo del rango de altura de los datos del punto de contacto. Los datos con un rango más pequeño de alturas tendrán un valor más preciso para la moda, mientras que

los datos con un rango mayor tendrán valores menos precisos.

Cover and Density Metric and Explanation	Illustration
<p>1st_cover_above3_20METERS.asc</p> <ul style="list-style-type: none"> o 1st: designates first returns are used in the numerator and the denominator of the density ratio o cover: designates that it is a density ratio o above3: designates that a cover cutoff of 3 meters was used when calculating the density ratio o 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of first returns} > 3 \text{ meters}}{\text{total \# of first returns in the pixel}} \right) * 100$</p>	
<p>all_cover_above3_20METERS.asc</p> <ul style="list-style-type: none"> o all: designates all returns are used in the numerator and the denominator of the density ratio o cover: designates that it is a density ratio o above3: designates that a cover cutoff of 3 meters was used when calculating the density ratio o 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of all returns} > 3 \text{ meters}}{\text{total \# of all returns in the pixel}} \right) * 100$</p>	
<p>all_1st_cover_above_mean_20METERS.asc</p> <ul style="list-style-type: none"> o all: designates that all returns above the mean are used in the numerator of the density ratio o 1st: designates that only first returns are used in the denominator of the density ratio o cover: designates that it is a density ratio o above_mean: designates that the mean height was used as the cover cutoff when calculating the density ratio o 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of all returns} > \text{mean height}}{\text{total \# of first returns in the pixel}} \right) * 100$</p>	
<p>1st_cover_above_mode_20METERS.asc</p> <ul style="list-style-type: none"> o 1st: designates first returns were used as the numerator and denominator of the density ratio o cover: designates that it is a density ratio o above_mode: designates that the mode height was used as the cover cutoff when calculating the density ratio o 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of first returns} > \text{mode height of all returns}}{\text{total \# of first returns in the pixel}} \right) * 100$</p>	

Figura 30 Métricas LiDAR de densidad explicadas en representaciones 3D

Además de las anteriores métricas, es posible obtener la mayor parte de las mismas en varios estratos de altura mediante el modificador /strata:[#, #, #, ...], donde # representa el límite superior para cada estrato. Los retornos se cuentan si su altura es \geq al límite inferior y $<$ al límite superior. El primer estrato contiene pulsos $<$ al primer límite. El último estrato contiene puntos \geq el último límite.

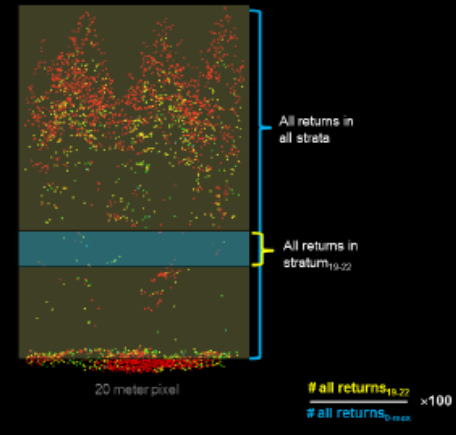
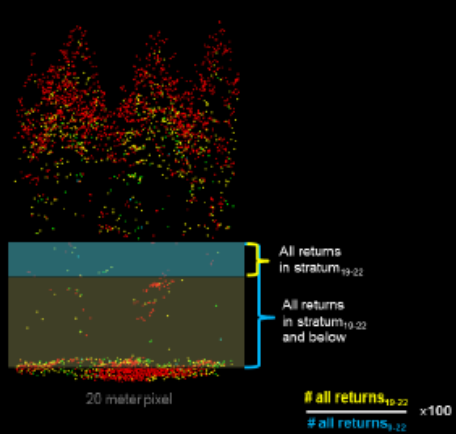
Strata Density Metric and Explanation	Illustration
<p>strata4_return_proportion_20METERS.asc</p> <ul style="list-style-type: none"> ○ strata4: specifies the stratum ○ return_proportion: indicates a relative density ratio of the point count in the specified stratum to the total points considered in each cell ○ 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of points in the strata}}{\text{total \# points of all strata in the cell}} \right) * 100$</p>	
<p>Rel_strata4_return_proportion_20METERS.asc</p> <ul style="list-style-type: none"> ○ Rel: designates a <i>normalized</i> relative density ○ strata4: specifies the stratum ○ return_proportion: indicates a relative density ratio of the point count in the specified stratum to the total points considered in each cell ○ 20METERS: designates the cell size at which the metric was calculated <p>Logic: $\left(\frac{\text{\# of points in the strata}}{\text{total \# points in the specified strata and below strata in the cell}} \right) * 100$</p>	

Figura 31 Métricas LiDAR de proporción de retornos por estrato

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

7. METODOLOGÍA ESTADÍSTICA

En este capítulo se describen las técnicas estadísticas empleadas en la estimación de la biomasa a partir de las métricas extraídas de los datos LiDAR. Partiendo del análisis exploratorio previo de las características de los datos, se describen a continuación las técnicas que permiten buscar correlaciones existentes entre las variables potencialmente predictoras y la biomasa y por último los distintos modelos de regresión más comúnmente empleados en la literatura científica.

7.1. Análisis de correlaciones

Técnica estadística que describe la relación entre dos o más variables, medidas dos a dos, en una muestra común de individuos. Para determinar la existencia de dicha relación, su fuerza, dirección y significación estadística, se utiliza el coeficiente de correlación de Pearson. Ambas variables estarán ajustadas a una escala métrica presuponiéndose que han sido obtenidas a partir de poblaciones normalmente distribuidas.

La correlación de Pearson se calcula mediante la siguiente ecuación:

$$R = \frac{(\sum xy/n) - (\bar{x}\bar{y})}{s_x s_y}$$

donde R es el coeficiente de correlación de Pearson, x e y los valores de las variables, n el número de observaciones, \bar{x} e \bar{y} son las medias de las variables y s_x

y S_y son las desviaciones estándar muestrales de las variables (Ebdon, 1982).

El resultado variará entre -1 y 1, siendo 0 la ausencia de correlación. El signo indicará si la correlación es directa (positiva) o indirecta (negativa). El nivel crítico (p) establecido para considerar las correlaciones como significativas se fija habitualmente en un nivel de significación o α en 0,05, siendo alfa la probabilidad de obtener un falso positivo en la prueba estadística, es decir, la probabilidad de rechazar una hipótesis nula siendo esta verdadera. Esto supone que se rechazará la hipótesis nula de independencia y se determinará que existe relación entre ambas variables cuando $p < 0,05$. Esta prueba de significación se efectúa mediante el estadístico t , que adopta la siguiente forma:

$$t = \frac{R\sqrt{n-2}}{\sqrt{1-R^2}}$$

donde R es el coeficiente de correlación de Pearson y n es el número de casos.

7.2. Modelos de regresión lineal

El análisis de regresión es una técnica estadística de empleada para predecir una variable (dependiente) a partir del conocimiento de una o más variables predictoras (independientes). Cada variable independiente es analizada y ponderada. Estas indican su grado de intervención en la explicación y su contribución relativa a la predicción.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

La hipótesis estructural básica es la linealidad en la relación entre variable dependiente y variables predictoras

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_j X_{ij} + \varepsilon_i, \quad i = 1, \dots, n$$

Donde:

Y_i Variable aleatoria dependiente

β_0 Coeficiente independiente de la regresión lineal

β_j Coeficientes de las variables predictoras

X_{ij} Valores conocidos de las variables predictoras

ε_i Perturbación aleatoria

La estimación más habitual de β es la denominada Mínimos Cuadrados Ordinarios (MCO) (OLS *Ordinary Least Squares*) y se realiza minimizando la suma de cuadrados de los residuos:

$$SCR(\beta) = \|Y - X\beta\|_2^2 = \sum_{i=1}^n \left(y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2$$

Derivando $SCR(\beta)$ e igualando a 0 se obtiene un sistema de ecuaciones normales: $X^{tr}X\beta = X^{tr}Y$. Como la función objetivo es convexa y diferenciable, de esta manera se obtiene un mínimo, y en el caso de que las columnas de X formen un conjunto linealmente independiente, la solución es única y viene dada por:

$$\hat{\beta}^{mco} = (X^{tr} X)^{-1} X^{tr} Y$$

Además, se establecen las siguientes hipótesis o supuestos para la perturbación:

- Esperanza nula de la perturbación
- Varianza de la perturbación constante (homocedasticidad)
- Normalidad de la perturbación
- Independencia de las perturbaciones entre sí

Cuando estas hipótesis se verifican es posible obtener expresiones que permiten analizar la significación estadística del modelo, y por lo tanto hacer inferencia (p.ej. pruebas de hipótesis e intervalos de confianza) para β o funciones de β .

Existen al menos dos razones por las cuales el estimador β^{mco} puede no resultar adecuado en ciertas situaciones (Tibshirani, 1996):

- baja precisión en las predicciones; el estimador a presenta poco sesgo pero gran variancia, lo cual se traduce en un pobre poder predictivo sobre nuevas observaciones.
- falta de interpretabilidad; si se utiliza un gran número de predictores (necesario para tener bajo sesgo ante un problema más o menos complejo), sería deseable determinar un pequeño subconjunto de estos con fuerte poder explicativo y predictivo.

Las dificultades asociadas con el primer punto se relacionan con el problema de invertir la matriz $X^{tr}X$: son tanto de tipo numéricas (problemas de redondeo que se propagan), como estadísticas (inflación de variancia). A su

Estimación de la biomasa potencial disponible de *Pinus radiata*
para usos energéticos en Bizkaia a partir de datos LiDAR
aerotransportados. Propuesta metodológica

vez, ambas desventajas del estimador por MCO están a menudo relacionadas con la existencia de predictores fuertemente correlacionados.

El caso extremo es cuando $p \gg n$ (muchas más variables que observaciones) algo que agrava estas dificultades ya que en ese caso el estimador no estará bien definido (el sistema de ecuaciones normales es indeterminado). Una situación que resulta cada vez más frecuente en muchos ámbitos de la ciencia (Genética, Bioinformática, Procesamiento de Señales, Econometría, etcétera), lo cual lo ha convertido en un área de investigación muy dinámica en los últimos años (Fan y Li, 2006, Hastie y otros, 2009, Li y Xu, 2009).

Las relaciones obtenidas al ajustar los modelos de regresión, en general, deben considerarse como una aproximación simple, dentro de un rango limitado de valores, a la variación conjunta de una serie de variables.

Este modelo presenta muchas ventajas desde el punto de vista de la estimación y la interpretación puesto que es posible incorporar en el mismo efectos no lineales en las variables ($X_j^2, \log X_j, \sqrt{X_j}$) y términos de interacción ($X_j \cdot X_k$) si se sospecha de antemano que mejorarán el modelo (a costa, eso sí, de incrementar el número de parámetros en el modelo). En ciertos casos, la transformación de variables permite que se cumplan las hipótesis del modelo de regresión lineal (Hair et al., 1999).

Este procedimiento busca pues maximizar la capacidad predictiva a partir de un conjunto de variables

independientes. Existen varias estrategias de modelización empleando para ello distintas técnicas de regresión con el doble objetivo de maximizar el grado de varianza explicada y el de minimizar el error en la predicción. En términos generales podemos denominarlas como técnicas de regresión múltiple lineal y técnicas de regresión curvilíneas.

7.2.1. Valoración de la capacidad explicativa y de la exactitud de la predicción

Para cuantificar la proporción de la varianza de la variable dependiente que es explicada por las independientes se utiliza el coeficiente de determinación (R^2), que es el cociente entre la suma de los cuadrados de la regresión y la suma total de cuadrados.

$$R^2 = \frac{\sum_{i=1}^n (Y_i' - \bar{Y})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

donde R_2 es el coeficiente de determinación, Y_i es el valor real para la observación i , \bar{Y} es la media de los valores reales, Y_i' es el valor pronosticado para la observación i (adaptado de Etxeverría, 1999; Hair et al., 1999).

El valor resultante es siempre positivo y varía entre 0 y 1, siendo potencialmente mayor el poder explicativo del modelo de regresión cuanto más alto sea (Ebdon, 1982; Etxeverría, 1999; Hair et al., 1999; Pardo y Ruiz, 2005).

Una variante del R^2 es el coeficiente de determinación corregido o ajustado (R^2_a), que tiene en cuenta el número de variables independientes incluidas en la ecuación de

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

regresión (Ecuación 6), corrigiendo el efecto de aumento de R^2 por la adición de variables independientes al modelo de regresión si éstas tienen escaso poder explicativo y/o si los grados de libertad se vuelven demasiado pequeños. De esta manera, el R^2 se hace más pequeño a medida que se tienen menos observaciones por variable independiente, resultando muy adecuado para comparar ecuaciones con un número diferente de variables independientes, tamaños muestrales diferentes o ambos (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

$$R_a^2 = R^2 - \left(\frac{p(1-R^2)}{n-p-1} \right)$$

donde R^2 es el coeficiente de determinación, p es el número de variables independientes y n es el número de casos (adaptado de Pardo y Ruiz, 2005).

Los residuos también pueden ser utilizados para valorar globalmente la exactitud de la predicción. De los distintos estadísticos que se utilizan más habitualmente se han seleccionado tres: el error estándar de la estimación (Se), el error cuadrático medio (RMSE) y el MAPE.

El primero permite una rápida comparación de los distintos modelos creados; el segundo, se ha seleccionado por ser muy utilizado en trabajos de estimación de parámetros forestales (p.e. Fazakas et al., 1999; Hagner, 2002; Pekkarinen, 2002; Mäkelä y Pekkarinen, 2004; Muukkonen y Heiskanen, 2005; Tuominen y Pekkarinen, 2005; Hall et al., 2006; Labrecque et al., 2006); el tercero se ha considerado

porque ofrece una medida del error en términos porcentuales.

Los tres se interpretan de igual forma: cuanto menor es el valor obtenido, mayor es la capacidad predictora del modelo de regresión.

$$S_e = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y'_i)^2}{n-2}}$$

donde S_e es el error estándar de la estimación, Y_i es el valor real para la observación i , Y'_i es el valor pronosticado para la observación i y n es el número de casos (adaptado de Pardo y Ruiz, 2005).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y'_i)^2}{n}}$$

donde $RMSE$ es el error cuadrático medio, teniendo los restantes términos idéntico significado que en la Ecuación 7 (adaptado de Pekkarinen, 2002).

$$RMSE_r = \frac{RMSE}{\bar{Y}} * 100$$

donde $RMSE_r$ es el error cuadrático medio relativo, $RMSE$ es el error cuadrático medio e \bar{Y} es la media de los valores reales (adaptado de Muukkonen y Heiskanen, 2005).

La valoración de la significación estadística se hace a dos niveles: a nivel de modelo y a nivel de los coeficientes de regresión de las variables independientes seleccionadas (incluida también la constante). Para

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

contrastar la hipótesis de que la cantidad de varianza explicada por el modelo es mayor que la varianza explicada por su media (es decir, que $R^2 > 0$) se utiliza el estadístico F (Ecuación 10), sobre el que se aplica la prueba de significación (normalmente $p < 0,05$) (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

$$F = \frac{\sum_{i=1}^n (Y'_i - \bar{Y})^2 / p}{\sum_{i=1}^n (Y_i - Y'_i)^2 / (n - p - 1)}$$

donde F es el estadístico F , Y'_i es el valor pronosticado para la observación i , \bar{Y} es la media de los valores reales, Y_i es el valor real para la observación i , p es el número de variables independientes y n es el número de casos (adaptado de Etxeverría, 1999; Hair et al., 1999).

Por su parte, el estadístico t y su prueba de significación permiten validar la hipótesis de que los coeficientes de regresión β_n y la constante β_0 calculados en el modelo son distintos de 0 y, por lo tanto, son representativos para toda la población (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

$$t_{\beta_n} = \frac{\beta_n}{S_{\beta_n}}$$

donde t_{β_0} es el estadístico t de la constante, β_0 es la constante y S_{β_0} es el error estándar de la constante (adaptado de Pardo y Ruiz, 2005).

$$t_{\beta_n} = \frac{\beta_n}{S_{\beta_n}}$$

donde t_{β_n} es el estadístico t del coeficiente de regresión $-\beta_n-$ de la variable independiente n y S_{β_n} es el error estándar de β_n (adaptado de Pardo y Ruiz, 2005).

7.2.2. **Supuestos del análisis de regresión lineal**

Ya se ha mencionado que el análisis de regresión ha de respetar un conjunto de supuestos subyacentes, algunos de los cuales deberán ser contrastados dos veces: en primer lugar, para las variables individuales que lo pueden integrar (dependiente e independientes) y en segundo lugar para el valor teórico del modelo una vez calculado, ya que éste actúa colectivamente sobre las variables utilizadas y, por tanto, debe cumplir los mismos supuestos que las variables individuales. Así pues, es necesario un examen previo de las variables para saber si se puede o no aplicar un modelo de regresión lineal, mientras que el examen posterior permite afirmar que éste es válido (Ebdon, 1982; Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

A continuación, se detalla cada uno de estos supuestos, indicando los métodos estadísticos utilizados para la evaluación de su cumplimiento y las soluciones que pueden ser aplicadas a las variables individuales en caso de incumplimiento.

a. Supuesto de normalidad. Aspecto fundamental del análisis de regresión, se refiere a la necesidad de que tanto la variable dependiente como las independientes se ajusten a una distribución normal. Si las variables consideradas se alejan en exceso de la normalidad, los test estadísticos de significación

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

de la regresión -t y F- carecen de valor, por lo que no se puede inferir la validez del modelo (Hair et al., 1999; Pardo y Ruiz, 2005).

La normalidad univariante de las variables del modelo de regresión lineal es fácil de contrastar mediante métodos gráficos y test estadísticos. Dentro de los primeros se encuentran el análisis del histograma y del gráfico de normalidad; destacando dentro de los segundos los test de simetría, la kurtosis y las pruebas de Kolmogorov-Smirnov (K-S) Shapiro-Wilk, Anderson-Darling (A-D), Lilliefors (L) o Jarque-Bera (J-B). En caso de que las variables no presenten normalidad pueden ser transformadas empleando, p. ej., la raíz cuadrada, logaritmos o la inversa o el cuadrado.

Otra posibilidad consiste en realizar una transformación Cox-Box, técnica utilizada para estabilizar la varianza, hacer que los datos se acerquen más a una distribución normal y mejorar la validez de las medidas de asociación, como la correlación de Pearson entre variables o la regresión

$$Y' = (Y + \lambda_2)^{\lambda_1}$$

lineal múltiple. Afecta tan solo a las variables de respuesta (variables predictoras) y pertenece a la clase de transformaciones denominadas "transformaciones de potencia", definidas por:

en la cual los datos son calculados en una potencia de λ_1 después de cambiarlo a una cierta cantidad λ_2 . Posteriormente, el parámetro de cambio λ_2 se fija igual a 0. Esta clase incluyen raíces cuadradas, logaritmos, recíprocos, y otras transformaciones comunes, que dependen sobre una potencia. Los ejemplos incluyen (tabla 18):

Tabla 18 Transformaciones Box-Cox

Potencia	Transformación	Descripción
$\lambda_1 = 2$	$Y' = Y^2$	Cuadrado
$\lambda_1 = 1$	$Y' = Y$	Datos sin Transformar
$\lambda_1 = 0.5$	$Y' = \sqrt{Y}$	Raíz Cuadrada
$\lambda_1 = 0.333$	$Y' = \sqrt[3]{Y}$	Raíz Cúbica
$\lambda_1 = 0$	$Y' = \ln(Y)$	Logaritmo
$\lambda_1 = -0.5$	$Y' = \frac{1}{\sqrt{Y}}$	Raíz Cuadrada Inversa
$\lambda_1 = -1$	$Y' = \frac{1}{Y}$	Recíproco

Normalmente, las distribuciones con asimetría positiva (a la derecha) se transforman de forma más efectiva con logaritmos, mientras que las que tienen asimetría negativa (a la izquierda) lo hacen con la raíz cuadrada. De todos modos, lo recomendado por la literatura es aplicar todas las transformaciones posibles para después seleccionar la variable transformada más apropiada (Hair et al., 1999; Pardo y Ruiz, 2005).

Aunque la normalidad de las variables utilizadas ayuda a obtener normalidad en el modelo de regresión,

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

ésta no queda garantizada, por lo que es necesario contrastarla a partir del estudio de la distribución de los residuos tipificados. Se considera bueno el modelo si estos residuos se ajustan a una distribución normal; para validar este extremo pueden usarse los mismos métodos que con las variables individuales (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

b. Supuesto de homocedasticidad. Hace referencia al supuesto de que la variable dependiente ha de tener nivel de varianza a lo largo del rango igual al de las variables independientes. Dicho de otra manera, analizar si la varianza de los residuos es constante. La mayoría de problemas de varianzas desiguales entre la variable dependiente y las dependientes (heterocedasticidad) se relaciona con diferencias en unidades entre ambas y/o a la ausencia de distribución normal en la dependiente, en las independientes o en ambos tipos (Hair et al., 1999).

La prueba de homocedasticidad entre la variable dependiente y cada una de las independientes que se van a utilizar en el modelo de regresión se contrastará mediante el test de Breusch-Pagan, que fija como hipótesis nula que los residuos son homocedásticos.

Otra posibilidad es la de analizar visualmente los gráficos de dispersión individuales que relacionan ambos tipos de variables. Este test gráfico también

es útil para detectar el incumplimiento de este supuesto una vez ajustado el modelo, usándose en este caso para su edición los residuos tipificados frente a los valores pronosticados (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

Los problemas de heterocedasticidad pueden solucionarse a través de la aplicación de las mismas transformaciones de datos que las usadas para conseguir la normalidad ya que, como se ha indicado, en muchas ocasiones la heterocedasticidad es resultado de la no-normalidad (Hair et al., 1999).

c. Supuesto de linealidad. El cumplimiento de este supuesto entre la variable dependiente y las independientes es básico, ya que, si no se cumple, carece de sentido utilizar el método de regresión lineal. La forma más común de evaluar la linealidad entre variables es el gráfico de dispersión, siendo también este método adecuado para evaluar este supuesto en el modelo de regresión obtenido mediante el uso de los valores tipificados de los residuos y de los valores pronosticados (Hair et al., 1999; Etxeverría, 1999).

En nuestro caso se empleará el test RESET de Ramsey, que indica si el valor esperado de la variable dependiente depende linealmente de las variables independientes, es decir, si las combinaciones no lineales de los valores ajustados ayudan a explicar

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

la variable de respuesta. La hipótesis nula parte de considerar que existe una dependencia lineal.

Las transformaciones comentadas arriba pueden resolver el problema de la ausencia de linealidad entre las variables, siendo también posible la utilización de variables adicionales polinómicas que representan los componentes no lineales, aunque su creación está condicionada a la existencia de una explicación teórica, ya que si no se corre el riesgo de encontrar buenos resultados que sean poco generalizables (Hair et al., 1999; Etxeverría, 1999).

d. Supuesto de independencia. Este supuesto refiere a que cualquiera de los errores obtenidos en la predicción no debe estar relacionado con el resto; es decir, los residuos constituyen una variable aleatoria que no responde a ningún patrón. Si esto sucede, significa que existe alguna relación sistemática en la variable dependiente que no ha sido explicada con el conjunto de independientes utilizadas (Hair et al., 1999; Etxeverría, 1999).

Para identificar este hecho se utiliza el gráfico de dispersión de los residuos tipificados frente a los valores pronosticados o frente a cualquier variable de secuencia habitual como, por ejemplo, el tiempo (Hair et al., 1999; Etxeverría, 1999).

Otra forma de evaluación es mediante la utilización del estadístico Durbin-Watson, aunque debemos señalar que este supuesto afecta especialmente en

conjuntos de datos que representan series temporales, mientras que, en los conjuntos de datos con una distribución espacial, como es el caso que nos ocupa carece de importancia. (Etxeverría, 1999; Pardo y Ruiz, 2005).

e. Supuesto de no-colinealidad. Refiere a la inexistencia de altas correlaciones entre las variables independientes utilizadas en el modelo. El incumplimiento de este supuesto es grave en el caso de que el objetivo del análisis de regresión sea encontrar variables explicativas puesto que da lugar a modelos afectados por colinealidad o multicolinealidad (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

Los efectos de la multicolinealidad se pueden clasificar en términos de explicación y estimación. Los primeros conciernen principalmente a la capacidad del procedimiento de regresión y a la capacidad del investigador para representar y comprender los efectos de cada variable en el valor teórico de regresión. Así, la presencia de multicolinealidad limita el tamaño del R^2 y hace más difícil añadir una predicción explicativa extra con variables adicionales. Además, la explicación del efecto que las variables independientes tienen en la dependiente se hace difícil debido a que sus efectos son mixtos o se confunden (niveles altos de varianza compartida y niveles bajos de varianza única).

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

Por su parte, los efectos sobre la estimación refieren al cálculo incorrecto de los coeficientes de regresión y de sus pruebas de significación estadística, hasta el punto de que resulte imposible calcular ningún coeficiente de correlación, de equivocarse el sentido de la relación entre las variables independientes y la dependiente o de hacerlos altamente inestables (pequeños cambios en los datos producen cambios muy grandes). Por otro lado, algunos métodos de selección de variables como la regresión por pasos no funcionan tan bien con multicolinealidad. (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

Sin embargo, hay autores que defienden que la multicolinealidad no afecta a cuán bien encaja el modelo cuando el objetivo de este no es analizar las relaciones causales sino producir predicciones. Si el modelo satisface los supuestos residuales y tiene un R^2 satisfactoria, incluso un modelo con multicolinealidad severa puede producir grandes predicciones, eso sí, la capacidad explicativa de los coeficientes de β quedaría muy comprometida (Clarke, Fokoue y Zhang, 2009).

En este estudio, a pesar de que el objetivo principal es obtener predicciones ajustadas y con un grado de exactitud elevado, se ha optado por respetar los supuestos de la regresión lineal clásica puesto que ayudará a comprender mejor las posibles relaciones existentes entre un conjunto muy heterogéneo de

métricas LiDAR y la biomasa de una conífera (VERICAT GRAU y PIQUÉ NICOLAU, 2013).

El método más simple de identificar la colinealidad es el examen de la matriz de correlación de las variables independientes. La existencia de una elevada correlación entre dos de ellas indica que, si se utilizan juntas en el modelo, se incumplirá el supuesto (Hair et al., 1999). A este respecto se debe señalar que no existe consenso acerca del grado máximo de correlación permisible entre variables independientes (Pardo y Ruiz, 2005).

Aunque se lleve a cabo este análisis, un modelo de regresión múltiple puede violar el supuesto de no-colinealidad, ya que la ausencia de correlación entre dos variables no garantiza la no-colinealidad en el modelo, que puede deberse a los efectos combinados de dos o más variables independientes.

Para evaluar la colinealidad del modelo de regresión ya ajustado se pueden utilizar tres estadísticos: el valor de tolerancia, el factor de inflación de la varianza (VIF) y los índices de condición (Hair et al., 1999; Etxeverría, 1999; Pardo y Ruiz, 2005).

En este trabajo se empleará el factor de inflación (VIF) de la varianza como medida de colinealidad entre variables predictoras, siendo 1 su valor óptimo, que implicaría la ausencia total de la misma (Kleinbaum, Kupper, Muller 1991).

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

Si se detectara multicolinealidad en el modelo a pesar de haber introducido variables independientes poco correlacionadas, existen varias opciones:

- (i) Omitir una o varias correlacionadas e identificar otras.
- (ii) Utilizar el modelo con las variables correlacionadas sólo para predecir, sin interpretar los coeficientes.
- (iii) Combinar linealmente predictores, como sumarlos o multiplicarlos.
- (iv) Utilizar las correlaciones simples entre cada variable independiente y la variable dependiente para entender la relación entre ambas.
- (v) Utilizar métodos más complejos, como RLM con factores PCA (Análisis de Componentes Principales) ortogonalizados Varimax (RCP - Regresión de Componentes Principales), - dado que los PC están linealmente no correlacionados (rotación de Varimax), no hay multicolinealidades entre los coeficientes de proyección-. Estos procedimientos dificultan sin embargo la posterior asignación espacial de los valores estimados al no poder contar con las variables transformadas como capas de información espacial.
- (vi) Utilizar métodos de búsqueda secuencial para ajustar el modelo (Hair et al., 1999; Etxeverría, 1999).

En cualquier caso, habrá que tenerse en cuenta que todos los remedios tendrán desventajas potenciales. Si es posible convivir con estimaciones de coeficiente menos precisas, o con un modelo que tenga un elevado R^2 , pero pocos predictores significativos, no hacer nada puede ser la decisión correcta porque no afectará el ajuste y proporcionará buenas estimaciones.

f. Ausencia de valores atípicos, Es siempre conveniente evaluar la ausencia de valores anómalos (*outliers*), para lo cual se pueden seguir varias estrategias basadas, en su mayoría, en el análisis de las distancias de los residuos estandarizados, como la distancia de Cook y el análisis de valores influyentes (DDFITs y Hat values).

Los **valores de apalancamiento** (Hat values) proporcionan información acerca de la posibilidad de que una observación incluya valores predictores inusuales en comparación con el resto de los datos. Los apalancamientos son una medida de la distancia entre los valores x de una observación y la media de los valores x para todas las observaciones. Un valor de apalancamiento grande indica que los valores x de una observación se encuentran lejos del centro de los valores x para todas las observaciones. Las observaciones con gran apalancamiento podrían ejercer una influencia considerable en el valor ajustado y, en consecuencia, sobre el modelo de regresión.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

Los valores de apalancamiento se encuentran entre 0 y 1. Un valor de apalancamiento mayor que $2p/n$ o $3p/n$, donde p es el número de predictores más la constante y n es el número de observaciones, es considerado un valor grande y se debe examinar.

La **distancia de Cook**, o D , es una medida general del impacto combinado de cada observación sobre los valores ajustados. Dado que D se calcula utilizando valores con apalancamiento y residuos estandarizados, considera la posibilidad de que una observación sea poco común con respecto tanto a los valores x como a los valores y .

Geométricamente, la distancia de Cook es una medida de la distancia entre los valores ajustados calculados con y sin la observación i ésima. Los valores grandes, los cuales son indicativos de observaciones poco comunes, pueden producirse porque la observación posee 1) un residuo grande y un apalancamiento moderado, 2) un apalancamiento grande y un residuo moderado o 3) un residuo y un apalancamiento grandes.

Algunos autores recomiendan comparar D con la distribución $F(p, n-p)$. Si D es mayor que el valor F en el percentil 50, entonces D es considerado extremo y se debe examinar. Otros recomiendan comparar las estadísticas D entre sí, identificando valores que sean extremadamente grandes en relación con los demás valores. Una sencilla manera de comparar los valores D consiste en ponerlos en una gráfica

temporal donde el eje x representa las observaciones, no un índice ni un período de tiempo.

Por su parte las medidas **DFITS** utilizan el apalancamiento y el residuo eliminado (studentizado) para calcular la diferencia entre el valor ajustado calculado con y sin la observación *i*ésima. DFITS representa, aproximadamente, el número de desviaciones estándar estimadas en que el valor ajustado cambia cuando la observación *i*ésima es eliminada de los datos. Algunos autores sugieren que una observación con un valor de DFITS mayor que la raíz cuadrada de $2p/n$ es influyente. Otros recomiendan comparar los valores DFITS entre sí, para identificar valores que sean extremadamente grandes con respecto a los demás. Al igual que con los valores D una posible manera de comparar DFITS consiste en graficar los valores de DFITS en una serie temporal, donde el eje x representa las observaciones.

g. Tamaño de la muestra. Además de estos supuestos, hay que tener en cuenta la influencia del tamaño de la muestra para la posterior generalización de los resultados. Así, con el fin de que el modelo no esté sobre-ajustado a la muestra y sea generalizable a toda la población, se recomienda un mínimo de entre 15 y 20 observaciones por cada variable independiente incluida (Hair et al., 1999).

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

7.2.3. Métodos de selección de variables

La mayoría de las técnicas estadísticas funcionan mal con alta dimensionalidad. El análisis de regresión común rápidamente se vuelve poco fiable cuando la dimensionalidad se hace alta, un fenómeno conocido como la "maldición de la dimensionalidad" o "efecto Hughes" (Hastie & Tibshirani 1990). Una consecuencia inmediata de esta "maldición de la dimensionalidad" es la limitación del número de variables utilizables en el análisis.

Existen varias estrategias para seleccionar las variables que conformarán el modelo de regresión. En esencia se podrían resumir en dos grandes enfoques: la especificación confirmatoria y el enfoque de búsqueda secuencial.

El primero resulta adecuado cuando se quiere utilizar el modelo de regresión para confirmar la existencia de una relación lineal entre la variable independiente y un conjunto preseleccionado de variables explicativas, siendo el investigador quien especifica las mismas; se parte de la hipótesis de que el conjunto de variables escogidas consigue la máxima capacidad predictiva mientras se mantiene la robustez del modelo.

El segundo enfoque implica hacer una búsqueda de variables para obtener el modelo más robusto, eficiente y estable posible (George, 2000). Una primera aproximación consiste en ajustar los 2^p modelos posibles y comparar los mejores de cada tamaño $k \in \{1, \dots, p\}$. La comparación se realiza a través de alguna medida que tenga en cuenta el ajuste a los datos de entrenamiento pero que penalice la complejidad

del modelo de forma que mantenga una buena capacidad predictiva (generalización sobre datos nuevos) (Yee & Mitchell, 1991).

Esta aproximación proporciona a priori un método menos subjetivo de selección de las variables para maximizar la predicción con el número más pequeño de ellas (Hair et al., 1999), evitando incumplir el supuesto de no-colinealidad (Etxeverría, 1999).

Dentro de este tipo de estrategias de selección de variables destacan, los *métodos del mejor subconjunto*, los cuales se llevan a cabo mediante algoritmos que utilizan estructuras anidadas de los distintos modelos (*leaps and bound*) (Furnival & Wilson, 1974). Sin embargo, estos son practicables solo si p no es demasiado grande (p.ej. $p \approx 40$, $2^{40} \approx 1.0995 \times 10^{12}$).

Cuando la búsqueda exhaustiva de todos los submodelos posibles no es viable por el tamaño de p , una segunda opción pasa por los métodos que buscan un "*buen camino*" a través del espacio de modelos.

Métodos secuenciales

Las técnicas más conocidas en estos casos son los métodos secuenciales o por pasos (*stepwise*), en los cuales al pasar de un modelo a otro se agregan o eliminan variables de una a una (*Forward Selection, Backward Elimination o Forward-Backward*). Estos métodos basan la selección de variables en varios criterios estadísticos (Pérez López, 2009):

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

Criterio de significación. Se incorporan al modelo de regresión sólo las variables que contribuyen al modelo de forma significativa. Son dos los criterios de significación más comúnmente utilizados:

Probabilidad de F. Una variable entra a formar parte si el nivel crítico asociado a su coeficiente de correlación parcial al contrastar la hipótesis de independencia es $<0,05$, queda fuera si es $>0,10$.

Valor de F. Una variable se introduce si el valor de F utilizado para contrastar la hipótesis de independencia es $>3,84$, queda fuera si es $<2,71$.

Criterio de tolerancia. Superado el criterio de significación, sólo se integra una variable si su nivel de tolerancia es mayor que el establecido por defecto (0,0001) y si su incorporación al modelo no hace que una de las previamente seleccionadas pase a tener un nivel de tolerancia inferior al nivel establecido.

Criterio de información: Los más empleados son el de Akaike (AIC), el C_p de Mallows, o los basados en criterios de información bayesianos. Dado un conjunto de modelos candidatos para los datos, el modelo preferido es el que proporciona los valores mínimos del criterio empleado, no sólo recompensan la bondad de ajuste, sino también incluye una penalización, que es una función creciente del número de parámetros estimados. Esta penalización disuade el sobreajuste

porque aumentar el número de parámetros en el modelo casi siempre mejora la bondad del ajuste.

El método de pasos sucesivos puede considerarse como mezcla de los métodos hacia delante y hacia atrás. Comienza seleccionando, al igual que el método hacia delante, la variable independiente que -además de superar los criterios de entrada expresados- tiene el valor de correlación -absoluto- más alto con la dependiente. Posteriormente, si existe, se selecciona la variable independiente que, además de superar los criterios de entrada, posee el coeficiente de correlación parcial más alto (en valor absoluto). Cada vez que se incorpora una nueva variable, las previamente seleccionadas son, al igual que en el método hacia atrás, nuevamente evaluadas para determinar si siguen cumpliendo o no los criterios de salida, siendo expulsadas del modelo si los cumple.

El proceso se detiene cuando no quedan variables independientes que superen los criterios de entrada y las seleccionadas no cumplen los criterios de salida (Etxeberria Murgiondo, 2007; Pérez López, 2009). Por último, señalar que, en este método de selección, la ratio entre observaciones y variables independientes para asegurar que el modelo no está sobre-ajustado y es generalizable es de 50 a 1 (Hair y Gómez Suárez, 2010).

Este tipo de métodos son calificados a menudo en la literatura como métodos *greedy*, que reemplazan la búsqueda de un óptimo global por consideraciones sucesivas de óptimos locales, con lo cual no garantizan la mejor

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

solución y ni siquiera la misma entre sus distintas variantes.

Por otro lado, otra desventaja de peso es su inestabilidad en el sentido de que pequeños cambios en el conjunto de datos pueden producir grandes modificaciones en los resultados, en particular en las variables seleccionadas (Breiman, 1996). Esto se debe principalmente a que realizan un proceso discreto de exploración del espacio de modelos (cada variable es seleccionada o descartada).

Otras consecuencias indeseadas que a menudo se le achacan a este método es que la R^2 obtenida suele estar muy sesgada hacia arriba, que los intervalos de confianza son, incorrectamente, demasiado estrechos, que proporciona coeficientes sesgados y excesivamente grandes, y que no soluciona, en caso de existir, el problema la colinealidad entre las variables predictoras.

Métodos de Regresión Regularizada: Ridge y LASSO (Least Absolute Shrinkage and Selection Operator)

Estas técnicas fueron propuestas originalmente en los años setenta, como un método para lidiar con el problema de colinealidad en modelos lineales estimado por mínimos cuadrados, incluso en entornos $p < n$ (Hoerl y Kennard, 1976).

Tal y como se señalaba en la fórmula de estimación por mínimos cuadrados $\hat{\beta}^{mco} = (X^tr X)^{-1} X^tr Y$, la potencial

inestabilidad de $\hat{\beta}^{mco}$ podría aliviarse agregando una pequeña constante $k > 0$ a cada término de la diagonal de $X^{tr}X$ antes de invertir la matriz. Este proceso es el denominado como *estimador ridge*.

$$\hat{\beta}^{ridge}(k) = (X^{tr}X + kI_p)^{-1} X^{tr}Y$$

El principal problema de la Regresión *Ridge* es la determinación del valor de k más adecuado. La elección de este parámetro supone encontrar un balance entre los componentes de sesgo y variancia del error cuadrático medio. En este sentido (y asumiendo un modelo lineal), cuanto mayor es k más grande es el sesgo, pero menor es la variancia del estimador, y la determinación final implicará un compromiso entre ambos términos (Izenman, 2013).

Un método de aproximación inicial consiste en graficar los coeficientes de regresión estimados en función de k , y elegir el menor valor del parámetro para el cual se estabilizan dichos coeficientes. Un método más automático, pero intensivo computacionalmente, consiste en estimar k mediante validación cruzada.

En general, la Regresión *Ridge* produce predicciones más precisas que los modelos obtenidos por MCO + selección "clásica" de variables. Sin embargo, si bien al aumentar k (mayor penalización) los coeficientes estimados se contraen hacia cero, ninguno de ellos vale exactamente cero por lo cual no se produce selección de variables.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

Todas las variables originales permanecen en el modelo final.

También motivado por el objetivo de encontrar una técnica de regresión lineal que fuera estable pero que realizara selección de variables, (Tibshirani, 1996) propuso LASSO (Least Absolute Shrinkage and Selection Operator), una técnica de regresión lineal regularizada, como Ridge, con una leve diferencia en la penalización.

Como en Ridge, los coeficientes β_j se contraen hacia cero con la diferencia de que algunos de ellos se anulan. Por lo que, LASSO produce estimación y selección de variables en forma continua y simultánea.

Métodos locales no paramétricos

A menudo en la construcción de modelos espaciales de carácter ambiental se suele asumir una respuesta unimodal y simétrica a los valores, algo que a menudo queda refutado (Austin y Smith, 1989; Yee y Mitchell, 1991). Este tipo de distribuciones multimodales o sesgadas a veces se tratan con funciones polinómicas de orden alto, pero esta estrategia aumenta en gran medida el riesgo de aparición de patrones que sólo son aplicables a los datos de aprendizaje en la formación de los modelos donde funcionan casi perfectamente, pero tienen una mala capacidad predictiva con nuevos datos.

Como se ha señalado, en los modelos de regresión lineal se asume que las relaciones entre la respuesta y los predictores son lineales, sin embargo, los efectos del mundo real son generalmente más complejos.

Este enfoque parte de la hipótesis de que, al modelar distribuciones de parámetros ambientales, como es el caso de la biomasa, la respuesta está relacionada con las variables predictoras de una manera no necesariamente lineal y local.

Los modelos locales no paramétricos pueden resultar adecuados bajo esta hipótesis, ya que utilizan una estrategia de selección y reducción de variables locales y son lo suficientemente flexibles como para permitir relaciones no lineales. De este tipo se ha testado el modelo MARS (Multivariate Adaptive Regression Splines, (Friedman, 1991)).

MARS es una herramienta de modelado flexible que automatiza la construcción de modelos predictivos para variables dependientes tanto continuas como binarias. Combina la regresión lineal clásica, la construcción matemática de *splines* y la división recursiva binaria para producir un modelo local donde las relaciones entre la respuesta y los predictores pueden ser lineales o no lineales. Para ello, MARS se aproxima a la función subyacente a través de un conjunto de *regresiones lineales adaptativas* por subconjuntos denominadas *funciones de base* (BF).

Los cambios en la pendiente de esas funciones de base ocurren en los puntos llamados nodos (*knots*). De este modo la línea de regresión se desdobra en estos, los cuales marcan el final de una región de datos y el comienzo de otra con diferente comportamiento de la función. Los nodos se establecen por pasos sucesivos hacia adelante y hacia atrás.

Estimación de la biomasa potencial disponible de *Pinus radiata*
para usos energéticos en Bizkaia a partir de datos LiDAR
aerotransportados. Propuesta metodológica

En primer lugar, se produce un modelo que claramente produce un sobreajuste de los datos. En etapas posteriores, los nodos que contribuyen menos a la eficiencia del modelo se descartan por pasos hacia atrás. El mejor modelo se selecciona mediante validación cruzada; proceso que aplica una penalización a cada término (*knot*) añadido al modelo para mantener valores de baja complejidad (menor cantidad de variables predictoras).

Esencialmente construye modelos flexibles ajustando regresiones lineales por piezas; es decir, la no linealidad de un modelo se aproxima mediante el uso de pendientes de regresión separadas en distintos intervalos del espacio de las variables predictoras.

Este enfoque es capaz de revelar patrones de datos y relaciones entre variables que son difíciles de detectar mediante otros métodos.

Dada una variable objetivo y un conjunto de variables predictoras candidatas, MARS automatiza los aspectos del desarrollo y el despliegue del modelo, incluyendo:

- Selección de variables al separar variables de predicción relevantes de irrelevantes.
- Transformar variables predictoras que muestren una relación no lineal con la variable dependiente.
- Determinar interacciones entre variables predictoras.
- Realiza pruebas automáticas para evitar sobreajustes.

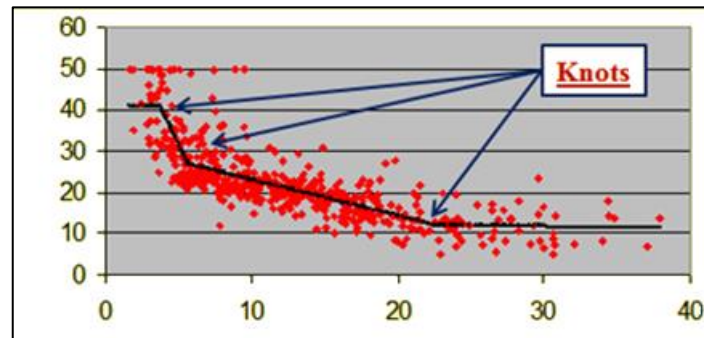


Figura 32 Modelo MARS (Fuente: <http://media.salford-systems.com/pdf/spm7/IntroMARS.pdf>)

Las variables a utilizar y los puntos finales de los intervalos para cada variable se encuentran a través de un procedimiento de búsqueda intensivo. Además de buscar variables una por una, MARS también busca interacciones entre variables, permitiendo considerar cualquier grado de interacción.

El modelo óptimo se selecciona en un proceso de dos etapas:

En la primera etapa MARS comienza con un modelo que consiste sólo en la ordenada en el origen de la recta de regresión (que es la media de los valores de respuesta). Luego agrega repetidamente funciones de base en pares al modelo. En cada paso encuentra el par de funciones de base que proporciona la máxima reducción en el error residual de la suma de los cuadrados. Las dos funciones de base del par son idénticas, excepto que para cada función se utiliza un lado diferente de una función bisagra en espejo. Cada nueva función de base consiste en un término que ya está en el modelo (que podría ser el término de

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

interceptación) multiplicado por una nueva función de bisagra. Una función de bisagra se define por una variable y un nudo, por lo que, para agregar una nueva función de base, MARS debe buscar entre todas las combinaciones siguientes:

- o términos existentes (llamados *términos padres* en este contexto)
- o todas las variables (para seleccionar una para la nueva función de base)
- o todos los valores de cada variable (para el nudo de la nueva función de bisagra).

Para calcular el coeficiente de cada término MARS aplica una regresión lineal sobre los términos. Este proceso de adición de términos continúa hasta que el cambio en el error residual es demasiado pequeño para continuar o hasta que se alcanza el número máximo de términos. El usuario es quien especifica el número máximo de términos antes de que comience el proceso de creación del modelo. La búsqueda en cada paso se realiza por *fuerza bruta*.

En el segundo paso procede con una técnica de paso hacia atrás. El pase hacia adelante generalmente construye un modelo de sobreajuste, tiene un buen ajuste a los datos usados para construir el modelo, pero no generalizará bien a los nuevos datos. Para construir un modelo con mejor capacidad de generalización procede con el paso hacia atrás en el que elimina los términos uno por uno, eliminando el término menos efectivo en cada paso hasta encontrar el mejor submodelo. Los subconjuntos de

modelos se comparan usando el criterio **GCV** (Generalized Cross Validation).

El GCV es una forma de regularización: balancea la bondad de ajuste con la complejidad del modelo. Busca estimar el desempeño de un modelo con nuevos datos, no en los datos de entrenamiento. Los datos nuevos usualmente no están disponibles en el momento de la construcción del modelo, por lo que en su usa GCV para estimar el rendimiento de los nuevos datos.

La suma de los cuadrados de los residuos (RSS) de los datos de entrenamiento es inadecuada para comparar los modelos, ya que el RSS siempre aumenta a medida que los términos MARS son eliminados, es decir, si se usan los RSS para comparar modelos, el método de paso hacia atrás elegirá siempre el modelo más grande, pero el modelo más grande generalmente no tiene el mejor rendimiento de generalización.

La fórmula para el GCV es

$$GCV = RSS / (N * (1 - N^{\circ} \text{ Efectivo de parámetros} / N)^2)$$

Donde RSS es la suma de los cuadrados de los residuos medida en los datos de entrenamiento y N es el número de observaciones.

El "N° Efectivo de parámetros" en el contexto de MARS se obtiene desde:

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

$$N^{\circ} \text{ Efectivo de parámetros} = N^{\circ} \text{ de términos MARS} + \text{Penalización} * (N^{\circ} \text{ de términos MARS} - 1) / 2$$

Donde el término " $(N^{\circ} \text{ de términos MARS} - 1) / 2$ " representa el número de nudos de la función de bisagra, por lo que la fórmula penaliza la adición de nudos.

Así, la fórmula GCV ajusta el RSS de los datos de entrenamiento para tener bajo control la flexibilidad del modelo. La flexibilidad se penaliza porque los modelos que son demasiado flexibles modelarán el ruido específico en los datos en lugar de sólo la estructura sistemática de los datos.

La validación cruzada generalizada se denomina así porque utiliza una fórmula para aproximar el error que se determinaría mediante una validación cruzada tipo LOO (Leave One Out).

Resumiendo, en la primera etapa, MARS construye un modelo excesivamente grande añadiendo "Funciones de base". A medida que se añaden las funciones de base, el modelo se vuelve más flexible y más complejo, y el proceso continúa hasta que se haya alcanzado un número máximo de funciones de base especificado por el usuario. En la segunda etapa, las funciones básicas se eliminan en orden de menor contribución al modelo hasta que se encuentre un modelo óptimo.

MARS hace una selección automática de variables (lo que significa que incluye variables importantes en el modelo

y excluye las sin importancia). Sin embargo, hay que tener en cuenta que la selección de variables puede resultar algo arbitraria especialmente en presencia de colinealidad y concurvidad ((Morlini, Statistica, Economia, & Parma, n.d.)

Con los modelos MARS, como con cualquier regresión no paramétrica, los intervalos de confianza de los parámetros y otras comprobaciones del modelo no se pueden calcular directamente (a diferencia de los modelos de regresión lineal).

Los posibles problemas de multicolinealidad que se pueden presentar empleando estas técnicas se pueden abordar mediante estrategias como utilizar en RLM las funciones de base generadas en MARS sin interacciones como variables predictoras.

7.3. Modelos de regresión no lineales

La regresión curvilínea o regresión no lineal es una técnica estadística que permite predecir una variable dependiente a partir de una o varias independientes cuando la relación entre estos dos tipos de variables no se ajusta a una línea recta. Como en la regresión lineal, la variable dependiente y las independientes deben estar medidas en escala métrica, siendo posible incluir variables categóricas entre las independientes mediante procedimientos especiales. Como resultado de la regresión no lineal se obtiene una ecuación que predice la variable dependiente proporcionando información sobre las variables independientes que la explican (Ebdon, 1982; Etxeverría, 1999).

Encontrar una ecuación que describa una relación no lineal es bastante más complejo que en el caso de la regresión lineal simple o múltiple, resultando más sencillo aplicar transformaciones a los datos de ambas variables para "linealizar" la relación y proceder mediante el análisis de regresión lineal (Ebdon, 1982; Etxeberria Murgiondo, 2007). Es por ello que en la presente investigación solamente se ajustaron modelos univariantes a partir de la función "estimación curvilínea" integrada en SPSS. Este programa ofrece, al efecto, once modelos diferentes de estimación de una variable dependiente a partir del uso de una única variable independiente, proporcionando todos los métodos gráficos y estadísticos necesarios para evaluar la bondad del ajuste obtenido y su significación estadística (Pérez López, 2009).

Tabla 19 Modelos de regresión curvilínea

Modelo de regresión	Expresión matemática
Lineal	$Y = \beta_0 + \beta_1 X$
Logarítmico	$Y = \beta_0 + \beta_1 \ln(X)$
Inverso	$Y = \beta_0 + \beta_1 (1/X)$
Cuadrático	$Y = \beta_0 + \beta_1 X + \beta_2 X^2$
Cúbico	$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 X^3$
Potencia	$Y = \beta_0 (X^{\beta_1})$
Compuesto	$Y = \beta_0 (\beta_1^X)$
S	$Y = e^{\beta_0 + \beta_1 (1/X)}$
Logístico	$Y = 1 / (1/c + \beta_0 \beta_1^X)$
Crecimiento	$Y = e^{\beta_0 + \beta_1 X}$
Exponencial	$Y = \beta_0 e^{\beta_1 X}$

donde Y es la variable dependiente, X la independiente, β_0 la constante, β_1 el coeficiente de regresión y c el parámetro que refiere al límite superior de la función logística (tabla 19).

Valoración de la exactitud de la predicción

Para comprobar la bondad del ajuste obtenido mediante la regresión curvilínea se utilizan los mismos estadísticos que en el caso de la lineal. Así, para cuantificar la proporción de la variable dependiente explicada por la independiente se utiliza R^2 , mientras que Se , $RMSE$ y ofrecen una visión global de la exactitud de la predicción.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

El examen de la significación estadística del modelo se hace también mediante el cálculo del estadístico F y el contraste de su significación ($p < 0,05$), mientras que el estadístico t (Ecuaciones 11 y 12) y su prueba de significación se utilizan para contrastar la validez del coeficiente de regresión y de la constante del modelo ajustado.

Método de ajuste del modelo

Al tratarse de modelos univariantes, la especificación del modelo de regresión se puede hacer únicamente de forma confirmatoria, de tal manera que la única cuestión a resolver es qué tipo de modelo curvilíneo se adapta mejor para describir la relación entre la variable dependiente y la independiente. Para ello resulta fundamental el análisis del patrón de puntos del gráfico de dispersión definido por las dos variables que previamente habían mostrado la existencia de una relación no lineal. En función de este patrón se explora la utilidad de dos o tres de los modelos recogidos en la Tabla 19. Los resultados obtenidos son analizados y comparados utilizando para ello los estadísticos señalados en el anterior punto (R^2 , Se , etcétera.). Finalmente, el modelo propuesto es aquel que, siendo significativo a todos los niveles, explica la mayor cantidad de varianza de la variable dependiente y presenta el error de estimación menor.

7.4. Métodos de validación de los modelos

En los casos en los que no se dispone de un conjunto de datos independiente para realizar la validación, se puede

subdividir la muestra inicial en varios subconjuntos, dando lugar a una validación cruzada. Para validar los modelos de regresión existen varias alternativas:

1.- Una de las aproximaciones más empleadas consiste en dividir el conjunto de la muestra en dos subconjuntos de manera aleatoria, empleando un porcentaje que suele rondar el 80% para el ajuste (muestra de aprendizaje), y el 20% restante para la validación (test). Esta división de la muestra se hace de forma aleatoria para que cualquier ordenación de los datos no afecte a los procesos de estimación y validación. Además, para valorar la robustez de los modelos, esta división aleatoria se puede realizar de manera recursiva, calculándose el respectivo modelo y la validación en cada una de ellas. La similitud de los coeficientes y de los estadísticos usados para valorar la exactitud del modelo serán los indicadores de que el modelo es independiente de la selección aleatoria de los datos.

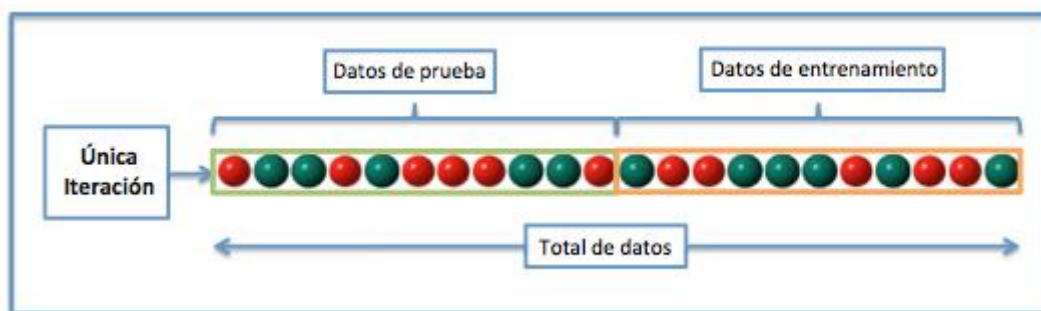


Figura 33 Validación Cruzada (Método de retención) By Joan.domenech91 (Own work) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], via Wikimedia Commons

2.- La validación cruzada de k veces: El conjunto de datos original se divide aleatoriamente en k subconjuntos mutuamente exclusivos (pliegues) de aproximadamente el

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

mismo tamaño. Se elimina un subconjunto y se utilizan los casos restantes para generar un modelo. Este modelo se aplica posteriormente a la sección eliminada y se calcula su error empírico. El proceso se repite con los nueve subconjuntos restantes y el error empírico medio se utiliza como la estimación final del error total del modelo.

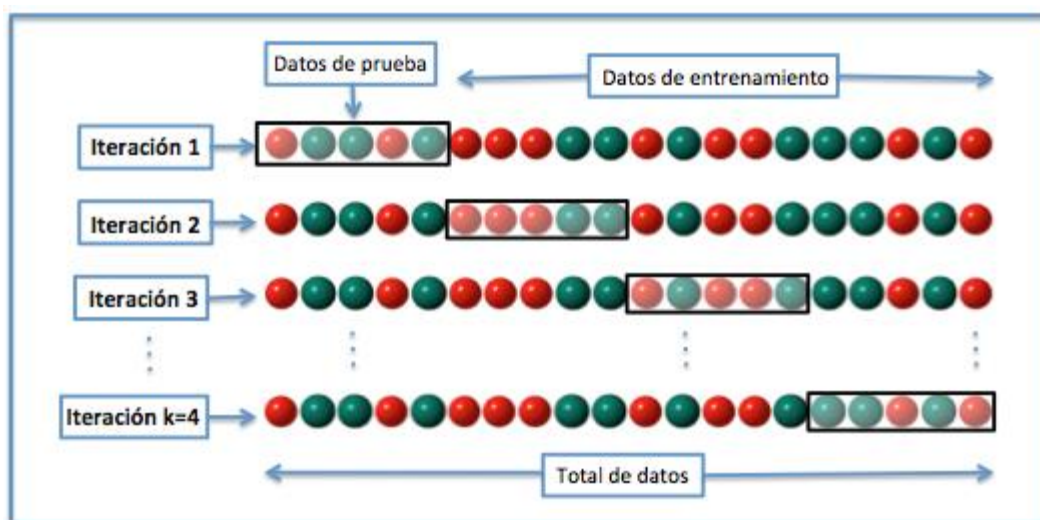


Figura 34 Validación cruzada de K iteraciones con K=4 By Joan.domenech91 (Own work) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], via Wikimedia Commons

El caso más extremo sería el "leave-one-out" o el "error de la validación cruzada dejando uno fuera" (Barron, 2014), que implica separar los datos de forma que para cada iteración tengamos una sola muestra para los datos de prueba y todo el resto conformando los datos de entrenamiento en el que se realizan tantas iteraciones como muestras (N) tenga el conjunto de datos. De forma que para cada una de las N iteraciones se realiza un cálculo de error. El resultado final lo obtenemos realizando la media aritmética de los N valores de errores obtenidos, según la fórmula:

$$E = \frac{1}{K} \sum_{i=1}^k E_i$$

La ausencia de superposición entre los conjuntos de datos usados para el ajuste del modelo y los empleados para calcular el criterio de validación garantiza la validez de esta práctica. Este método exige más cálculos que una validación simple, pero aprovecha todas las observaciones disponibles para el ajuste del modelo.

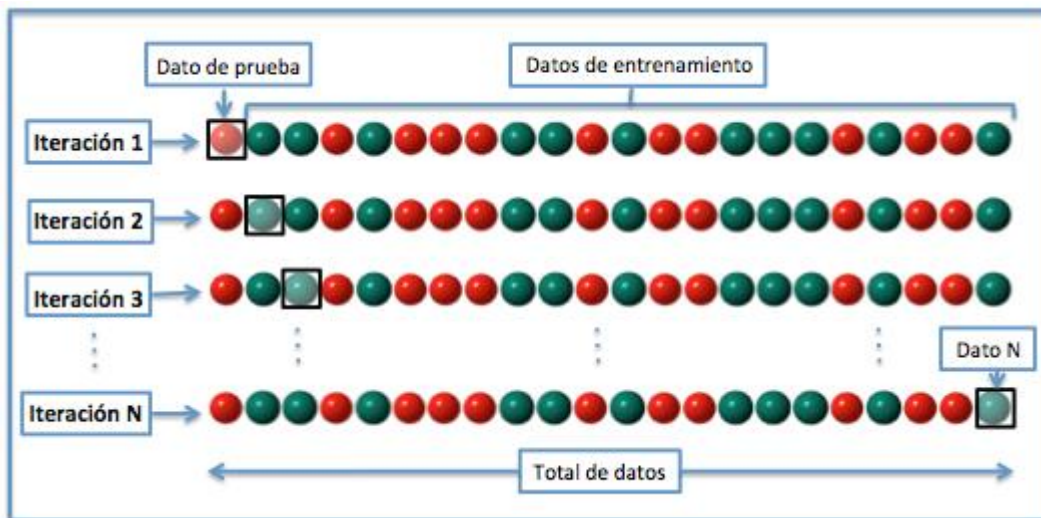


Figura 35 Validación cruzada dejando uno fuera (LOO) By Joan.domenech91 (Own work) [CC BY-SA 3.0 (<http://creativecommons.org/licenses/by-sa/3.0>)], via Wikimedia Commons

El procedimiento consta de los siguientes pasos:

1. Subdividir el conjunto de datos S_n en K subconjuntos de datos de tamaños aproximadamente iguales: $S_n^{(1)}, \dots, S_n^{(k)}$
2. Para k que va de 1 a K :
 - a. Ajustar el modelo a partir del conjunto de datos sin su k -ésima parte:

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

$$\frac{S_n}{S_n^{(k)}} = S_n^{(1)} U \dots U_{S_n^{(k-1)}} U_{S_n^{(k+1)}} U \dots U_{S_n^{(k)}}$$

- b. Calcular el criterio de validación de dicho modelo ajustado tomando la parte restante $S_n^{(k)}$
3. Calcular el promedio de los K criterios de validación calculados.

Los criterios de validación más utilizados son:

- El Error de la Raíz Cuadrada de la Media (RMSE)
- Error porcentual absoluto medio (MAPE)
- R^2 ajustada

8. ANÁLISIS DE LOS RESULTADOS DE ESTIMACIÓN

En este capítulo se describe la aplicación de las técnicas descritas en el anterior y se analizan los resultados obtenidos.

Tras obtener las métricas nubes de puntos LiDAR para cada parcela seleccionada, se ha desarrollado una estrategia para identificar el conjunto de variables con más capacidad de explicar la varianza y de predecir con el menor error posible la biomasa. El objetivo de esta fase de calibración del modelo es que, cumpliendo con todos los supuestos del análisis de regresión, el modelo elegido emplee el menor número de variables predictivas posibles (eficiencia del modelo) maximizando a la vez el ajuste (R^2) y minimizando el error (SE). Este proceso se ha abordado siguiendo el siguiente esquema metodológico:

El primer lugar, previo a análisis posteriores, se han analizado las principales estadísticas descriptivas de cada variable para detectar posibles anomalías en el conjunto de los datos y establecer una primera visión de las características de cada variable (tabla 20). Al no detectarse ningún valor anómalo en ninguna de las métricas LiDAR se procedió con el siguiente paso, que consiste en testar la normalidad de las variables.

Tabla 20 Codificación de la variables del estudio

VARIABLES	
Acrónimo	Nombre
BIOA	Biomasa_aerea
EMAX	Elev maximum
EMEA	Elev mean
EMOD	Elev mode
ESTD	Elev stddev
EVAR	Elev variance
ECV	Elev CV
EIQ	Elev IQ
ESKE	Elev skewness
EKUR	Elev kurtosis
EAAD	Elev AAD
EMME	Elev MAD median
EMMOD	Elev MAD mode
EL1	Elev L1
EL2	Elev L2
EL3	Elev L3
EL4	Elev L4
ELCV	Elev L CV
ELSK	Elev L skewness
ELKU	Elev L kurtosis
EP01	Elev P01
EP05	Elev P05
EP10	Elev P10
EP20	Elev P20
EP25	Elev P25
EP30	Elev P30
EP40	Elev P40
EP50	Elev P50
EP60	Elev P60
EP70	Elev P70
EP75	Elev P75
EP80	Elev P80
EP90	Elev P90
EP95	Elev P95
EP99	Elev P99
CRR	Canopy relief ratio
ESQM	Elev SQRT mean SQ

ECMC	Elev CURT mean CUBE
P1RA4	Percentage first returns above 4
PARA4	Percentage all returns above 4
AA4T1	(All returns above 4) / (Total first returns)*100
P1AM	Percentage first returns above mean
P1AMO	Percentage first returns above mode
PAAM	Percentage all returns above mean
PAAMO	Percentage all returns above mode
AAMT1	(All returns above mean) / (Total first returns)*100
AAMOT1	(All returns above mode) / (Total first returns)*100
ES9	Elev strata (below 9.00) return proportion
ES914	Elev strata (9.00 to 14.00) return proportion
ES1419	Elev strata (14.00 to 19.00) return proportion
ES1924	Elev strata (19.00 to 24.00) return proportion
ES2429	Elev strata (24.00 to 29.00) return proportion
ES3439	Elev strata (34.00 to 39.00) return proportion
ES3944	Elev strata (39.00 to 44.00) return proportion
ES4954	Elev strata (49.00 to 54.00) return proportion

8.1. Análisis de la normalidad

La normalidad de las variables, incluida la variable dependiente, se ha comprobado mediante diferentes test (Shapiro-Wilk, Anderson-Darling, Lilliefors, Jarque-Bera y Kolmogorov-Smirnov) para descartar o, en caso de considerar necesario, transformar las variables que no superen al menos en uno de los test. La hipótesis nula H_0 de todos ellos supone la existencia de normalidad para una $p > 0,05$ (tabla 21).

Tabla 21 Pruebas de normalidad de las variables. (En rojo variables que no han superado al menos un test de normalidad)

Prueba Variable	Shapiro- Wilk	Anderson- Darling	Lilliefors	Jarque- Bera	Kolmogorov- Smirnov
BIOA_MONTERO	0,003	0,016	0,349	0,051	0,363
BIOA_CANGA	0,006	0,011	0,025	0,025	0,300
BIOA_PEREZ	0,018	0,029	0,019	0,061	0,273
EMAX	0,455	0,424	0,388	0,799	0,777
EMEA	0,028	0,070	0,445	0,115	0,812

EMOD	0,015	0,052	0,025	0,100	0,299
ESTD	0,447	0,231	0,310	0,655	0,722
EVAR	< 0,0001	< 0,0001	0,000	< 0,0001	0,052
ECV	0,002	0,010	0,001	< 0,0001	0,109
EIQ	0,006	0,084	0,164	< 0,0001	0,582
ESKE	0,015	0,007	0,041	0,035	0,358
EKUR	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001
EAAD	0,331	0,163	0,271	0,365	0,691
EMME	0,002	0,085	0,187	< 0,0001	0,609
EMMOD	< 0,0001	< 0,0001	0,002	< 0,0001	0,121
EL1	0,028	0,070	0,445	0,115	0,812
EL2	0,508	0,265	0,392	0,762	0,780
EL3	0,000	< 0,0001	0,000	< 0,0001	0,058
EL4	0,000	0,000	< 0,0001	0,004	0,042
ELCV	0,041	0,107	0,015	0,007	0,247
ELSK	0,006	0,007	0,088	0,043	0,474
ELKU	0,016	0,001	0,001	0,110	0,081
EP01	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001
EP05	< 0,0001	< 0,0001	< 0,0001	< 0,0001	0,001
EP10	< 0,0001	< 0,0001	< 0,0001	0,000	0,028
EP20	0,000	0,000	0,005	0,013	0,172
EP25	0,001	0,001	0,007	0,024	0,192
EP30	0,001	0,003	0,004	0,034	0,162
EP40	0,005	0,014	0,054	0,057	0,398
EP50	0,019	0,056	0,509	0,093	0,847
EP60	0,042	0,112	0,381	0,124	0,773
EP70	0,077	0,155	0,340	0,163	0,745
EP75	0,097	0,185	0,487	0,183	0,835
EP80	0,099	0,185	0,436	0,204	0,807
EP90	0,129	0,239	0,393	0,254	0,781
EP95	0,099	0,165	0,591	0,248	0,885
EP99	0,047	0,111	0,276	0,249	0,695
CRR	0,026	0,050	0,223	0,103	0,647
ESQM	0,051	0,117	0,543	0,143	0,863
ECMC	0,075	0,160	0,583	0,172	0,882
P1RA4	0,003	0,003	0,007	0,030	0,193
PARA4	0,003	0,003	0,006	0,014	0,179
AA4T1	0,758	0,507	0,393	0,618	0,781
P1AM	0,643	0,272	0,418	0,540	0,796
P1AMO	0,117	0,506	0,791	0,023	0,958
PAAM	0,074	0,026	0,055	0,149	0,399
PAAMO	0,049	0,336	0,373	0,002	0,767
AAMT1	0,794	0,665	0,610	0,646	0,894
AAMOT1	0,001	0,102	0,235	< 0,0001	0,658
ES9	< 0,0001	< 0,0001	0,000	0,007	0,064
ES914	< 0,0001	< 0,0001	< 0,0001	0,000	0,027
ES1419	< 0,0001	< 0,0001	0,007	0,017	0,187
ES1924	< 0,0001	< 0,0001	< 0,0001	0,002	0,004
ES2429	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001
ES3439	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001
ES3944	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001
ES4954	< 0,0001	< 0,0001	< 0,0001	< 0,0001	< 0,0001

Como se puede observar en la tabla 21, son varias las variables que no han superado al menos uno de los test de normalidad, por lo que, en vista de que la mayoría de las variables han superado los test de normalidad, se ha optado por dejarlas fuera del grupo de variables candidatas a formar parte de los modelos de estimación.

En tercer lugar, se ha analizado la correlación y la fuerza de las relaciones lineales entre cada variable explicativa y la variable biomasa. Dado que el conjunto de variables predictoras candidatas era muy elevado y ante posibles problemas de multicolinealidad y sobreajuste del modelo, el siguiente paso en la selección de variables ha consistido en comprobar cómo se relacionan las variables predictoras con la variable dependiente y las correlaciones cruzadas entre todas las variables candidatas.

Para ello se generó una matriz de correlación con todas las variables descartando aquellas que, atendiendo al valor del momento-producto de Pearson, presentan índices menores a +/- 0,5 o no significativas al 5%.

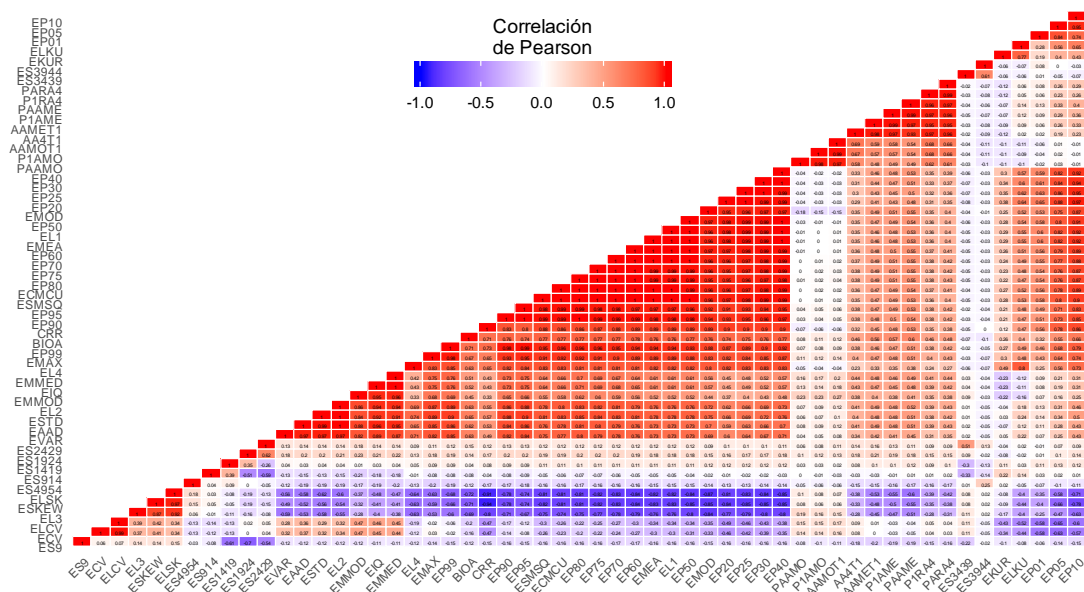


Figura 36 Coeficiente de correlación de Pearsons entre variables

Las variables descartadas fueron (tabla 22):

Tabla 22 Variables descartadas por falta de significación estadística en el C. de Correlación de Pearsons

Variable
ECV
ELKU
P1RA4
PARA4
AA4T1
ES3944
ES4954
PAAMO
AAMOT1

Con las restantes se ha procedido a hacer una segunda selección de variables en base a los criterios de

eficiencia, minimización del error y maximización del grado de ajuste que procedemos a explicar a continuación.

En una segunda aproximación se procedió a verificar las relaciones lineales entre la biomasa y las variables predictoras.

Primero se calcularon los diagramas de dispersión entre la variable biomasa y cada una de las variables significativas tal y como se recomienda en (Picard, Saint-André y Henry, 2012). (Vid. ANEXO X).

En segundo lugar, se calcularon las R^2 y el Error Medio Cuadrático de todos los modelos de regresión con una variable predictora (tabla 23).

Tabla 23 Grado de ajuste y error de los modelos de regresión con una sola variable

Variabes	R2	SE	MAD
EP60	0,63	2587,84	45,21
EP50	0,62	2609,25	45,05
EP70	0,62	2623,86	45,60
EP75	0,62	2639,18	45,77
ECMC	0,62	2651,96	46,02
EP80	0,62	2662,93	46,02
ESQM	0,62	2665,56	45,33
EP40	0,61	2671,25	45,95
EL1	0,61	2713,43	46,02
EMEA	0,61	2713,43	46,02
EP90	0,61	2727,82	46,83
EMOD	0,61	2736,87	45,63
EP30	0,60	2770,65	44,91
EP25	0,59	2848,51	46,01
EP95	0,59	2870,25	47,84
EP20	0,57	3006,52	46,97
ELSK	0,54	3162,78	47,95

EP99	0,54	3176,87	48,19
ESKE	0,53	3242,67	48,27
CRR	0,52	3310,98	49,69
EMAX	0,48	3590,08	50,06
EL3	0,48	3592,39	52,03
PAAM	0,46	3719,85	52,28
ESTD	0,46	3742,56	53,52
P1AM	0,43	3973,89	54,12
EL2	0,43	3986,09	56,65
EAAD	0,41	4124,63	56,96
AAMT1	0,40	4128,23	57,48
EIQ	0,27	5044,61	58,83
EMME	0,27	5077,15	58,71

Como se puede observar ninguna variable alcanza por si sola valores de coeficiente de determinación que pudieran considerarse aceptables (al menos 0,7).

También se observa que las variables que se encuentran en los primeros puestos son variables relacionadas con la altura del dosel, que, como hemos visto, presentan entre ellas altas correlaciones. También destacan entre los primeros puestos dos variables relacionadas con la densidad de la cubierta como son ECMC y ESQM. En un tercer grupo podemos observar la presencia de variables relacionadas con la estructura vertical como son las variables que cuantifican el sesgo en la distribución de valores de los datos LiDAR en parcela: ELSK y ESK. Por último, aparece otro grupo de variables con una menor capacidad explicativa de la varianza (menos del 50%) entre las que destacan las relacionadas con la densidad de cubierta: PAAM, P1AM y AAMT1.

La transformación logarítmica (Tabla 24) tampoco mejora estos resultados, por lo que se decidió no transformarlas, a pesar de que en varios estudios ha dados buenos resultados, tanto en lo que respecta a la mejora de la R^2 como a la hora de resolver problemas como la heterocedasticidad (Torre Tojal, 2016; Valbuena-Rabadán, 2014; Valbuena-Rabadán, Santamaría-Peña y Sanz-Adán, 2016).

Tabla 24 Grado de ajuste y error de los modelos de regresión con una sola variable. Variables log-transformadas

Variables Log-transformadas	R²	SE	MAD
LN_EP50	0,64	0,094	0,34
LN_EP60	0,64	0,094	0,34
LN_EP70	0,63	0,095	0,34
LN_EP40	0,63	0,096	0,35
LN_ESQM	0,63	0,096	0,35
LN_ECMC	0,63	0,096	0,35
LN_EP75	0,63	0,096	0,35
LN_EL1	0,63	0,097	0,35
LN_EMEA	0,63	0,097	0,35
LN_EP80	0,63	0,097	0,35
LN_EP90	0,62	0,099	0,35
LN_EP30	0,62	0,099	0,35
LN_EP25	0,61	0,102	0,35
LN_EP95	0,60	0,103	0,36
LN_EMOD	0,60	0,103	0,34
LN_EP20	0,59	0,107	0,36
LN_EP99	0,57	0,112	0,36
LN_EMAX	0,53	0,122	0,37
LN_CRR	0,52	0,124	0,38
LN_ESTD	0,50	0,130	0,38
LN_EL2	0,48	0,136	0,39
LN_EAAD	0,46	0,140	0,39
LN_PAAM	0,44	0,145	0,42
LN_P1AM	0,42	0,150	0,42
LN_AAMT1	0,41	0,153	0,42
LN_EMME	0,37	0,164	0,42
LN_EIQ	0,37	0,164	0,42

Esta primera aproximación pone de manifiesto la necesidad de continuar buscando combinaciones de variables para intentar mejorar la capacidad predictiva del modelo puesto que una sola variable aporta poca capacidad predictora al modelo.

8.2. Selección de variables

En este punto se ha considerado necesario recurrir a métodos semiautomáticos de selección de variables, habida cuenta de que las combinaciones posibles de en el modelo resultan inabordables de forma manual: p.ej. para combinaciones de hasta 5 variables las variaciones superarían los 17 millones (V_{30-5}). El máximo de variables a incluir en una muestra de 151 casos para evitar aceptar una falsa hipótesis nula estaría en torno a 7-8 variables independientes para una R^2 de 0,8 (Cohen, 1988).

En este sentido las alternativas son variadas, tal como se ha comentado más arriba. Estos son los métodos que se han ensayado y comparado utilizando para ello los siguientes programas estadísticos (Minitab, SPSS, R, Salford Predictive Modeler, XLSTAT):

- Métodos Stepwise (Incluida en todos los paquetes estadísticos. Ensayo con SPSS)
- Método de mejores subconjuntos con distintos criterios:
 - o De prevención del sobreajuste (SPSS)
 - o R^2 Ajustada (SPSS, XLSTAT, Minitab)
 - o Error Medio Cuadrático (XLSTAT)

- Métodos LASSO y Ridge (R - Lasso y Salford Predictive Modeler)
- Método de selección de variables MARS (R - "earth package")

Para todas las alternativas se han analizado en primera instancia los siguientes parámetros que sirven para para evaluar el grado de **ajuste**, de **sobreajuste**, la **exactitud**, la **multicolinealidad** y la **significación** de los coeficientes del modelo.

El **grado de ajuste** se ha evaluado considerando los valores de R^2 y R^2 *ajustada*. Se considera que los valores deben acercarse a 1 y *no ser inferiores a 0,7*.

Para evaluar el **grado de sobreajuste** se han tenido en cuenta la R^2 *predicha* mediante validación cruzada LOO en la que se calcula el R^2 *ajustada* eliminando sistemáticamente cada observación del conjunto de datos, estimando la ecuación de regresión y determinando qué tan bien predice el modelo la observación removida. Si el valor de la R^2 *predicha* difiere mucho de la R^2 *ajustada* del conjunto de entrenamiento, el modelo no podría considerarse generalizable fuera de la muestra, por lo que se consideraría aceptable una **desviación de entre un 3-5%** entre el ajuste de los datos de entrenamiento y los de validación.

Para evaluar la **exactitud** se han seleccionado las siguientes medidas de error:

- Error Estándar de la Estimación, SE

$$S_e = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y'_i)^2}{n-2}}$$

donde S_e es el error estándar de la estimación, Y_i es el valor real para la observación i , Y'_i es el valor pronosticado para la observación i y n es el número de casos.

Expresa la exactitud en las mismas unidades que los datos. Es por lo tanto una medida de exactitud dependiente de la escala y no puede utilizarse para hacer comparaciones entre conjuntos de datos que estén en diferentes escalas.

- Error porcentual absoluto medio (MAPE, Mean Absolute Percentage Error) Expresa la exactitud como un porcentaje del error.

$$\frac{\sum |(y_t - \hat{y}_t) / y_t|}{n} \times 100, (y_t \neq 0)$$

Los valores que se considerarían aceptables están entre un 0,2-0,25, es decir, que el pronóstico, en promedio está errado en un entre un 20 y un 25%.

Para evaluar la **multicolinealidad** se ha recurrido al Valor de Inflación de la Varianza (VIF) que, como ha quedado dicho, debería moverse en **valores entre 1-2**.

La **significación** de los coeficientes se ha evaluado a partir del estadístico t y sus correspondientes **valores críticos al 5%**.

En el caso de los modelos que superan los umbrales mínimos establecidos para cada parámetro se ha procedido a realizar las pruebas *ex post* que determinan la robustez del mismo: linealidad, normalidad, homocedasticidad e independencia.

8.2.1. Modelos con métodos *Stepwise*

De todos los modelos que han sido seleccionados mediante el método *stepwise* para cada variable dependiente, en los que las variables seleccionadas han entrado o salido en base a los criterios de probabilidad de F para entrar $\leq 0,050$ y probabilidad de F para eliminar $\geq 0,100$, los modelos que ofrece mayores ajustes son los que incluyen las variables: P1AM, ECMC, EL3, EL2, EMEA, EP40, CRR y EMME. Estas aparecen en el mejor modelo de las dependientes de Montero y de Canga. En el caso de Pérez-Cruzado se repiten las variables EP40, EL2 y CRR y aparecen PAAM, ESTD Y EP50 (Tablas 25,26,27,28), figuras 37,38,39).

Se trata de modelos en los que aparecen variables relacionadas directamente con métricas de altura de los retornos (percentiles de altura, L2 y desviación estándar) y dos relacionadas con la densidad de cabida cubierta P1AM Y PAAM.

Tabla 25 Método Stepwise

STEPWISE						
Mejor Modelo	R ²	R ² aj.	R ² Test	SE	ANOVA	
					f	p
MONTERO						
1	0,764	0,751	0,733	42,553	57,445	0,000
CANGA						
1	0,803	0,792	0,777	38,312	72,457	0,000
PÉREZ-CRUZADO						
1	0,774	0,764	0,747	37,684	81,977	0,000

Importancia del predictor

Destino: BIOA_MONT

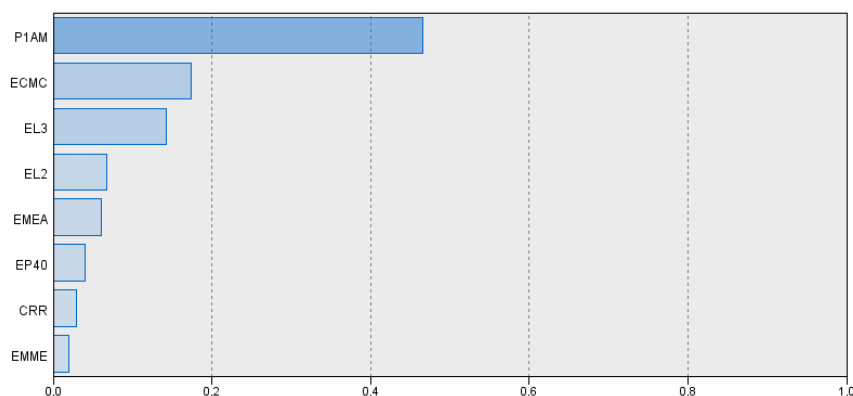


Figura 37 Importancia de los predictores para Montero (Método Stepwise)

Tabla 26 Estadísticos de los coeficientes para Montero (Stepwise)

Coeficientes ^a													
Modelo		Coeficientes no estandarizados		Coeficiente s estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
		B	Error estándar	Beta	t		Limite inferior	Limite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1	(Constante)	-5,388	27,535		-0,196	0,845	-59,819	49,044					
	EMEA	-183,553	55,849	-12,029	-3,287	0,001	-293,956	-73,149	0,729	-0,266	-0,134	0,000	8058,492
	EMME	-22,005	20,333	-0,236	-1,082	0,281	-62,199	18,189	0,514	-0,090	-0,044	0,035	28,498
	EL2	-214,025	48,131	-2,013	-4,447	0,000	-309,171	-118,879	0,637	-0,350	-0,181	0,008	123,334
	EL3	-305,211	72,690	-1,016	-4,199	0,000	-448,905	-161,517	-0,677	-0,332	-0,171	0,028	35,189
	EP40	-74,406	34,189	-5,017	-2,176	0,031	-141,990	-6,822	0,733	-0,180	-0,089	0,000	3197,277
	CRR	197,550	89,250	0,265	2,239	0,027	23,096	372,003	0,681	0,185	0,091	0,118	8,439
	ECMC	269,609	43,320	18,273	6,224	0,000	183,973	355,244	0,737	0,463	0,254	0,000	5185,668
	P1AM	3,160	0,375	0,425	8,425	0,000	2,419	3,901	0,663	0,577	0,344	0,654	1,530

a. Variable dependiente: BIOA_MONT

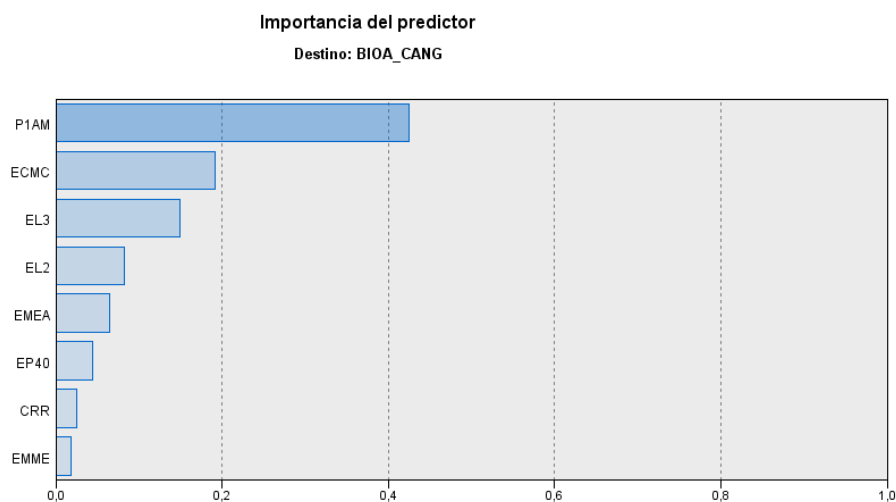


Figura 38 Importancia de los predictores para Canga (Método Stepwise)

Tabla 27 Estadísticos de los coeficientes para Canga (Stepwise)

Coefficientes^a

Modelo	Coeficientes no estandarizados		Coeficiente s estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar	Beta	t		Limite inferior	Limite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1 (Constante)	-21,326	24,791		-0,860	0,391	-70,334	27,681					
EMEA	-171,249	50,284	-11,381	-3,406	0,001	-270,651	-71,848	0,782	-0,275	-0,127	0,000	8058,492
EMME	-23,363	18,306	-0,254	-1,276	0,204	-59,552	12,825	0,520	-0,106	-0,048	0,035	28,498
EL2	-204,698	43,334	-1,953	-4,724	0,000	-290,382	-119,034	0,655	-0,369	-0,176	0,008	123,334
EL3	-282,147	65,446	-0,952	-4,311	0,000	-411,522	-152,773	-0,700	-0,340	-0,160	0,028	35,189
EP40	-71,097	30,781	-4,862	-2,310	0,022	-131,946	-10,247	0,787	-0,190	-0,086	0,000	3197,277
CRR	177,160	78,455	0,241	2,230	0,027	20,091	334,228	0,724	0,184	0,083	0,118	8,439
ECMC	255,689	39,003	17,573	6,556	0,000	178,588	332,791	0,788	0,482	0,244	0,000	5185,668
P1AM	2,843	0,338	0,388	8,420	0,000	2,176	3,511	0,651	0,577	0,313	0,654	1,530

a. Variable dependiente: BIOA_CANG

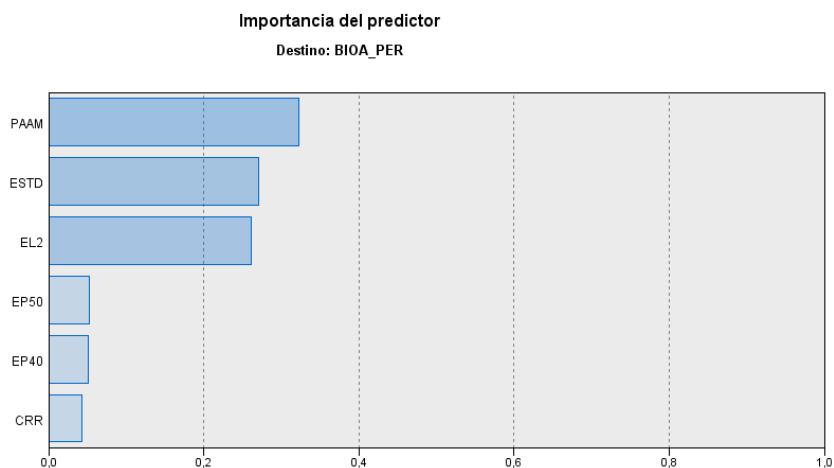


Figura 39 Importancia de los predictores para Pérez-Cruzado (Método Stepwise)

Tabla 28 Estadísticos de los coeficientes para Pérez-Cruzado (Stepwise)

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar	Beta	t		Límite inferior	Límite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1 (Constante)	-59,503	18,769		-3,170	0,002	-96,601	-22,405					
ESTD	243,038	36,312	4,590	6,693	0,000	171,265	314,812	0,648	0,487	0,265	0,003	298,998
EL2	-447,860	68,248	-4,627	-6,562	0,000	-582,757	-312,963	0,619	-0,480	-0,260	0,003	316,190
EP40	-63,561	21,836	-4,707	-2,911	0,004	-106,720	-20,401	0,789	-0,236	-0,115	0,001	1662,974
EP50	64,476	21,927	4,954	2,940	0,004	21,134	107,817	0,793	0,238	0,117	0,001	1804,722
CRR	193,753	73,165	0,286	2,648	0,009	49,137	338,370	0,738	0,215	0,105	0,135	7,396
PAAM	3,033	0,416	0,366	7,293	0,000	2,211	3,855	0,658	0,519	0,289	0,625	1,601

a. Variable dependiente: BIOA_PER

A pesar de que todos los modelos muestran que el estadístico f y el t de los coeficientes β son significativos y que los coeficientes de determinación y error son aceptables (tabla 25), las altas correlaciones que existen entre casi todas las variables, provocan que el valor diagnóstico de la colinealidad (VIF - Inflación de la Varianza) esté muy por encima de lo recomendable en todas las variables (tablas 26, 27, 28). Esto induce a pensar que estos modelos no puedan darse por buenos.

8.2.2. Modelos con Métodos de Mejores Subconjuntos

Este método comprueba "todos los modelos posibles" para seleccionar el mejor según el criterio de mejores subconjuntos. Los criterios pueden ser los mencionadas arriba de:

- R^2 corregida se basa en el ajuste del conjunto de entrenamiento, y se ajusta para penalizar modelos excesivamente complejos.
- Criterio de prevención sobreajustado que se basa en el ajuste (error cuadrado medio) del conjunto de prevención sobreajustado. El conjunto de prevención sobreajustado es una submuestra aleatoria de aproximadamente el 30% del conjunto de datos original que no se utiliza para enseñar el modelo.

8.2.2.1. Criterio de mejor R^2 Ajustada

El método de selección de modelos lineales según el criterio de mejor subconjunto de R^2 ajustada se basa en el ajuste del conjunto de entrenamiento, y se ajusta para penalizar modelos excesivamente complejos.

Los resultados ofrecidos por este método se pueden observar en las Tablas 29-32 y en los gráficos de importancia de variables (figuras 41-43). Al igual que sucedía con el método anterior, las variables seleccionadas para las biomásas de Montero y Canga han resultado ser las mismas, aunque en distinto orden de importancia. En este caso, de

nuevo la variable que con mayor fuerza explica la varianza de las variables dependientes en ambos modelos es una variable relacionada con la densidad de cubierta del dosel. ECMC. A continuación, aparecen variables relacionadas con la altura (EL2, EL3, EL1 Y EP40) seguidas por PlAM, EMAX, EP70, EP60 y AAMT1, con una importancia relativa mucho menor. Llama la atención el elevado número de variables implicadas en el modelo, teniendo en cuenta que los resultados del ajuste son peores que los del método *Stepwise*, ofreciendo este un menor número de variables de predicción (R^2 ajustada de 0,696 y 0,733).

En el caso de la biomasa aérea estimada a partir de las ecuaciones de Pérez-Cruzado, el número de variables, siendo menor, ofrece unos mejores resultados de ajuste y exactitud (R^2 ajustada de 0,764). En cuanto al orden de variables, se repite el esquema anterior, con una variable de densidad que destaca por encima de las demás: PAAM, seguida por variables de altura (ESTD, EL2, EP50, EP40) y una de densidad (CRR).

En todos los modelos los estadísticos f y t son significativos, pero nos encontramos de nuevo con un problema de colinealidad de las variables predictoras.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Tabla 29 Método mejores subconjuntos. Criterio R^2

Mejor Modelo	R^2	R^2 aj.	R^2 test	SE	ANOVA	
					f	p
MONTERO						
1	0,715	0,696	0,644	46,963	39,209	0,000
CANGA						
1	0,749	0,733	0,680	43,426	46,744	0,000
PÉREZ-CRUZADO						
1	0,774	0,764	0,747	37,684	81,977	0,000

Importancia del predictor

Destino: BIOA_MONT

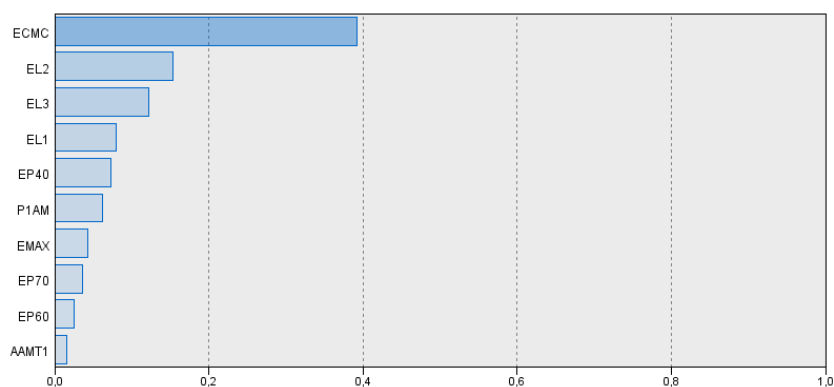


Figura 40 Importancia de los predictores para Montero (Mejor subconjunto R^2)

Tabla 30 Estadísticos de los coeficientes para Montero (Mejor subconjunto R^2)

Modelo	Coeficientes ^a											Estadísticas de colinealidad		
	Coeficientes no estandarizados		Coeficiente estandarizados	t	Sig.	95,0% intervalo de confianza para B		Correlaciones			Tolerancia	VIF		
	B	Error estándar				Limite inferior	Limite superior	Orden cero	Parcial	Parte				
1	(Constante)	10,703	19,230		0,557	0,579	-27,313	48,719						
	EMAX	-1,557	2,034	-0,139	-0,766	0,445	-5,579	2,464	0,652	-0,064	-0,034	0,061	16,330	
	EL1	131,741	55,098	8,634	2,391	0,018	22,817	240,666	0,729	0,197	0,108	0,000	6439,127	
	EL2	76,386	23,227	0,719	3,289	0,001	30,468	122,304	0,637	0,267	0,148	0,042	23,581	
	EL3	-227,908	111,510	-0,758	-2,044	0,043	-448,356	-7,461	-0,677	-0,170	-0,092	0,015	67,987	
	EP40	-82,743	36,189	-5,580	-2,286	0,024	-154,287	-11,199	0,733	-0,189	-0,103	0,000	2941,209	
	EP60	97,942	38,224	7,077	2,562	0,011	22,375	173,508	0,744	0,211	0,115	0,000	3767,680	
	EP70	-142,437	34,431	-10,629	-4,137	0,000	-210,505	-74,369	0,741	-0,329	-0,186	0,000	3260,139	
	P1AM	2,837	3,056	0,381	0,928	0,355	-3,206	8,879	0,663	0,078	0,042	0,012	83,397	
	AAMT1	-0,030	2,766	-0,004	-0,011	0,991	-5,497	5,438	0,649	-0,001	0,000	0,012	83,398	

a. Variable dependiente: BIOA_MONT

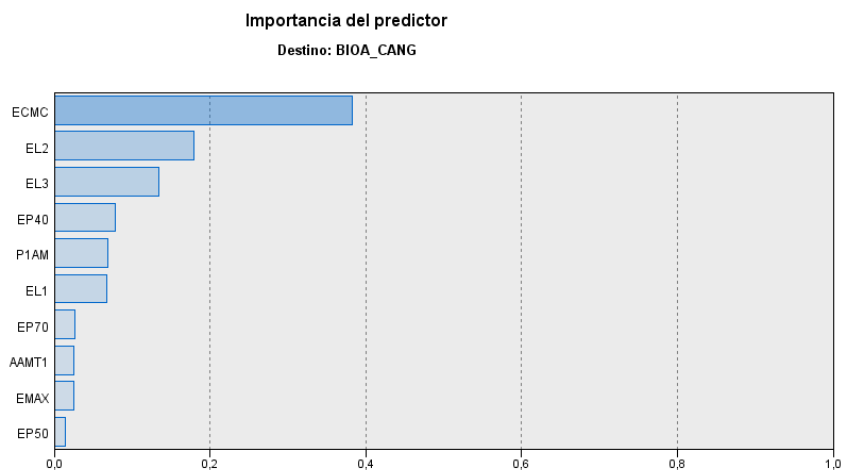


Figura 41 Importancia de los predictores para Canga (Mejor subconjunto R²)

Tabla 31 Estadísticos de los coeficientes para Canga (Mejor subconjunto R²)

Coefficientes^a

Modelo	Coeficientes no estandarizados		Coeficiente s estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad		
	B	Error estándar	Beta	t		Limite inferior	Limite superior	Orden cero	Parcial	Parte	Tolerancia	VIF	
1	(Constante)	-7,470	17,731		-0,421	0,674	-42,523	27,582					
	EMAX	-0,794	1,875	-0,072	-0,424	0,673	-4,501	2,913	0,697	-0,036	-0,018	0,062	16,228
	EL1	143,070	50,632	9,508	2,826	0,005	42,974	243,165	0,782	0,232	0,119	0,000	6359,416
	EL2	60,346	21,237	0,576	2,842	0,005	18,362	102,330	0,655	0,233	0,120	0,043	23,055
	EL3	-263,462	100,344	-0,889	-2,626	0,010	-461,834	-65,090	-0,700	-0,216	-0,111	0,016	64,386
	EP40	-104,038	37,730	-7,114	-2,757	0,007	-178,627	-29,449	0,787	-0,226	-0,116	0,000	3738,923
	EP50	56,249	36,629	3,990	1,536	0,127	-16,165	128,662	0,792	0,128	0,065	0,000	3792,373
	EP70	-90,644	24,067	-6,859	-3,766	0,000	-138,222	-43,066	0,790	-0,302	-0,159	0,001	1862,847
	P1AM	3,541	2,832	0,483	1,250	0,213	-2,058	9,139	0,651	0,105	0,053	0,012	83,740
	AAMT1	-1,008	2,566	-0,152	-0,393	0,695	-6,080	4,064	0,633	-0,033	-0,017	0,012	83,943

a. Variable dependiente: BIOA_CANG

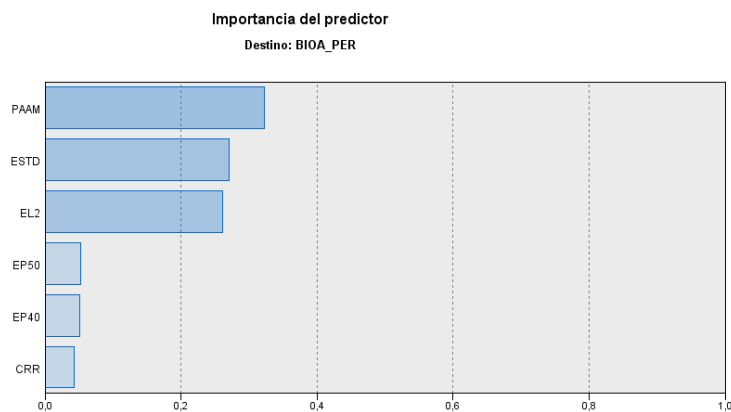


Figura 42 Importancia de los predictores para Pérez-Cruzado (Mejor subconjunto R²)

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Tabla 32 Estadísticos de los coeficientes para Pérez-Cruzado (Mejor subconjunto R2)

Modelo	Coeficientes ^a												
	Coeficientes no estandarizados		Coeficiente estandarizados	t	Sig.	95,04 intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad		
	B	Error estándar				Beta	Limite inferior	Limite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1	(Constante)	-59,503	18,769										
	ESTD	243,038	36,312	4,590	6,693	0,000	171,265	314,812	0,648	0,487	0,265	0,003	298,998
	EL2	-447,960	68,248	-4,627	-6,562	0,000	-582,757	-312,963	0,619	-0,480	-0,260	0,003	316,190
	EP40	-63,561	21,836	-4,707	-2,911	0,004	-106,720	-20,401	0,789	-0,236	-0,115	0,001	1662,974
	EP50	64,476	21,927	4,954	2,940	0,004	21,134	107,817	0,793	0,238	0,117	0,001	1804,722
	CER	193,753	73,165	0,286	2,648	0,009	49,137	338,370	0,738	0,215	0,105	0,135	7,396
	PAAM	3,033	0,416	0,366	7,293	0,000	2,211	3,855	0,658	0,519	0,289	0,625	1,601

a. Variable dependiente: BIOA_PER

8.2.2.2. Criterio de prevención del sobreajuste

El criterio de prevención sobreajustado se basa en el error cuadrado medio del conjunto de prevención sobreajustado. El conjunto de prevención sobreajustado es una submuestra aleatoria de aproximadamente el 30% del conjunto de datos original que no se utiliza para enseñar el modelo.

En este método observamos que los datos de ajuste, tanto del R^2 ajustado como del grupo test son algo mejores que los del método de R^2 ajustada, pero siguen produciéndose problemas de colinealidad (Tablas 33-36). Las variables seleccionadas se han reducido para todas las fórmulas de biomasa, pero siguen siendo las variables de densidad las que proporcionan mayores correlaciones parciales. En todos los casos es la variable PlAM la más fuerte, seguida de otra variable relacionada con la densidad (ECMC), y por primera vez aparece una variable que tiene relación con la estructura vertical del dosel forestal (ESKE) (figuras 44-46). Todos los modelos superan los test f y t .

Tabla 33 Método mejores subconjuntos. Criterio de prevención del sobreajuste

Mejor Modelo	R ²	R ² aj.	R2 Test	SE	ANOVA	
					f	p
MONTERO						
1	0,701	0,688	0,660	47,566	56,240	0,000
CANGA						
1	0,754	0,740	0,710	42,813	54,487	0,000
PEREZ CRUZADO						
1	0,743	0,728	0,702	40,450	51,237	0,000

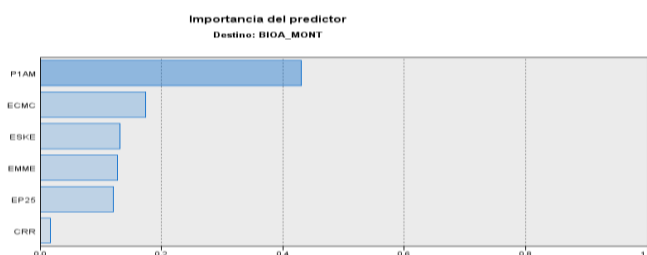


Figura 43 Estadísticos de los coeficientes para Montero (Mejor subconjunto Prevención del sobreajuste)

Tabla 34 Importancia de los predictores para Montero (Mejor subconjunto Prevención del sobreajuste)

Modelo	Coeficientes no estandarizados		estandarizados		t	Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad		
	B	Error estándar	Beta	t			Limite inferior	Limite superior	Orden cero	Parcial	Parte	Tolerancia	VIF	
1	(Constante)	86,300	49,601		1,740	0,084	-11,740	184,340						
	ESKE	-74,401	19,990	-0,540	-3,722	0,000	-113,512	-34,890	-0,706	-0,296	-0,170	0,059	10,117	
	EMME	-59,117	21,477	-0,633	-2,753	0,007	-101,568	-16,667	0,514	-0,224	-0,125	0,039	25,446	
	EP25	-36,309	13,241	-2,266	-2,742	0,007	-62,481	-10,137	0,714	-0,223	-0,125	0,003	328,884	
	CFR	-185,552	126,047	-0,249	-1,472	0,143	-434,694	63,591	0,681	-0,122	-0,067	0,073	13,779	
	EGMG	43,557	13,186	2,952	3,303	0,001	17,495	69,620	0,737	0,265	0,151	0,003	384,502	
	PIAM	2,891	0,419	0,389	6,893	0,000	2,062	3,720	0,663	0,498	0,314	0,653	1,531	

a. Variable dependiente: BIOA_MONT

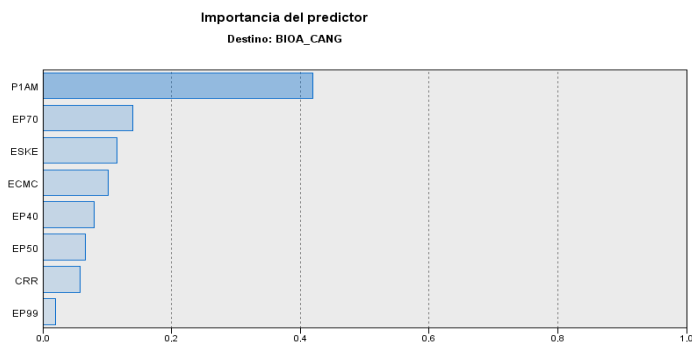


Figura 44 Importancia de los predictores para Canga (Mejor subconjunto Prevención del sobreajuste)

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Tabla 35 Estadísticos de los coeficientes para Canga (Mejor subconjunto Prevención del sobreajuste)

Modelo	Coeficientes no estandarizados		Coeficiente s estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar	Beta	t		Límite inferior	Límite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1 (Constante)	88,703	45,296		1,958	0,052	-0,838	178,244					
ESKE	-63,123	15,182	-0,464	-3,291	0,001	-101,041	-25,204	-0,734	-0,266	-0,137	0,087	11,469
EP40	-70,859	24,793	-4,846	-2,858	0,005	-119,880	-21,858	0,787	-0,233	-0,119	0,001	1661,024
EP50	53,649	35,427	6,643	2,643	0,009	23,616	163,681	0,752	0,217	0,110	0,000	3649,880
EP70	-62,096	16,505	-4,699	-3,762	0,000	-84,722	-39,469	0,790	-0,301	-0,157	0,001	901,376
EP99	-5,588	3,765	-0,488	-1,484	0,140	-13,030	1,855	0,739	-0,124	-0,062	0,016	62,562
CRP	-273,223	115,976	-0,372	-2,358	0,020	-502,287	-44,158	0,724	-0,194	-0,098	0,070	14,373
ECMC	56,930	17,937	3,513	3,174	0,002	21,471	92,389	0,788	0,257	0,132	0,001	878,312
FLAM	2,524	0,383	0,344	6,591	0,000	1,767	3,281	0,691	0,484	0,274	0,635	1,575

a. Variable dependiente: BIOA_CAMG

Importancia del predictor

Destino: BIOA_PER

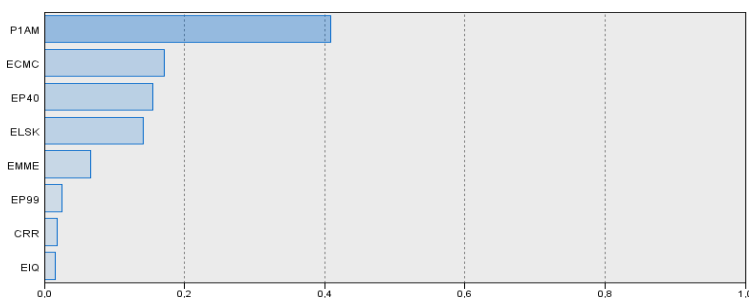


Figura 45 Importancia de los predictores para Pérez-Cruzado (Mejor subconjunto Prevención del sobreajuste)

Tabla 36 Estadísticos de los coeficientes para Pérez-Cruzado (Mejor subconjunto Prevención del sobreajuste)

Modelo	Coeficientes no estandarizados		Coeficiente s estandarizados		Sig.	95,0% intervalo de confianza para B		Correlaciones			Estadísticas de colinealidad	
	B	Error estándar	Beta	t		Límite inferior	Límite superior	Orden cero	Parcial	Parte	Tolerancia	VIF
1 (Constante)	84,978	45,392		1,872	0,063	-4,755	174,710					
EIQ	16,875	19,011	0,407	0,888	0,376	-20,706	54,457	0,481	0,074	0,038	0,009	116,243
EMME	-90,844	40,675	-1,068	-2,233	0,027	-171,251	-10,438	0,478	-0,184	-0,095	0,008	126,214
ELSK	-418,050	113,356	-0,585	-3,688	0,000	-642,133	-193,967	-0,739	-0,296	-0,157	0,072	13,887
EP40	-57,390	14,950	-4,250	-3,838	0,000	-86,934	-27,826	0,789	-0,307	-0,163	0,001	676,619
EP99	-8,042	3,692	-0,761	-2,178	0,031	-15,340	-0,744	0,733	-0,180	-0,093	0,015	67,402
CRP	-181,621	111,207	-0,268	-1,633	0,105	-401,456	38,214	0,738	-0,136	-0,070	0,067	14,831
ECMC	77,309	18,949	5,755	4,080	0,000	39,851	114,766	0,786	0,324	0,174	0,001	1098,008
FLAM	2,275	0,368	0,336	6,180	0,000	1,547	3,003	0,625	0,460	0,263	0,613	1,631

a. Variable dependiente: BIOA_PER

8.2.2.3. Métodos LASSO y Ridge

Como se ha comentado en el capítulo anterior ambas técnicas tiene como propósito principal mejorar el ajuste del modelo lidiando con el problema de la colinealidad, empleando para ello una constante $k > 0$ a cada término de la diagonal de $X^{tr}X$ antes de invertir la matriz. Esto supone encontrar un balance entre los componentes de sesgo y variancia del error cuadrático medio.

El principal problema que plantea esta técnica es que si bien al aumentar k (mayor penalización) los coeficientes estimados se contraen hacia cero, ninguno de ellos vale exactamente cero por lo cual no se produce selección de variables.

El programa Salford Predictive Modeler® permite estimar el parámetro k mediante validación cruzada y los mejores modelos de entre todos los posibles empleando criterios basados en la exactitud:

Además, permite mantener bajo control tanto el número máximo de variables predictoras (seleccionadas entre las que obtienen una mayor puntuación), como la colinealidad entre las mismas al permitir fijar un máximo de correlación entre las variables candidatas.

Se ha optado por comparar los resultados de ajuste con y sin restricción del número de variables predictoras, fijando en 5 el número de variables del segundo test, con un máximo de correlación entre ellas de +/- 0,6. Los primeros se codifican con el sufijo 1 y los segundos con

el sufijo 2. En ambos casos se ha establecido una validación cruzada OOL y como criterios de selección de mejor modelo el que ofrece un mejor R^2 de GVC (Tabla 37).

Una de los problemas que acompaña a los modelos Ridge 1 es que al incluir la totalidad de los predictores la colinealidad se hace patente en todos los modelos (figuras 47-49). Al aplicar el criterio restrictivo del máximo de variables admisibles y del máximo de correlación de 0,6 el problema se reduce, pero los resultados de ajuste empeoran. (Tabla 37).

Algo parecido pasa con los modelos Lasso. Los modelos sin restricción de variables siguen incluyendo un número de variables con correlaciones muy fuertes entre ellas. Al aplicar los criterios restrictivos mencionados la pérdida de ajuste y la mejora en el error es, sin embargo, mejor que en el método ridge2. De todos los modelos posibles el que mejor ajuste ofrece sin comprometer la colinealidad es el de Canga-Lasso2 con un coeficiente de determinación de 0,717, un RMSE de 41,78 y un MAPE del 23%.

Tabla 37 Resultados de ajuste de los métodos Ridge y Lasso

Ecuación alométrica	Modelo	R^2	R^2 ajustado	R^2 Test	RMSE	MAPE
MONTERO	Ridge 1	0,724	0,724	0,656	44,615	0,209
	Ridge 2	0,670	0,673	0,635	48,725	0,221
	Lasso 1	0,707	0,707	0,645	45,953	0,216
	Lasso 2	0,667	0,668	0,625	48,953	0,227
CANGA	Ridge 1	0,730	0,712	0,675	43,547	0,222
	Ridge 2	0,711	0,712	0,678	45,022	0,236
	Lasso 1	0,789	0,789	0,701	38,440	0,198
	Lasso 2	0,716	0,717	0,674	44,628	0,228
PÉREZ-CRUZADO	Ridge 1	0,708	0,708	0,682	41,797	0,225
	Ridge 2	0,704	0,704	0,659	42,074	0,229
	Lasso 1	0,716	0,716	0,670	41,212	0,224
	Lasso 2	0,706	0,706	0,664	41,888	0,231

Importancia del predictor (Montero)
Ridge 1

Variable	Score	
PIAM	100,00	
AMT1	97,58	
PAAM	94,03	
EL3	59,61	
ESTD	52,73	
ESKE	52,71	
ELSK	43,59	
EL2	37,91	
EAAD	36,73	
CRR	26,85	
EMOD	25,91	
EP40	25,69	
EP20	25,28	
EP30	25,14	
EP60	24,61	
ESQM	24,57	
EP60	24,36	
EP25	24,21	
ECMC	23,96	
EP75	23,79	
EMEA	23,64	
EL1	23,58	
EP70	23,52	
EP90	23,12	
EP80	22,59	
EP95	21,58	
EMME	21,27	
EP99	19,93	
EIQ	9,97	
EMAX	8,71	

Importancia del predictor (Montero)
Ridge 2

Variable	Score	
EP60	100,00	
PIAM	94,20	
EL3	59,62	

Importancia del predictor (Montero)
Lasso 1

Variable	Score	
ESTD	100,00	
EMME	54,78	
ESKE	36,35	
PAAM	36,02	
EP60	31,94	
EMAX	18,03	
AMT1	17,72	
PIAM	12,42	
ECMC	7,18	
EL2	4,59	
CRR	3,93	
EP20	1,68	
ELSK	1,60	
EL3	0,03	

Importancia del predictor (Montero)
Lasso 2

Variable	Score	
EP60	100,00	
PAAM	85,91	
ELSK	31,54	

Figura 46 Selección de variables para Montero (Métodos Ridge y LASSO)

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

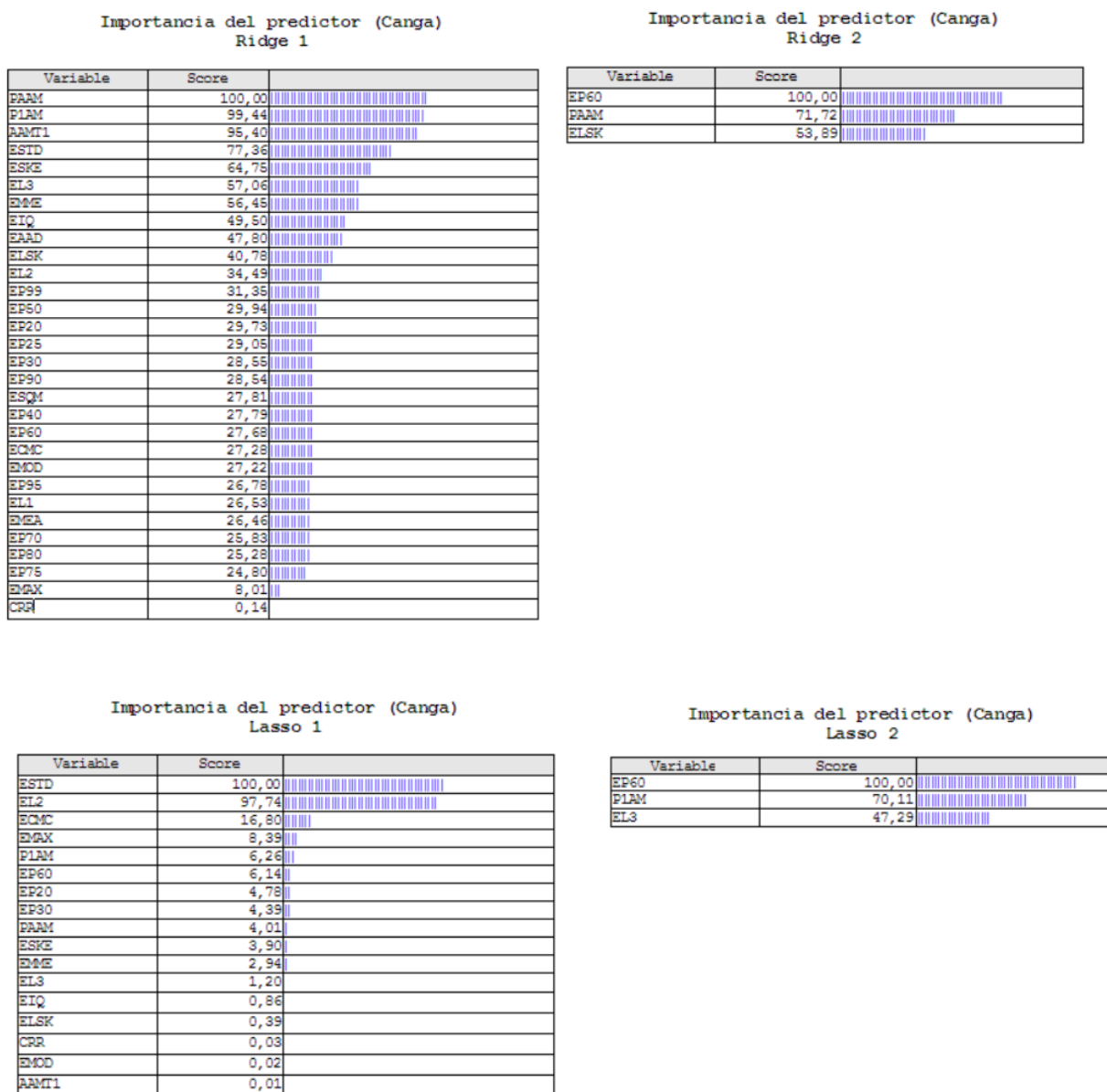


Figura 47 Selección de variables para Canga (Métodos Ridge y LASSO)

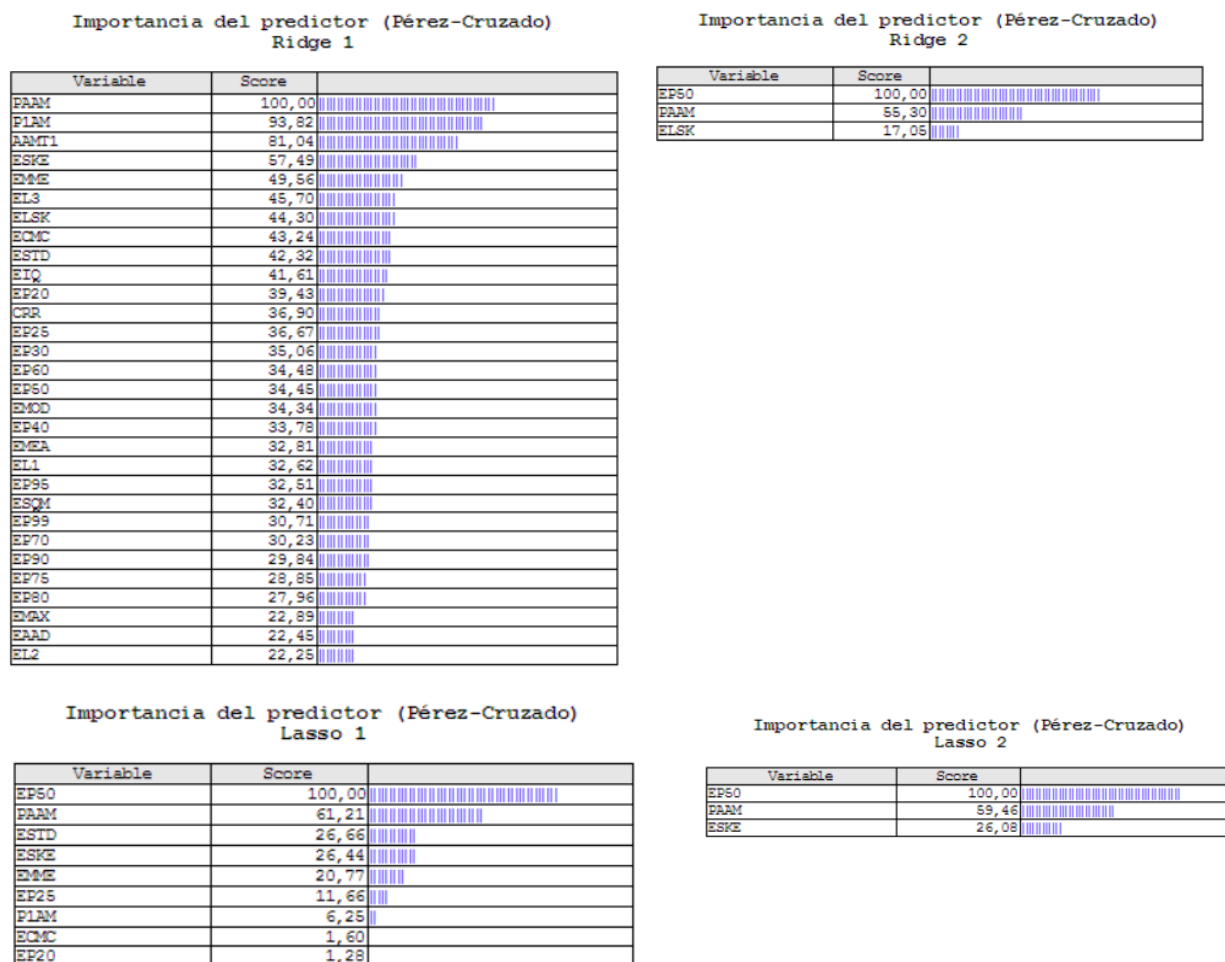


Figura 48 Selección de variables para Pérez-Cruzado (Métodos Ridge y LASSO)

8.2.2.4. Método MARS

Como se ha comentado arriba, los modelos MARS combinan la regresión lineal clásica, la construcción matemática de *splines* y la división recursiva binaria para producir modelos donde las relaciones entre la variable dependiente y las predictoras pueden ser lineales o no lineales. Para ello, se aproxima a la función subyacente a través de un conjunto de *regresiones lineales adaptativas* por

subconjuntos denominadas *funciones de base* (BF) (Friedman, 1991; Solomon, Oatley y McGarry, 2007).

$$f(X) = \beta_0 + \sum_{m=1}^M \beta_m h_m(X),$$

Donde:

$h_m(X)$ es una función del conjunto C de funciones candidatas o un producto de dos o más de dichas funciones.

β son los coeficientes estimados minimizando la suma residual de cuadrados (regresión lineal estándar). Estos pueden considerarse como los pesos que representan la importancia de la variable.

Por lo tanto en los modelos MARS los coeficientes β de la regresión lineal tradicional no ponderan directamente las variables sino una función de las mismas del tipo $\max(0, x-n)$ (Friedman, 1991).

Se muestran a continuación los resultados de ajuste y exactitud del mejor modelo seleccionado por MARS del conjunto de las 30 variables candidatas a partir de dos criterios diferentes: uno en el que se pondera el R^2 de GVC (Generalized Cross Validation), equivalente a la R^2 de una validación cruzada LOO y de un criterio, y un segundo en el que se prima el mínimo RMSE del conjunto de test.

En la tabla 38 podemos observar que los resultados de ajuste son mejores que los métodos anteriores en todas las combinaciones, por lo que a priori se podría considerar

que son la mejor alternativa. Sin embargo, hay que tener en cuenta que uno de los problemas que presenta este tipo de modelos está en la dificultad de comprobar la colinealidad de las funciones de base. Para comprobar esa cuestión, se ha considerado cada función de base como una variable independiente de un modelo de regresión clásico. Los resultados de la significación de los estadísticos t de los coeficientes β y la infracción de la varianza de cada función de base variables se muestran en la tabla 40.

Tabla 38 ajuste y error de los modelos MARS

Criterios	R ²	R ² ajustado	R ² Test	RMSE	RMSE Test	MAPE
MONTERO						
GVC	0,738	0,731	0,716	44,212	53,552	0,215
RMSE TEST	0,708	0,704	0,695	42,557	48,068	0,203
CANGA						
GVC	0,798	0,788	0,770	36,476	43,970	0,195
RMSE TEST	0,783	0,777	0,769	39,634	42,221	0,197
PÉREZ-CRUZADO						
GVC	0,797	0,788	0,750	34,818	43,244	0,190
RMSE TEST	0,764	0,758	0,758	37,556	40,564	0,211

Las variables incluidas en las funciones de base incluyen variables que muestran altas correlaciones: EP50, EL3 y EP40 en las funciones de base GVC_CANGA_BF6, GVC_CANG_BF7, GVC_CANG_BF10, GVC_PER_BF1, GVC_PER_BF6, GVC_PER_BF12, por lo que, como cabía esperar, los valores VIF de estas son demasiado altos. Por otro lado, las funciones son todas significativas al 95% excepto las GVC_CANGA_BF6, GVC_CANG_BF7, GVC_CANG_BF10 (Tablas 39 y 40).

De nuevo, de la aplicación de este método para seleccionar variables, es una de las alternativas de la ecuación de (Canga et al., 2013) la que proporciona mejores resultados: la RMSE_TEST, que alcanza a explicar el casi el 78% de la

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

varianza con un RMSE de 39,63 y un MAPE del 19,7% (Tabla 40).

Tabla 39 Variables y Funciones de Base de los Modelos MARS

Criterios	Variables	Funciones / Modelo														
MONTERO																
GVC	<p>Importancia del predictor (Montero) MARS (GVC)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>ELSK</td> <td>100,00</td> </tr> <tr> <td>PAAM</td> <td>69,38</td> </tr> <tr> <td>EMME</td> <td>85,45</td> </tr> </tbody> </table>	Variable	Score	ELSK	100,00	PAAM	69,38	EMME	85,45	$BF1 = \max(0, ELSK - 0,022998)$ $BF2 = \max(0, 0,022998 - ELSK)$ $BF3 = \max(0, PAAM - 2,08944)$ $BF4 = \max(0, EMME - 2,7) * BF1$ $Y = 45,5635 + 696,901 * BF2 + 3,73736 * BF3 + 1091,41 * BF4$						
Variable	Score															
ELSK	100,00															
PAAM	69,38															
EMME	85,45															
RMSE Test	<p>Importancia del predictor (Montero) MARS (TEST RMSE)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>ELSK</td> <td>100,00</td> </tr> <tr> <td>PAAM</td> <td>74,12</td> </tr> </tbody> </table>	Variable	Score	ELSK	100,00	PAAM	74,12	$BF2 = \max(0, 0,022998 - ELSK)$ $BF3 = \max(0, PAAM - 2,08944)$ $Y = 50,9329 + 643,087 * BF2 + 3,81927 * BF3$								
Variable	Score															
ELSK	100,00															
PAAM	74,12															
CANGA																
GVC	<p>Importancia del predictor (Canga) MARS (GVC)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>PAAM</td> <td>100,00</td> </tr> <tr> <td>ESKE</td> <td>83,56</td> </tr> <tr> <td>EP50</td> <td>24,58</td> </tr> <tr> <td>EP40</td> <td>20,88</td> </tr> <tr> <td>EL3</td> <td>19,07</td> </tr> </tbody> </table>	Variable	Score	PAAM	100,00	ESKE	83,56	EP50	24,58	EP40	20,88	EL3	19,07	$BF2 = \max(0, PAAM - 33,1299)$ $BF3 = \max(0, 33,1299 - PAAM)$ $BF5 = \max(0, -0,267409 - ESKE)$ $BF6 = \max(0, EP50 - 15,92)$ $BF7 = \max(0, 15,92 - EP50)$ $BF8 = \max(0, EL3 - 0,173319)$ $BF10 = \max(0, EP40 - 4,488)$ $Y = 527,445 + 5,1926 * BF2 - 2,98638 * BF3 + 131,63 * BF5 + 34,755 * BF6 - 37,5443 * BF7 + 113,299 * BF8 - 33,7937 * BF10$		
Variable	Score															
PAAM	100,00															
ESKE	83,56															
EP50	24,58															
EP40	20,88															
EL3	19,07															
RMSE Test	<p>Importancia del predictor (Canga) MARS (TEST RMSE)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>PAAM</td> <td>100,00</td> </tr> <tr> <td>ESKE</td> <td>70,28</td> </tr> <tr> <td>EP60</td> <td>53,52</td> </tr> </tbody> </table>	Variable	Score	PAAM	100,00	ESKE	70,28	EP60	53,52	$BF1 = \max(0, EP60 - 4,754)$ $BF2 = \max(0, PAAM - 33,1299)$ $BF3 = \max(0, 33,1299 - PAAM)$ $BF5 = \max(0, -0,267409 - ESKE)$ $Y = 125,411 + 4,32866 * BF1 + 5,42746 * BF2 - 3,0291 * BF3 + 97,878 * BF5$						
Variable	Score															
PAAM	100,00															
ESKE	70,28															
EP60	53,52															
PÉREZ-CRUZADO																
GVC	<p>Importancia del predictor (Pérez-Cruzado) MARS (GVC)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>PIAM</td> <td>100,00</td> </tr> <tr> <td>ESKE</td> <td>89,91</td> </tr> <tr> <td>EP50</td> <td>87,57</td> </tr> <tr> <td>EL3</td> <td>85,07</td> </tr> <tr> <td>EP40</td> <td>82,34</td> </tr> <tr> <td>EPX</td> <td>89,42</td> </tr> </tbody> </table>	Variable	Score	PIAM	100,00	ESKE	89,91	EP50	87,57	EL3	85,07	EP40	82,34	EPX	89,42	$BF1 = \max(0, EP50 - 4,595)$ $BF5 = \max(0, -0,267409 - ESKE)$ $BF6 = \max(0, EP40 - 13,582)$ $BF12 = \max(0, 26,97 - EMAX)$ $BF19 = \max(0, PIAM - 2,18474)$ $BF47 = \max(0, EL3 - 0,050378)$ $Y = -133,882 + 21,1437 * BF1 + 115,717 * BF5 - 18,7089 * BF6 + 8,50399 * BF12 + 2,43443 * BF19 + 187,81 * BF47$
Variable	Score															
PIAM	100,00															
ESKE	89,91															
EP50	87,57															
EL3	85,07															
EP40	82,34															
EPX	89,42															
RMSE Test	<p>Importancia del predictor (Pérez-Cruzado) MARS (Test RMSE)</p> <table border="1"> <thead> <tr> <th>Variable</th> <th>Score</th> </tr> </thead> <tbody> <tr> <td>PIAM</td> <td>100,00</td> </tr> <tr> <td>ESKE</td> <td>72,89</td> </tr> <tr> <td>EP50</td> <td>85,26</td> </tr> </tbody> </table>	Variable	Score	PIAM	100,00	ESKE	72,89	EP50	85,26	$BF1 = \max(0, EP50 - 4,595)$ $BF2 = \max(0, PAAM - 39,9088)$ $BF3 = \max(0, 39,9088 - PAAM)$ $BF5 = \max(0, -0,267409 - ESKE)$ $Y = 142,311 + 4,24634 * BF1 + 6,69484 * BF2 - 2,93319 * BF3 + 89,1224 * BF5$						
Variable	Score															
PIAM	100,00															
ESKE	72,89															
EP50	85,26															

Tabla 40 Coeficientes y VIF de las funciones de base de los modelos MARS

Funciones MARS	β	SE	t	p	VIF
GVC MONTERO					
Constante	43,4	13,6	3,19	0,002	
GVC_MONT_BF1	-24,5	87,4	-0,28	0,780	1,79
GVC_MONT_BF2	685,9	57,5	11,93	0,000	1,37
GVC_MONT_BF3	3,859	0,492	7,84	0,000	1,63
GVC_MONT_BF4	460	120	3,84	0,000	1,30
RMSE Test MONTERO					
Constante	50,9	11,4	4,48	0,000	
RMSE_MONT_FB2	643,1	58,7	10,96	0,000	1,30
RMSE_MONT_FB3	3,819	0,461	8,28	0,000	1,30
GVC CANGA					
Constante	201,1	73,7	2,73	0,007	
GVC_CANG_BF2	5,62	1,18	4,77	0,000	1,38
GVC_CANG_BF3	-3,130	0,582	-5,38	0,000	1,90
GVC_CANG_BF5	115,9	18,7	6,19	0,000	3,25
GVC_CANG_BF6	4,85	7,96	0,61	0,544	85,88
GVC_CANG_BF7	-8,37	7,36	-1,14	0,257	46,36
GVC_CANG_BF8	132,9	44,3	3,00	0,003	1,14
GVC_CANG_BF10	-2,19	7,61	-0,29	0,774	191,68
RMSE Test CANGA					
Constante	125,4	12,2	10,32	0,000	
RMSE_CANG_BF1	4,329	0,896	4,83	0,000	2,91
RMSE_CANG_BF2	5,43	1,20	4,53	0,000	1,36
RMSE_CANG_BF3	-3,029	0,560	-5,41	0,000	1,67
RMSE_CANG_BF5	97,9	16,1	6,07	0,000	2,29
GVC PÉREZ-CRUZADO					
Constante	-133,9	39,0	-3,43	0,001	
GVC_PER_BF1	21,14	3,77	5,61	0,000	59,47
GVC_PER_BF5	115,7	17,2	6,73	0,000	3,22
GVC_PER_BF6	-18,71	4,09	-4,57	0,000	32,43
GVC_PER_BF12	8,50	2,26	3,77	0,000	11,38
GVC_PER_BF19	2,434	0,309	7,88	0,000	1,48
GVC_PER_BF47	187,8	39,5	4,75	0,000	1,97
RMSE Test PÉREZ-CRUZADO					
Constante	142,3	12,5	11,38	0,000	
RMSE_PER_BF1	4,246	0,903	4,70	0,000	2,98
RMSE_PER_BF2	6,69	2,57	2,60	0,010	1,14
RMSE_PER_BF3	-2,933	0,431	-6,81	0,000	1,52
RMSE_PER_BF5	89,1	15,8	5,63	0,000	2,38

8.2.2.5. Resumen de los métodos de selección de variables

La tabla 41 muestra un resumen de los métodos empleados para la selección de variables y de los principales parámetros de los modelos resultantes. La última columna muestra los principales problemas que presentan: colinealidad, grado de ajuste, exactitud y sobreajuste. Ya se ha dicho que los problemas de colinealidad hacen inviables los modelos desde la perspectiva de la regresión clásica. Su influencia es más cuestionable desde la perspectiva de la minería de datos y el Machine Learning en el que lo principal es la capacidad predictiva del modelo y no su capacidad de explicar las relaciones existentes entre variables independientes y la variable dependiente (Clarke et al., 2009).

Como se ha mencionado anteriormente, en este estudio se ha optado por adoptar la perspectiva clásica de la regresión lineal. Por este motivo, los modelos que presentaban colinealidad han sido descartados y, en su caso, reformulados para incluir los predictores que más contribuyen al ajuste y menos correlaciones parciales presentan. Por este motivo se han descartado todos los modelos con VIF superior a 3.

El resto de modelos presentan algún tipo de problema relacionado con el ajuste y la exactitud (valores de R^2 por debajo de 0,7 y MAPE por encima del 20%. Los modelos sin multicolinealidad no presentan graves problemas de sobreajuste (no llegan al 5%).

De entre todos los modelos hemos destacado dos, que son los que menos problemas presentan en términos globales. Ambos son modelos que emplean la biomasa estimada por la ecuación alométrica de (Canga et al., 2013). El Lasso y el MARS RMSE_TEST. Aparentan ser, por lo tanto, lo suficientemente robustos para evaluarlos en una tercera fase donde se analizarán posibles interacciones entre las variables predictoras (Aiken, West y Reno, 1991; Jaccard y Turrisi, 2003; Jaccard, Wan y Turrisi, 1990).

Tabla 41 Resumen del rendimiento de los modelos

Nº	Modelo	ECUACION ALOMETRICA	R ² AE (**)	R ² Test	Sobreajuste	SE	MAPE	VIF	Tipo de desajuste (***)
1	Stepwise	MONTERO	0,751	0,733	1,80%	42,553	19,9%	16638,43	a
		CANGA	0,792	0,777	1,50%	38,312	19,8%	16638,43	a
		PÉREZ-CRUZADO	0,764	0,747	1,70%	37,684	20,3%	4091,88	a
4	MS-R ²	MONTERO	0,696	0,644	5,30%	46,963	21,0%	16682,86	a, b, d
		CANGA	0,733	0,680	5,30%	43,426	21,0%	16024,92	a, d
		PÉREZ-CRUZADO	0,764	0,747	1,70%	37,684	20,4%	681,98	a
5	MS-PS	MONTERO	0,688	0,660	2,90%	47,566	22,9%	127,37	a, b
		CANGA	0,740	0,710	3,10%	42,813	22,6%	897,56	a
		PÉREZ-CRUZADO	0,728	0,702	2,70%	40,45	22,7%	264,35	a
6	Ridge ₂	MONTERO	0,673	0,635	3,80%	48,725	21,6%	2,29	b
		CANGA	0,712	0,678	3,40%	45,022	23,2%	2,92	
		PÉREZ-CRUZADO	0,704	0,659	4,50%	42,074	22,9%	2,98	
9	Lasso ₂	MONTERO	0,668	0,625	4,30%	48,953	21,6%	2,92	b
		CANGA	0,717	0,674	4,30%	44,628	21,7%	2,92	
		PÉREZ-CRUZADO	0,706	0,664	4,20%	41,888	23,1%	2,95	
10	MARS GVC	MONTERO	0,731	0,716	1,50%	44,212	20,5%	1,52	
		CANGA	0,788	0,770	1,80%	36,476	20,2%	47,37	a
		PÉREZ-CRUZADO	0,788	0,750	3,80%	34,818	19,5%	18,32	a
12	MARS RMSE Test	MONTERO	0,704	0,695	0,90%	42,557	21,6%	1,30	
		CANGA	0,777	0,769	0,80%	39,634	19,7%	2,05	
		PÉREZ-CRUZADO	0,758	0,758	0,00%	37,556	21,3%	2,00	

a	Colinealidad
b	Ajuste
c	Exactitud
d	Sobreajuste

(*) Mejor Subconjunto

(**) R² Ajustada del conjunto de datos de entrenamiento

(***) El color señala el principal problema

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

8.3. Selección final de modelos

En esta fase se han evaluado las posibles interacciones entre las variables surgidas del modelo Lasso 2 y las funciones de base de MARS RMSE-Test que mejoren los resultados han proporcionado que los anteriores modelos sin interacciones. Los resultados pueden observarse en la tablas 42 y 43. El único modelo que ha mejorado incorporando algún tipo de interacción entre variables es el Lasso 2 concretamente la potencia cuadrada de la variable ESKE. En el caso del modelo MARS ninguna interacción entre Funciones de base a proporcionado mejores valores de ajuste y exactitud.

Tabla 42 Rendimiento de los modelos seleccionados

Nº	Modelo	Variables	R ² AE (**)	R ² Test	Sobreajuste	SE	MAPE	VIF
1	RLM (2 interacciones)	EP60 ELSK PAAM ELSK*ELSK	0,772	0,759	1,68%	40,17	21,2%	2,58
2	MARS	BF1=max(0, EP60-4,754) BF2=max(0, PAAM-33,1299) BF3=max(0,33,1299-PAAM) BF5=max(0,-0,267409-ESKE)	0,777	0,769	0,80%	39,63	19,7%	2,05

Tabla 43 Estadísticos de los coeficientes y colinealidad de los modelos seleccionados

Variables	β	SE	t	p	VIF
Modelo RLM (Regresión Lineal Múltiple) con 2 interacciones					
EP60	5,45	1,03	-1,44	,000	3,72
ELSK	-129,3	59,1	5,31	,030	3,82
PAAM	3,566	0,452	-2,19	,000	1,66
ELSK*ELSK	1263	200	7,89	,000	1,12
Modelo MARS					
BF1=max(0, EP60-, 754)	4,329	0,896	4,83	,000	2,91
BF2=max(0, PAAM-33,1299)	5,43	1,20	4,53	,000	1,36
BF3=max(0,33,1299-PAAM)	3,029	0,560	-5,41	,000	1,67
BF5=max(0,-0,267409-ESKE)	97,9	16,1	6,07	,000	2,29

El siguiente paso ha consistido en comprobar los supuestos estructurales de la regresión lineal de ambos modelos. Los resultados de la comprobación de dichos supuestos estructurales pueden observarse en la tabla 44. Ambos modelos superan todos los test, excepto el de la existencia de valores atípicos:

- El test de normalidad de Shapiro-Wilk, que plantea como hipótesis nula que la muestra proviene de una población normalmente distribuida (Shapiro and Wilk 1965).
- El test de Breusch-Pagan, que fija como hipótesis nula que los residuos son homocedásticos, es decir, que la varianza de los residuos es constante.
- El test de Durbin-Watson, para desechar posibles dependencias entre los datos que partiendo de la hipótesis nula de que la autocorrelación es cero.
- El ya explicado análisis de la colinealidad en el modelo, a partir del factor de inflación de la varianza (VIF) como medida de colinealidad, siendo 1 su valor óptimo, que implicaría la ausencia total de la misma (Kleinbaum, Kupper, Muller 1991).
- Para verificar es la linealidad del modelo se ha comprobado el test RESET de Ramsey, que indica si el valor esperado de la variable dependiente depende linealmente de las variables independientes, es decir, si las combinaciones no lineales de los valores ajustados ayudan a explicar la variable de respuesta.

La hipótesis nula parte de considerar que existe dependencia lineal.

- Por último, se ha verificado la ausencia de valores atípicos a partir del test de Bonferroni, que para un nivel de confianza del %95 por ciento, señala si existen valores atípicos estadísticamente significativos. La hipótesis nula de la inexistencia de valores atípicos en la muestra.

Tabla 44 Test de validación de los supuestos estructurales de los modelos seleccionados

Modelo	MAPE	R ² Ajustada	R ² Test	Shapiro-Wilk	Breusch-Pagan	Durbin-Watson	VIF	RESET	Bonferroni
RLM con 2 interacciones	21,2%	0,772	0,759	0,234	0.124	1,760	2,581	0,907	0,000
MARS (aciones)	19,7%	0,777	0,769	0,101	0,873	1,808	2,011	0,768	0,000

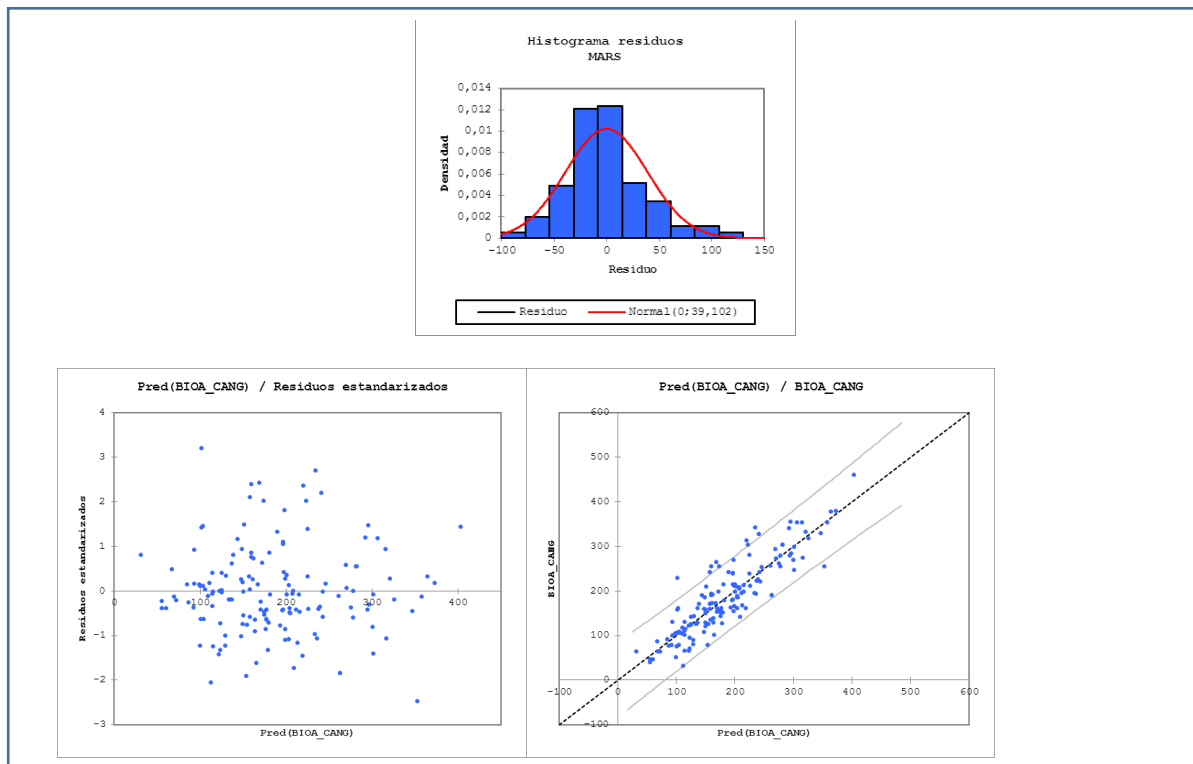


Figura 49 Graficas de normalidad, homocedasticidad y linealidad de los residuos (MARS)

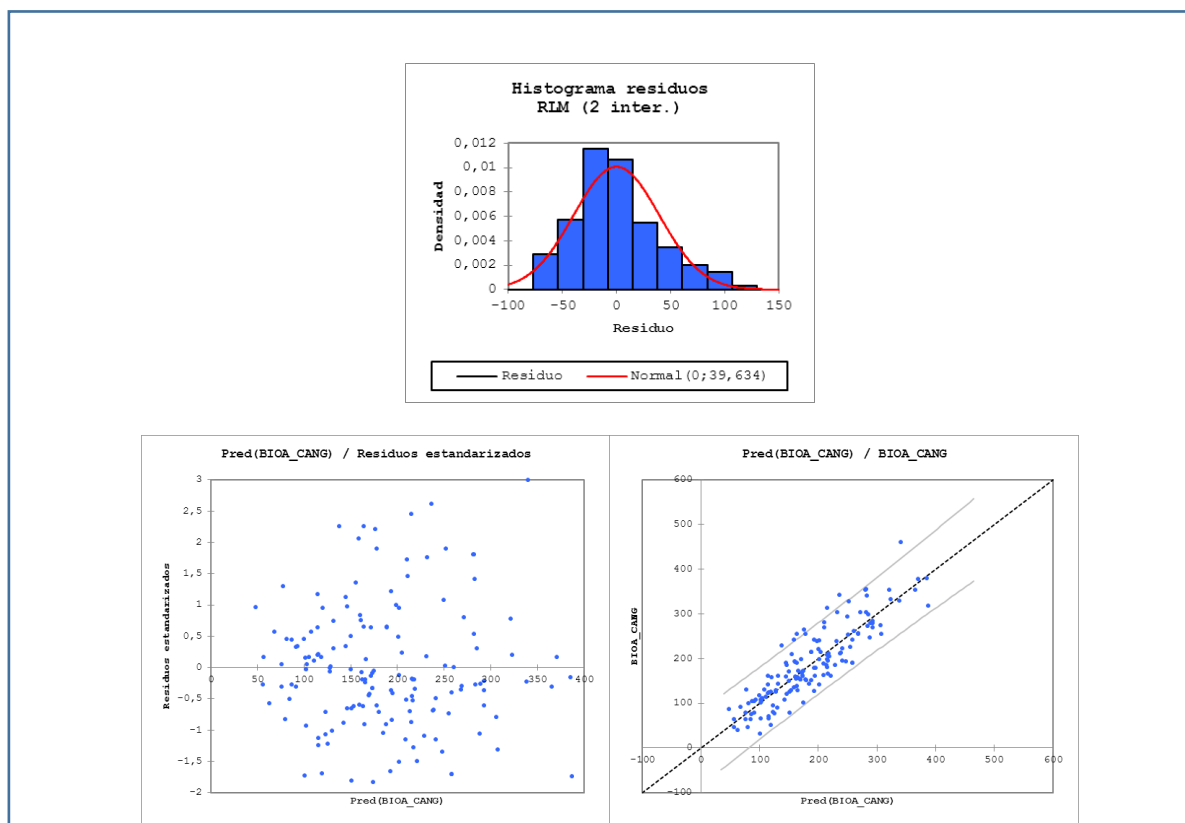


Figura 50 Graficas de normalidad, homocedasticidad y linealidad de los residuos (Regresión Lineal)

8.4. Tratamiento de los valores anómalos

El último paso necesario para seleccionar el modelo definitivo es la detección y, en caso de ser necesario, el tratamiento de los valores anómalos. Si bien no existe una definición exacta de un valor atípico o de una observación atípica, los valores atípicos son observaciones que no siguen el patrón de las otras observaciones.

Esto no es normalmente un problema si el valor atípico es simplemente una observación extrema extraída de la cola de una distribución normal, pero si los resultados atípicos son fruto de un error de medición o alguna otra violación de los supuestos ordinarios de la regresión por mínimos

cuadrados estándar, se puede llegar a comprometer la validez de los resultados.

La probable existencia de errores no sistemáticos en la toma de coordenadas GPS del centro de las parcelas del IFN4 hace que la posibilidad de encontrar valores atípicos entre los datos de las variables predictoras no sea descartable.

Evaluar y corregir dichos errores mediante una nueva toma de datos de las coordenadas del centro de las parcelas con trabajo de campo y usando técnicas geodésicas adecuadas sería lo deseable, pero como se ha comentado anteriormente, muy costoso tanto en tiempo como en recursos materiales y fuera del alcance de este estudio.

Se ha optado por abordar los posibles problemas siguiendo los siguientes pasos:

1º Detectar posibles valores anómalos en los datos prestando especial atención a los valores atípicos en las variables predictoras

2º Comprobar las parcelas clasificadas como atípicas consultando las ortofotos del 2008 y 2011 e intentar detectar si se trata de errores groseros tales como grandes diferencias entre los datos del inventario y la información registrada por el sensor, nº de rebotes ilógicos en comparación con la biomasa estimada con los datos de campo, etcétera...

2º Comprobar si se trata de errores claramente debidos a una incorrecta geolocalización de las parcelas. Para ello

se empleará un método de delineación de copas individuales a partir de los datos LiDAR para crear una capa de polígonos por cada parcela, que se superpondrá sobre la ubicación de los árboles geolocalizados en la fase de campo del IFN4.

En uno y otro caso, las parcelas que muestren claras discrepancias serán descartadas del análisis.

1.- Detección de valores anómalos:

Para abordar este análisis se ha recurrido al análisis de distintos gráficos en los que se relacionan los residuos, las distancias de Cook y los valores de apalancamiento (Fig. X). A partir de ellos se han detectado 11 parcelas con valores anómalos, de las cuales 8 están presentes en ambos modelos, (tabla 45).

Tabla 45 Parcelas con valores atípicos

Modelo RLM	Modelo MARS
0045	0053
0121	0121
0213	0213
0571	0571
0711	0711
0723	0723
0749	0749
0763	0763
0767	0767
0795	0993
1171	1093

2.- Una vez detectadas las anomalías a nivel de estadísticas del modelo se ha procedido a analizar las posibles causas mediante los métodos descritos arriba:

Para comprobar la hipótesis de errores de posicionamiento en las parcelas detectadas como anómalas se ha realizado una prueba de verificación consistente en contrastar la ubicación de los árboles en parcela del IFN4 y las copas de árboles delineados a partir de datos LIDAR.

Para la delineación de copas se ha seguido una aproximación adaptada de la de Hyypä et al. (2001). Este enfoque encuentra máximos locales dentro de un Modelo de Alturas de Dosel rasterizado (CHM - Canopy Height Model), los designa como ápices de árbol y luego usa un método de árbol de decisión para hacer crecimientos de coronas individuales alrededor de esos máximos locales.

La base teórico-cuantitativa de este enfoque para el análisis de la estructura geométrica dentro del modelo de superficie del dosel se encuentra en la denominada *morfología matemática*. Se trata de una secuencia de transformaciones de imagen morfológicas binarias y en escala de grises que se han empleado para aislar los árboles individuales que componen la superficie del dosel.

Aunque originalmente fue desarrollado para el análisis de imágenes binarias de dos dimensiones se ha extendido a las imágenes en escala de grises en tres dimensiones, donde los valores de escala de grises representan la intensidad u otro atributo de píxel, como la elevación de una superficie.

En el contexto de la morfología, los conjuntos representan las formas que conforman una imagen binaria o de escala de grises. Los conjuntos en dos dimensiones describen el

primer plano de la imagen y en tres dimensiones que pueden describir la variación dentro de una superficie.

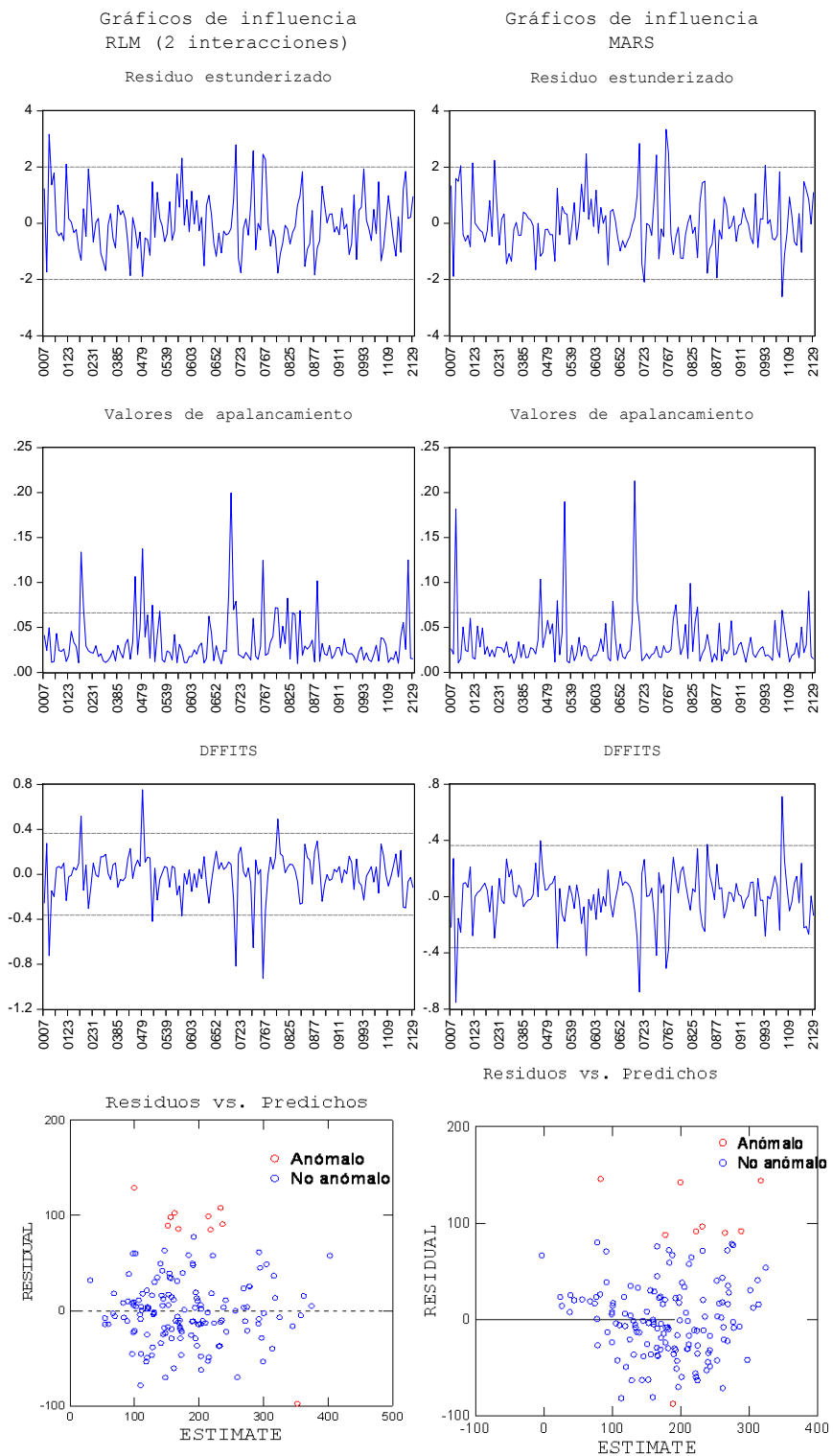


Figura 51 Gráficas de valores anómalos

El objetivo de cualquier operación morfológica es obtener información relativa a la estructura geométrica de una imagen mediante el sondeo de la imagen con otro conjunto, de tamaño y forma especificada, conocido como un *elemento estructurante*.

El tamaño y la forma del elemento estructurante se elige de acuerdo con el tipo de información de forma que se extrae de la imagen. En términos formales, una operación morfológica es una transformación de la imagen con el elemento de estructuración que sirve como parámetro para la transformación.

El resultado de una sola transformación (operación morfológica), llevado a cabo con un elemento estructurante dado, transmite información relativa al contenido de la forma de la imagen original. Variando el tamaño del elemento de estructuración puede dar lugar a diferentes transformaciones de imagen y por lo tanto puede proporcionar aún más información sobre contenido de la imagen.

Las operaciones morfológicas básicas son la **dilatación** (dilata) y la **erosión** (erode). Si una imagen se representa como un conjunto A y un elemento de estructuración como otro (más pequeño) conjunto B, el resultado de la dilatación de la imagen por un elemento estructurante B puede ser considerado como que muestra las áreas en que el elemento estructurante B encaja en el conjunto A (Soille, 1999). En términos teóricos, si A y B son subconjuntos de espacio d-dimensional, la dilatación de un conjunto A por B se define como:

$$A \oplus B = \bigcup_{b \in B} A_b$$

En el procesamiento de imágenes la operación de dilatación a menudo se denomina "relleno", "ampliar" o "crecer".

La operación dual a la dilatación es la erosión. Usando la notación anterior, la erosión de un conjunto A estructurando un conjunto de elementos B mostrará aquellas áreas donde el elemento estructurante encaja el conjunto A. En términos formales, la erosión de un conjunto A por un elemento estructurante B se define como:

$$A \ominus B = \bigcap_{b \in B} A_{-b}$$

En el procesamiento de imágenes el operador de la erosión a menudo se denomina "contracción" o "reducir". En la práctica, las operaciones de dilatación y erosión se utilizan juntas; por ejemplo, una erosión seguida de una dilatación constituye otra operación morfológica denominada **abertura** (opening).

$$A \circ B = (A \ominus B) \oplus B$$

El efecto práctico de las aberturas morfológicas es eliminar detalles de la imagen que son más pequeños que el elemento estructurante sin distorsionar la estructura geométrica de elementos no suprimidos. Por lo tanto, las aberturas tienden a romper istmos estrechos y eliminar pequeñas islas dentro de una imagen binaria (Haralick et al., 1987).

La morfología de escala de grises consiste en la ampliación de estas nociones de conjuntos en dos dimensiones a funciones en tres dimensiones. Requiere la definición de la superficie superior de un conjunto y el umbral de una función.

Para un conjunto A en el espacio tridimensional, donde se consideran las dos primeras coordenadas (x, y) como que constituye el dominio espacial y la coordenada z indicando la superficie, la superficie superior $T [A]$ de un conjunto es el valor más alto de z tal que $(x, y) \in A$.

El umbral de una función f , denotado como $U [f]$, es un conjunto formado por la superficie f y todo lo situado por debajo de la superficie. Para una función dada (imagen en escala de grises) f y un elemento estructurante tridimensional k , la dilatación de escala de grises de f por k se define como la superficie de la dilatación de sus umbras:

La erosión de escala de grises de una función f por un elemento de estructuración k se define como la superficie de la erosión de sus umbras:

Por tanto, podemos definir una abertura (*opening*) en escala de grises de una función f estructurando el elemento k como:

La operación de apertura en escala de grises puede interpretarse geoméricamente como pulsar el elemento estructurante contra su superficie deslizándolo debajo de toda la superficie. La apertura de la superficie por el elemento estructurante es el punto más alto alcanzado por cualquier parte del elemento estructurante a medida que se desliza debajo de la superficie. (Haralick, Sternberg y Zhuang, 1987) (Haralick et al., 1987) (Reutebuch, Schreuder y Andersen, 2001)

El procedimiento sigue los siguientes pasos:

(i) Se aplica un filtro de paso bajo al CHM rasterizado para suavizar la superficie y reducir el número de máximos locales.

(ii) Los máximos locales se localizan usando una ventana móvil circular; un píxel del CHM es etiquetado como máximo local si su valor es mayor que todos los demás valores en la ventana, siempre que sea mayor que una altura mínima por encima del suelo.

(iii) Cada máximo local es etiquetado como "región inicial" alrededor de la cual puede hacerse recrecer una copa de árbol; se extraen las alturas de los cuatro píxeles vecinos del CHM y estos píxeles se añaden a la región si su distancia vertical desde el máximo local es menor que un porcentaje de la altura máxima local y menor que alguna diferencia máxima, definidas ambas por el usuario; este procedimiento se repite para todos los vecinos de las

celdas incluidas en la región, y así sucesivamente hasta que no se añadan más píxeles a la región.

(iv) De cada región identificada, se extraen los ecos de primer retorno LiDAR

(v) Se aplica una envolvente convexa 2D a estos puntos, y los polígonos resultantes se convierten en las copas de árboles finales.

Este proceso no es completamente automático, ya que el tamaño de la ventana móvil, la altura de corte de árboles pequeños/arbustos y los umbrales de diferencia de porcentaje y absolutos de altura deben ser establecidos por el usuario.

Delineadas las copas de árbol individual se han confrontado con la ubicación de los árboles apeados en el inventario forestal de las 12 parcelas que presentan datos atípicos.

En todas ellas se ha observado que:

- Por un lado, el grado de correspondencia entre las copas delineadas y los árboles inventariados es, en todas ellas, bajo o muy bajo; probablemente consecuencia de la poca precisión en la adquisición de las coordenadas del centro de la parcela.
- Por otro lado, se evidencia una acusada heterogeneidad de las masas forestales en las que se insertan. (figura X y ANEXO X)

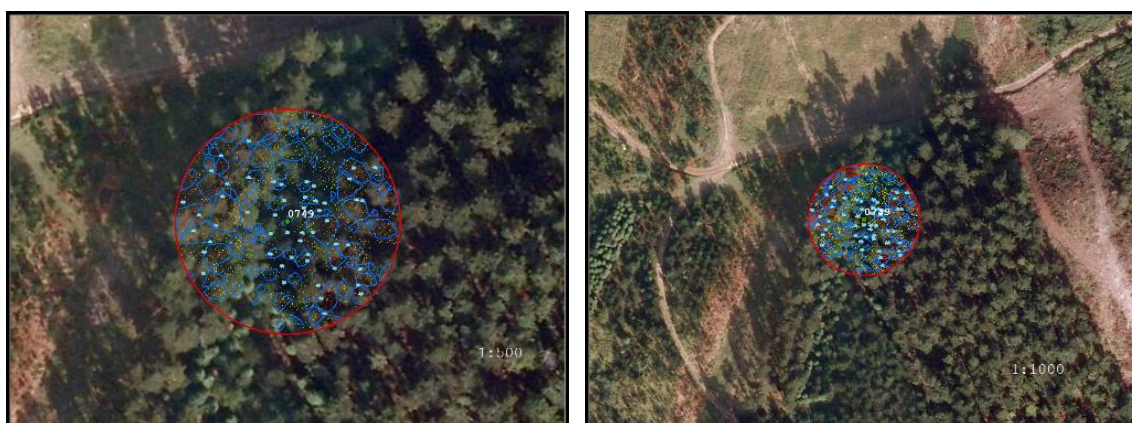


Ilustración 52 Ejemplo de parcelas con datos anómalos

Así pues, se estima que la unión de estos dos factores ha podido ser la causante de que los datos LiDAR y los de biomasa aérea no tengan un buen ajuste en estas parcelas, lo que habría provocado que se muestren como anómalas en el contexto de los modelos de predicción.

Esto nos lleva a tomar la solución menos costosa y más rentable en términos de dotarle de robustez al modelo: eliminar dichas parcelas.

8.4.1. Estimación de modelos sin parcelas anómalas

Una vez eliminadas las parcelas con valores anómalos se han vuelto a estimar los modelos observándose una clara mejoría tanto en el ajuste como en el error (tablas 47 y 48)

Comprobado que el grado de ajuste mejora considerablemente con el recorte de valores atípicos se ha procedido a validar los supuestos estructurales. Como se puede observar en la 48 y en los gráficos de residuos (figuras 53 y 54) ambos modelos han superado los condicionantes de normalidad de los residuos, homocedasticidad, linealidad,

ausencia de autocorrelación y colinealidad mediante los test de inferencia explicados arriba (Breiman, 1996).

Las diferencias en cuanto al rendimiento han resultado significativas tanto en ajuste como en exactitud, por lo que se ha optado por el modelo MARS, a pesar de que la posterior imputación de valores de predicción y la determinación de los intervalos de confianza y de predicción, al conjunto de celdas ocupadas por *P. radiata* en la fase posterior resulte más complicada.

Tabla 46 Test de validación de los supuestos estructurales de los modelos seleccionados sin valores anómalos

Modelo	F	Pr > F	MAPE	SE	R ² Ajustada	R ² Test	Shapiro-Wilk	Breusch-Pagan	Durbin-Watson	VIF	RESET
RLM SIN ATÍPICOS	151,24	<0,0000	0,326	34,370	0,810	0,800	0,076	0,393	2,381	2,722	0,627
SIN ATÍPICOS	214,11	<0,000	0,329	30,634	0,860	0,854	0,252	0,938	1,970	2,075	0,107

Tabla 47 Coeficientes y VIF de las funciones de base de los modelos seleccionados sin anómalos

Variables	β	SE	t	p	Intervalos de confianza (95%)		VIF
					Inferior	superior	
Modelo RLM-2 interacciones SIN ATÍPICOS							
Constante	-20,5	17,3	-1,19	0,237	-54,7	13,7	
ELSK	-125,9	54,6	-2,31	0,023	-233,8	-18,0	4,16
ELSK ²	1304	176	7,39	0,000	955	1652	1,13
EP60	5,255	0,926	5,68	0,000	3,425	7,085	3,93
PAAM	3,279	0,399	8,21	0,000	2,489	4,068	1,67
Modelo MARS-2 interacciones SIN ATÍPICOS							
c	117,25	9,67	12,13	0,000	98,12	136,37	
BF1	4,267	0,711	6,00	0,000	2,861	5,672	2,92
BF2	5,803	0,941	6,17	0,000	3,943	7,664	1,37
BF3	-2,737	0,446	-6,13	0,000	-3,620	-1,854	1,68
BF5	107,0	13,3	8,07	0,000	80,7	133,2	2,33

Tabla 48 Ecuaciones finales

Modelo RLM-2 interacciones SIN ATÍPICOS
BIOMASA AÉREA = -20,5-125,9*ELSK+1304*ELSK ² +5,255*EP60+3,279*PAAM
Modelo MARS SIN ATÍPICOS
BIOMASA AÉREA = 117,25+4,267*BF1+5,803*BF2-2,737*BF3+107,0*BF5

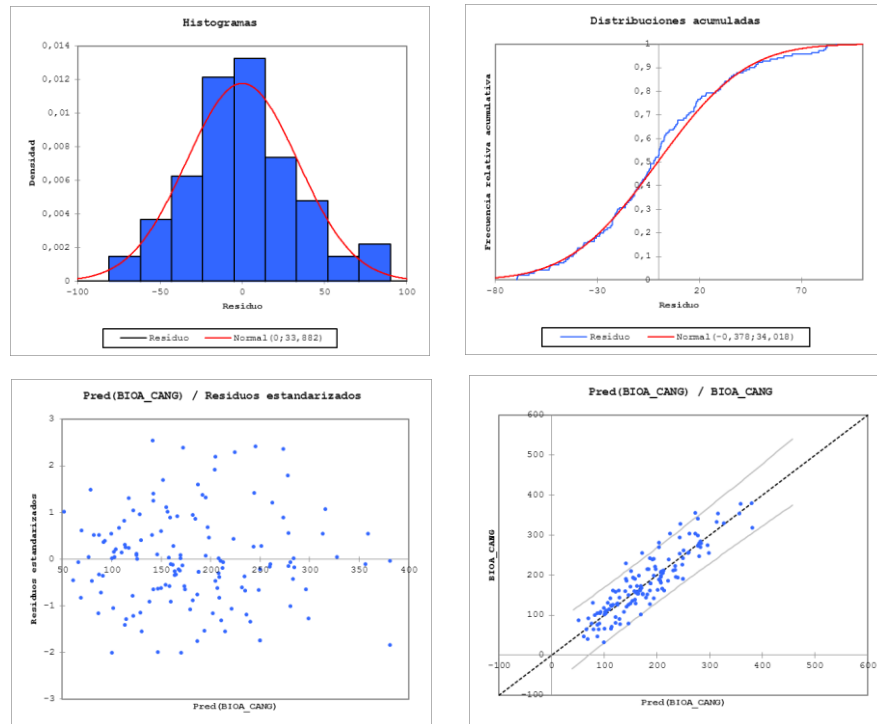


Figura 53 Graficas de normalidad, homocedasticidad y linealidad de los residuos (Regresión Lineal)

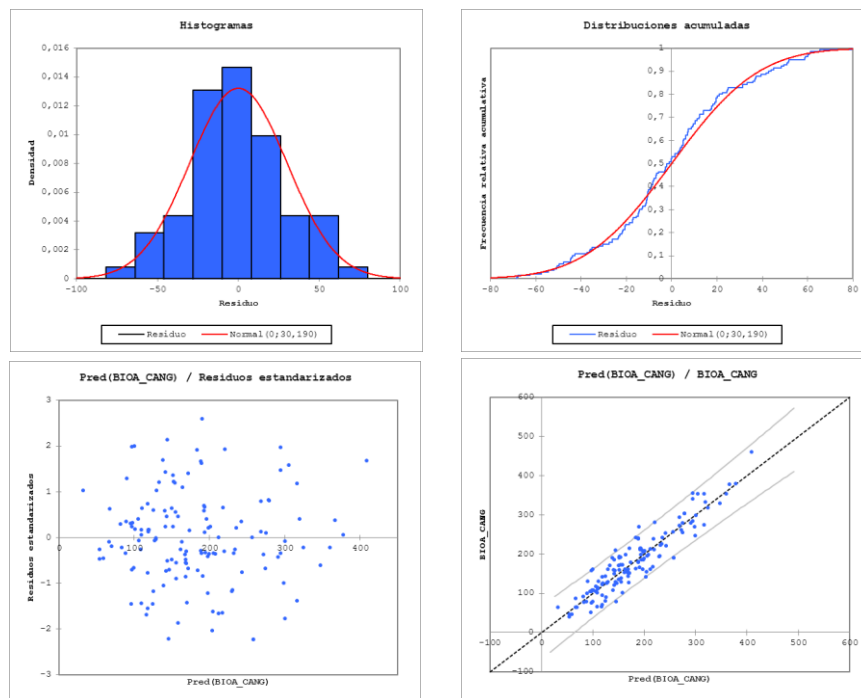


Figura 54 Graficas de normalidad, homocedasticidad y linealidad de los residuos (MARS)

8.5. Cartografiado de las estimaciones y del grado de incertidumbre

Una vez ajustado y validado el modelo definitivo, este se ha empleado para extrapolar las estimaciones para todas las masas de *P. radiata* en estado de Latizal/Fustal cartografiadas durante el IFN4. Para ello se ha generado una malla regular de celdas de superficie igual a la superficie de las parcelas del Inventario, es decir de 0,1936 Has.

Para cada una de estas celdas de 44m de lado se han calculado las métricas que son empleadas como las variables predictoras del modelo definitivo mediante el comando *GridStats* de FUSION que a su vez genera las correspondientes capas raster en un formato ASCII fácilmente incorporable al programa GIS. Este proceso ha supuesto el tratamiento de más de 8000 ficheros .LAS de puntos LiDAR, con una media de 2.639.654 puntos y un tamaño medio de más de 80 MB.

Para poder abordar semejante volumen de información se ha recurrido a crear un conjunto de programas de procesamiento por lotes para, primero segmentar en bloques (Tiles) la totalidad del territorio de Bizkaia y posteriormente estos bloques se han recompuesto para conformar capas únicas para cada variable.

El siguiente paso ha consistido en incorporar las capas generadas en el programa ArcGIS, comprobar la inexistencia de anomalías groseras en los datos espaciales y, a continuación, mediante álgebra de mapas, implementar la ecuación predictiva del modelo, sustituyendo los datos de las variables originales por los nuevos datos generados

en FUSION y producir de esta manera los datos estimados de biomasa/Ha para cada celda (Ene, L. T. et al., 2012; Mitchard et al., 2013; Swatantran, Dubayah, Roberts, Hofton y Blair, 2011).

8.5.1. Incertidumbre de las predicciones

Un modelo predictivo debe incorporar, además de las estimaciones de la variable de respuesta, el grado de incertidumbre que incorpora cada una de ellas. Para ello habitualmente se recurre a la estimación de los intervalos de confianza y los intervalos de predicción para cada una de las celdas.

Esto es posible dado que se han podido confirmar todas las hipótesis de los supuestos estructurales del modelo de regresión lineal, lo que permitirá calcular el grado de incertidumbre mediante las siguientes fórmulas (Peña, 2002).

Intervalos de confianza:

$$\mu_y = Y' \pm tS_{yx} \sqrt{\frac{1}{n} + \frac{(X - \bar{X})^2}{\sum X^2 - \frac{(\sum X)^2}{n}}}$$

Intervalos de predicción:

$$\mu_y = Y' \pm tS_{yx} \sqrt{1 + \frac{1}{n} + \frac{(X - \bar{X})^2}{\sum X^2 - \frac{(\sum X)^2}{n}}}$$

donde: Y'=predicción de Y para un valor dado de X

X =valor dado de X

\bar{x} =media de los valores de X

n =número de observaciones

S_{xy} =error estándar de estimación

t =valor de t para $\alpha/2$ y $\phi=n-2$

La diferencia entre un intervalo de predicción y un intervalo de confianza radica en el error estándar. El error estándar para un intervalo de confianza toma en cuenta la incertidumbre debido a la toma de muestras. La línea se calcula a partir de la muestra será diferente de la línea que habría sido calculada si tuviera la totalidad de la población, el error estándar toma esta incertidumbre en cuenta.

El error estándar para un intervalo de predicción en un individuo de observación tiene en cuenta la incertidumbre debido al muestreo, pero toma también en cuenta la variabilidad de los individuos en torno a la predicción de la media. La SE para la predicción del intervalo será más ancho que el intervalo de confianza y, por tanto, el intervalo de predicción será más ancho que el intervalo de confianza (Peña, 2002).

Mientras que las estimaciones de biomasa para cada celda del modelo raster a partir de la ecuación de regresión resulta fácilmente abordable desde el propio programa GIS mediante herramientas de álgebra de mapas, el cálculo de los intervalos de confianza y de predicción resulta muy complejo, puesto que cada valor de estimación de la variable dependiente requiere de la aplicación de las expresiones citadas. Para poder llevar a cabo dichos

cálculos se ha requerido recurrir a un procedimiento que pasamos a explicar a continuación:

- En primer lugar, las capas raster se han convertido a un formato de entidades de tipo punto y se ha asignado a cada registro resultante un identificador unívoco compuesto por la concatenación de las coordenadas xy de cada uno. Esto, como se verá más adelante permitirá reincorporar la información tratada externamente al programa GIS sin perder sus atributos de georreferenciación.
- En el siguiente paso, se han exportado las capas de las variables predictoras, a sendos ficheros en los que las variables están dispuestas en columnas y los casos en filas para poder ser así procesado en un programa estadístico capaz de realizar los cálculos estadísticos descritos en las ecuaciones X e Y.
- Estos ficheros compuestos por la variable predictora calculada para cada una de las celdas del modelo raster mediante el comando *Gridstat* de FUSION, se han unido en un único fichero a partir del identificador único mencionado.
- Las variables de este nuevo fichero, compuesto por más de 1,5E+6 registros, se han utilizado como valores de imputación para calcular los Intervalos de Confianza y de Predicción además del Error Estándar para cada valor estimado de biomasa.
- Estos valores en forma de nuevas variables se han importado de nuevo al programa ArcGIS y, empleando el identificador único señalado, unidos a uno de los ficheros de entidades de punto que compartía dicho identificador.

- Por último, este fichero se ha convertido de nuevo al formato ráster para elaborar los mapas de error de estimación y los de intervalos de confianza y predicción. (Mapas X e Y)

En la figura X y en el mapa del Anexo X se resumen los datos de la incertidumbre en las predicciones de la biomasa predicha por el modelo de estimación.

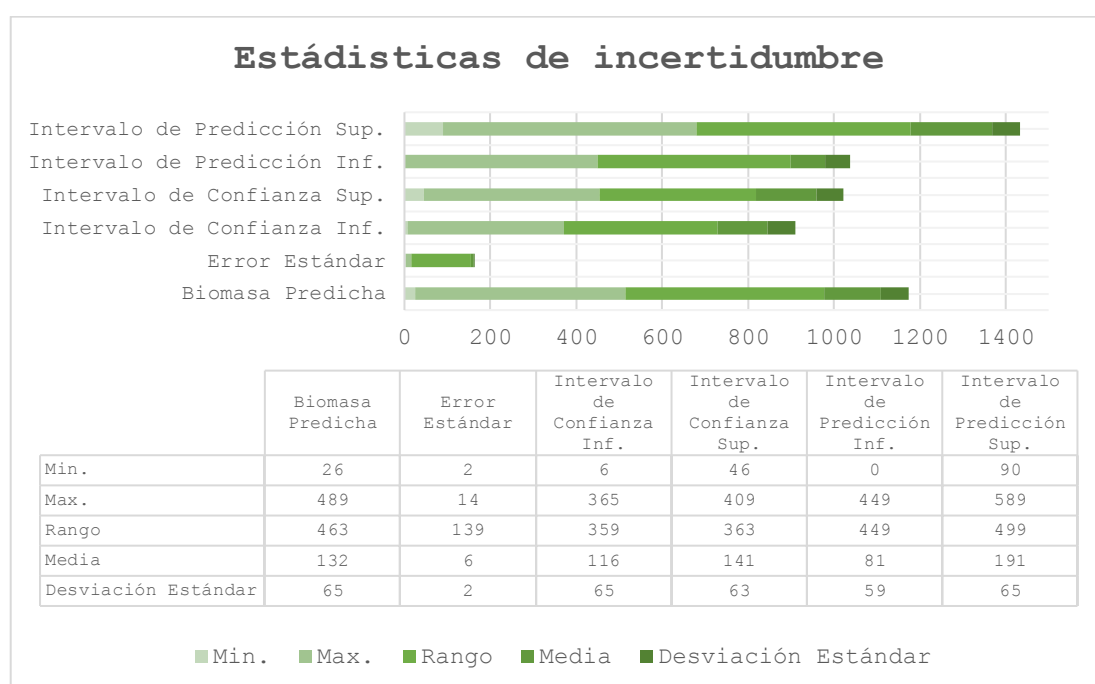


Figura 55 Gráfico de la incertidumbre de las predicciones

8.6. Discusión de los resultados de los modelos de estimación

Los resultados de este trabajo indican que los datos LiDAR pueden ser usados para hacer estimaciones con precisión de la variable biomasa aérea de la especie forestal *P. radiata*. Se demuestra que existe una buena correlación entre la distribución de parámetros relacionados con la altura y la densidad de retornos medida por los sensores y la biomasa calculada a partir de mediciones dasométricas en las parcelas de campo del Inventario Forestal Nacional para el caso del territorio de Bizkaia.

Los resultados obtenidos por procedimientos de ajuste de modelos de regresión lineal y de validación cruzada, están en el rango de las obtenidas por estudios semejantes en cuanto a la precisión y exactitud, ya que explican el 81% de la varianza de la muestra con un error estándar de 36 t/ha, entorno al 21% del valor medio muestral.

Estos valores son similares a los resultados obtenidos por (Andersen, McGaughey y Reutebuch, 2005; Næsset y Gobakken, 2008) con coeficientes de determinación de 0,860 para una variable log-transformada. Estos últimos, siendo similares a los obtenidos en este estudio, han sido aceptados por el U.S. Forest Service como modelos de aplicación en la zona Pacífico-Noroeste (Estados de Oregón y Washington).

En consonancia con estos resultados, para un entorno semejante al área de estudio, (Gonzalez-Ferreiro, Dieguez-Aranda y Miranda, 2012) estimaron la biomasa para *P. radiata* en una muestra compuesta por 54 parcelas de 225 m²

de superficie, distribuidas en una zona de 36 km² en Galicia. En este caso el mejor modelo resultó ser exponencial, utilizando como variables de entrada un percentil de altura, la desviación estándar de la altura y su asimetría. El valor obtenido para el coeficiente de determinación del modelo fue 0,74 y el RMSE de 40,469 ton/ha, semejantes a los obtenidos en este estudio para una zona geográficamente relativamente cercanas y de características similares (en ambos estudios la biomasa media es de \approx 150 ton/ha) y con una densidad semejante de retornos.

Para un estudio realizado para la comarca de Arratia-Nervión (Torre, 2016) obtuvo unos resultados de 0,74 R² y un RMSE de 0,26 (en unidades logarítmicas) empleando un modelo con transformación lognormal de la biomasa. Para la realización de este estudio partió de datos LiDAR con una densidad media de retornos de 0,5 p/m². Las variables predictoras incluidas en el modelo fueron el percentil 95 de la altura de los primeros retornos y la densidad de puntos 3era capa de las 10 en la que había segmentado los datos LiDAR de cada parcela. Cada capa abarcaría una décima parte de la diferencia entre el percentil 95 de la altura y la altura mínima.

De manera análoga, comparando las variables independientes finalmente incluidas en los modelos desarrollados en el presente estudio con las obtenidas por (Andersen et al., 2005; Næsset y Gobakken, 2008) se observa la presencia en casi todos de percentiles de la altura, una aproximación a la fracción de cabida cubierta a través de LIDAR (PoF)

y a los coeficientes de asimetría, relacionados con la estructura vertical de la foresta.

Los resultados también resultan comparables a los de estudios como (Hall, Burke, Box, Kaufmann y Stoker, 2005) que estimaban la biomasa con valores R^2 de 0,74 y un SE de 0,2 (en unidades logarítmicas) mediante una regresión de tipo exponencial. En su caso, se trataba de una zona ubicada en Colorado, USA, para las especies *P. ponderosa* y *Pseudotsuga menziesii* en 14 parcelas con un área de 0,32 ha y una biomasa media medida en campo de 105 ton/ha. Los datos LiDAR tenían una densidad de retornos de 1.23 puntos/m².

(Næsset, 2011) estimó la biomasa para una zona del sur de Noruega de 960 km² de extensión a partir de 35 parcelas de 1000 m² y con una densidad de puntos LiDAR de entre 0,9 y 1,2 puntos/m². Utilizó como variables predictoras de la biomasa el percentil 50 de las alturas de los primeros retornos y el porcentaje de puntos de vegetación por encima de los 2m, del cómputo total de últimos retornos. El modelo empleado es de tipo exponencial con un RMSE de 19,5 y un valor para R^2 de 0,76.

Se ha comprobado que existen distintas posibles fuentes de error, que explicarían las diferencias entre las estimaciones y las observaciones. En primer lugar, hay que considerar que la biomasa aérea de cada pie y su posterior integración en t/ha se ha obtenido a partir de ecuaciones alométricas diferentes, que a pesar de mostrar ajustes por encima del 90% entre componentes de la biomasa y valores

dasométricos como el DN y la altura, introducen inevitablemente una fuente de error inicial a los modelos de estimación:

Por un lado, las elaboradas por el equipo de Montero de Burgos, recogidas en la Monografía 13 (forestal) del INIA, que, como se preveía ha proporcionado los peores valores de ajuste en los modelos iniciales, probablemente debido a las dimensiones y a la heterogeneidad ecológica de los ambientes forestales en los que se obtuvieron las muestras. Otro elemento que ha podido influir en el peor ajuste es que todas las ecuaciones de este estudio tan solo incluyen como variable dasométrica el DN, dejando fuera otros como la altura.

Por otro las elaboradas específicamente para la especie objeto de estudio y en un ambiente bioclimático semejante en el estudio de Pérez-Cruzado que, a pesar de proporcionar ajustes globales entre la biomasa y las variables dasométricas superiores al 90% han proporcionado unos resultados algo peores que los obtenidos a partir de la tercera ecuación alométrica.

La tercera ecuación alométrica analizada y que ha resultado dar los mejores valores de ajuste y exactitud en la predicción es la obtenida en el estudio de (Canga et al., 2013). En este, al igual que el anterior la especie estudiada es el *P. radiata*

en un ambiente semejante al del área de estudio y unos valores de ajuste semejantes, de en torno a 96%.

Por último, hay que señalar que ambos casos la biomasa aérea de los pies mayores incluidos en las parcelas muestrales pertenecientes a especies diferentes al *P. radiata*, se han obtenido, cuando esto ha sido posible, a partir de las ecuaciones de (Montero González et al., 2006). Algunas pocas parcelas contenían pies cuya especie no disponía de ecuación alométrica y por lo tanto no han podido ser integrados en la suma total de la biomasa aérea de esa parcela.

Otra fuente posible de propagación de error en los modelos de estimación podría asociarse con la correspondencia geográfica de las parcelas de campo con los datos LiDAR recortados, la cual no se puede garantizar en todos los casos. Pueden existir desplazamientos, que aun siendo de poca magnitud, provocan diferencias de la distribución de alturas entre ambas ubicaciones y, teniendo en cuenta que los modelos predictores están basados en dicha distribución, una ubicación inexacta genera una segunda fuente de sesgo en los modelos y en por lo tanto en las predicciones.

CAPÍTULO X

CARTOGRAFÍA E INVENTARIO

9. CARTOGRAFIA E INVENTARIO DE LA BIOMASA

En este capítulo se muestran y analizan los resultados de la estimación de la biomasa forestal en sus tres niveles: Biomasa Forestal Potencial Total y Biomasa Forestal Potencial Accesible a nivel provincial, y la Biomasa Forestal Potencial Disponible a las escalas provincial, comarcal y municipal.

9.1. Cartografía e inventario a nivel provincial de la Biomasa Forestal Potencial Total y el CO² almacenado.

Con el modelo de estimación desarrollado en el capítulo anterior hemos dado el primer paso de la metodología para la estimación de la biomasa forestal aérea potencialmente accesible de *P. radiata* para usos energéticos en Bizkaia: estimar la biomasa aérea potencial total y su localización espacial en términos de t/ha.

El siguiente paso necesario es obtener de la cantidad acumulada de biomasa aérea potencial total de *P. radiata* de Bizkaia a escala provincial, para lo cual será necesario transformar los datos cartografiados - en los que cada píxel está expresado en t/ha - en datos en el que se exprese la cantidad de biomasa aérea de cada píxel para poder posteriormente agregarlos a nivel provincial. Para ello se ha empleado el módulo Spatial Analyst de ArcGIS, concretamente la herramienta Raster Calculator, donde se ha multiplicado la capa obtenida por el área del píxel de 44 m de lado equivalente al área de las parcelas muestrales

de 25m de radio, concretamente por 0,1936 ha. El resultado se representa en la cartografía de la Figura X.

Tal y como se observa en esta figura y en la gráfica de distribución (fig X), el valor máximo de biomasa contenido en un píxel es de ≈ 75 t, la media es de ≈ 25 t, el valor mínimo es de ≈ 5 t y la desviación estándar de ≈ 12 t.

Tal y como se recoge en el capítulo de objetivos y metodología, uno de objetivos secundarios de esta tesis consiste en realizar una estimación de la cantidad de dióxido de carbono almacenado por la parte aérea de las masas forestales de *P. radiata* en Bizkaia. Para ello se ha seguido el método propuesto por (Ibáñez, Vayreda y Gracia, 2002; Kollmann, 1959) en (Montero González et al., 2006). Según este método, esta se puede conocer a partir de la cuantificación de la biomasa seca de una especie arbórea. Según (Kollmann, 1959) la composición de la madera es similar en las distintas especies leñosas, así como también dentro de un mismo árbol, en sus diversas partes, tronco y ramas. Este autor indica que todas las maderas contienen, aproximadamente, un 50% de carbono. El Panel Intergubernamental sobre el Cambio Climático (Houghton, 1996) recomienda, en el caso de no existir datos específicos, considerar también este porcentaje.

Utilizando los datos sobre porcentaje de carbono en la madera que ofrece el Centre de Recerca Ecològica i Aplicacions Forestals (CREAF) (Ibáñez et al., 2002), completados para las especies sin datos con el valor que adelanta Kollmann, se obtiene el peso de carbono acumulado

en cada especie forestal. Para el *P. radiata* resulta ser de $\approx 49,7\%$ de carbono.

Mediante la proporción entre el peso de la molécula de CO_2 y el peso del átomo de C que la compone se halla la relación que se emplea para obtener los kg de CO_2 equivalente a partir de la cantidad de carbono presente en la biomasa ($44/12 = 3,67$).

Multiplicando los valores de biomasa por el contenido en carbono y por la relación *molécula de CO_2 -peso átomo C* se obtiene los valores de CO_2 acumulado por cada especie. En el caso del Pino radiata este sería de 1,824 t de CO_2 por tonelada de biomasa. (Montero González et al., 2006).

De la estimación de la BFPT aérea de *P. radiata* en masas de latizal y fustal de Bizkaia se deduce que estas almacenaban en 2011 $\approx 13.368.260$ t de CO_2 , que en términos comparativos equivaldría al 17,14% del almacenado por esta especie en España y al 32% del almacenado en la CAE (Montero González et al., 2006). Por poner alguna referencia que nos permita visualizar lo que representa esta cantidad señalaremos dos ejemplos ilustrativos: una tonelada de CO_2 a unos 10 grados centígrados tendría un volumen de unos 534 m³, lo que corresponde aproximadamente al volumen de una piscina de 10 m x 25 m x 2 m. Es aproximadamente lo que emite para alimentarse, calentarse, desplazarse un europeo medio en algo más de un mes (Knörr, 2008).

9.2. Estimación de la Biomasa Forestal Potencial Accesible.

Tal y como ha quedado definida en el capítulo de metodología, la Biomasa Forestal Potencial Accesible (BFPA) sería la resultante de aminorar la total considerando restricciones de índole ecológica, económica o de mecanización y de eficiencia en la recogida.

Entre los factores de aminoramiento que se mencionaban se han tenido en cuenta los siguientes:

9.2.1. Limitantes y condicionantes relacionados con la pendiente:

Se ha considerado con (Cabrera et al., 2011) que el factor de limitación de las pendientes viene ligado a la productividad climática ($\text{m}^3/\text{ha}/\text{año}$) de acuerdo con el índice de Paterson (Tabla 50). Para Bizkaia según este índice, empleado para la elaboración del Mapa de la Productividad Potencial Forestal de España, (Palomares y Serrano, 2000) las clases que podemos encontrar son las que aparecen en la leyenda del mapa de la figura X.

En zonas de altas productividades (clases I y II) se ha fijado un máximo de pendiente del 50%. Para el resto de zonas de productividad se ha fijado el límite en el 30%. En total, tan solo 15 de las casi 55.000 ha de *P. radiata* han quedado afectadas por esta limitación en toda Bizkaia.

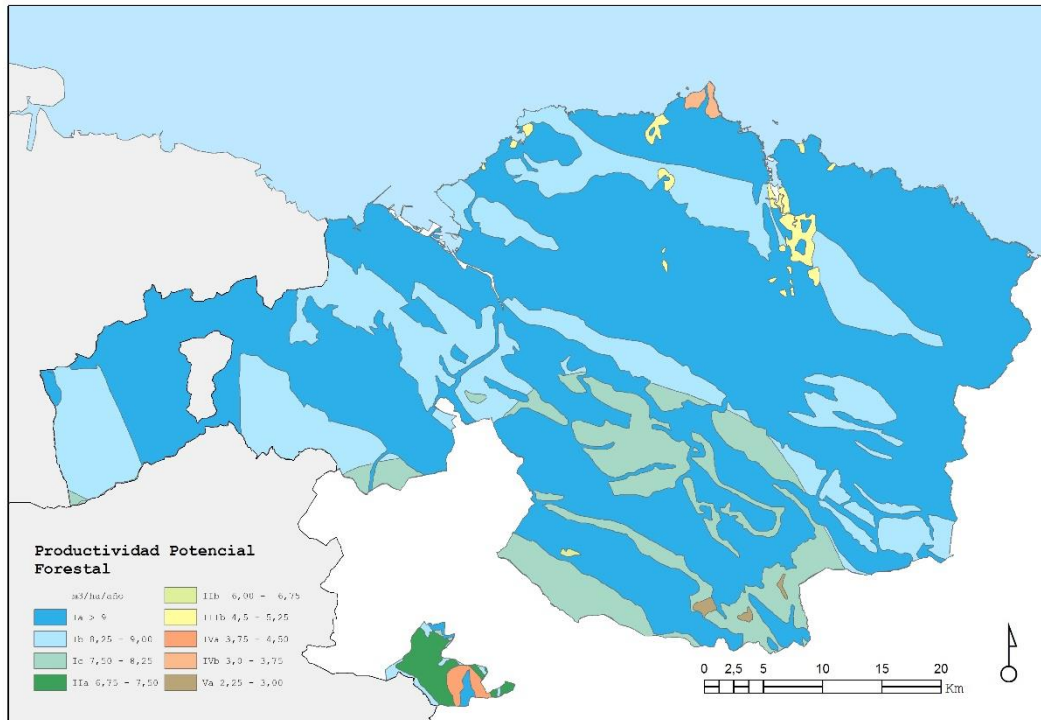


Figura 56 Mapa de productividad forestal (Palomares y Serrano, 2000) Elaboración propia

Tabla 49 Zonas de productividad de acuerdo con los valores alcanzados por el índice de Paterson (Palomares y Serrano, 2000)

ETIQUETA	PRODUCTIVIDAD-CLIMÁTICA (m ³ /ha/año)	CLASPROD
Ia	> 9,00	1
Ib	8,25 - 9,00	2
Ic	7,50 - 8,25	3
IIa	6,75 - 7,50	4
IIb	6,00 - 6,75	5
IIIa	5,25 - 6,00	6
IIIb	4,50 - 5,25	7
IVa	3,75 - 4,50	8
IVb	3,00 - 3,75	9
Va	2,25 - 3,00	10
Vb	1,50 - 2,25	11
VIa	1,00 - 1,50	12
VIb	0,50 - 1,00	13
VII	< 0,50	14

$$I = v.f.p.g / 12.a$$

donde:

I = valor del índice

v = temperatura media de mes más cálido en °C

a = rango anual de temperaturas, estimado por la diferencia entre la media de las máximas del mes más cálido y la media de las mínimas del mes más frío, expresadas en °C

p = Precipitación media anual, en mm

f = factor de insolación, estimado mediante el cociente $2.500/(INSOL+1000)$

en el que INSOL es la insolación media anual, expresada en horas.

g = Duración del periodo vegetativo. De acuerdo con el criterio de Gaussen, se considera como mes activo para la vegetación forestal aquel en el que la precipitación (expresada en mm) iguala o supera al doble de la temperatura media de ese mes (en °C) siempre que ésta sea igual o superior a 6°C.

La productividad potencial, en m³ de madera /ha/año está ligada a este índice mediante la expresión logarítmica.

$$PROD = 5,3 \log I - 7,4$$

9.2.2. Limitantes y condicionantes relacionadas con la distancia a desembosque:

El coste de extracción de la biomasa depende tanto de la naturaleza del terreno como de la distancia de saca o desembosque (Richardson et al., 2006). Por ello se propone modelar el efecto que tiene la accesibilidad de una zona para el aprovechamiento de la biomasa residual mediante la distancia entre la zona de corta de la biomasa y las pistas, caminos o carreteras donde habitualmente se sitúan las zonas de acopio.

Los datos de la relación productividad-distancia de desembosque se obtuvieron a partir de los resultados

mostrados en los trabajos de (Richardson et al., 2006; Sanz y Piñeiro, 2003). En ellos se analiza la variación de la productividad en función de la capacidad de carga de la maquinaria y de la distancia de desembosque y de la especie y la forma de preparación del residuo.

Del estudio de estos datos se concluyó que la rentabilidad era máxima cuando las distancias de desembosque eran ≤ 100 m, disminuyéndose de una forma proporcional hasta llegar a distancias ≥ 3000 m, en las que el beneficio es nulo. De esta forma se asignaron factores de ponderación de 1 a 0, siendo 1 la distancia ≤ 100 m, reduciéndolo proporcionalmente hasta el valor 0 para distancias ≥ 3000 m (Tabla 51).

Tabla 50 Valores de minoración de la biomasa en función de la distancia al desembosque

Distancia a desembosque	Minoración
0 - 100	1
100 - 200	0,97
200 - 300	0,93
300 - 400	0,9
400 - 500	0,86
500 - 600	0,83
600 - 700	0,79
700 - 800	0,76
800 - 900	0,72
900 - 1000	0,69
1000 - 1500	0,6
1500 - 2000	0,43
2000 - 2500	0,26
2500 - 3000	0,07
≥ 3000	0

Para la obtención de este índice se utilizó como información base la capa vectorial que contiene los ejes de la red viaria extraída de la Base Topográfica Armonizada 1:5000 de Geoeuskadi (Infraestructura de Datos Espaciales (IDE) de Euskadi) de la que se emplearon las categorías que contienen la red de pistas forestales, caminos y carreteras de Bizkaia. A partir de esta capa se calculó una capa de distancias euclidianas en metros a la vía de desembosque más cercana. Después de analizar el rango de distancias reales a la red viaria en Bizkaia se ha comprobado que la distancia máxima es de 1900 m por lo que el factor de minoración potencial más bajo que se pudo aplicar fue el de 0,43. Posteriormente se reclasificaron los intervalos de distancia señalados asignándoles los factores de ponderación señalados. Por último, se aplicaron estas ponderaciones a la capa de biomasa surgida en al paso anterior.

En términos absolutos, la minoración de la BFPA resultante de aplicar este factor ha sido de en torno 267500 t, aproximadamente el 1% del total.

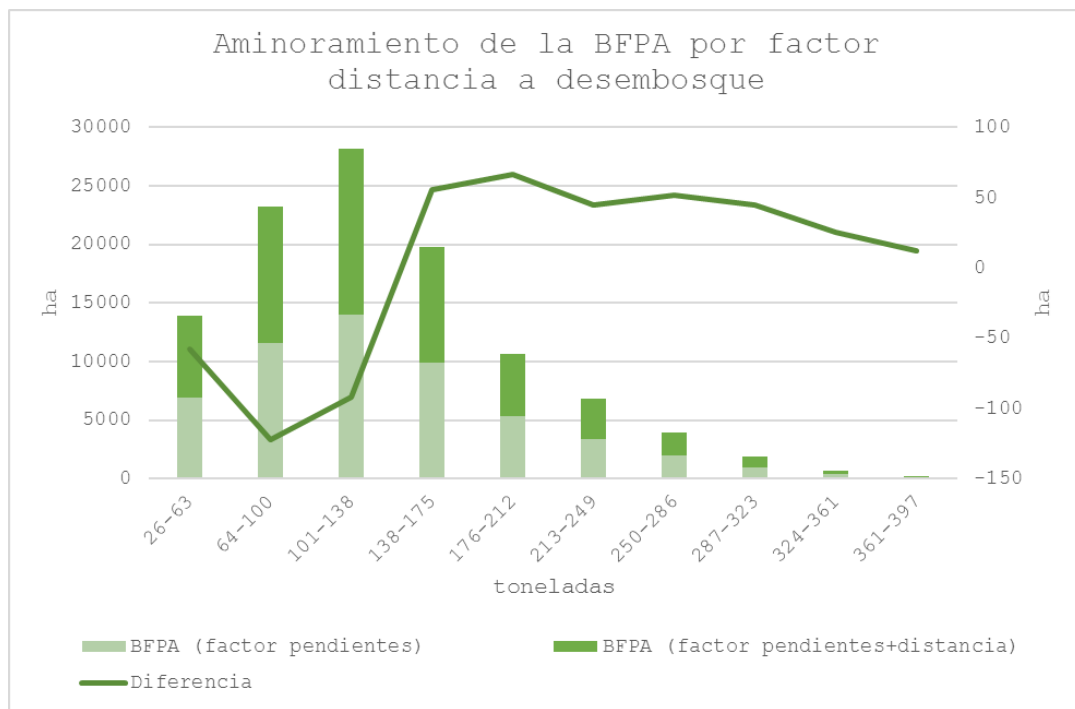


Figura 57 Aminoramiento de la Biomasa Forestal Potencial Accesible. Factor distancia.

En cuanto a cómo ha afectado a las diferentes parcelas teniendo en cuenta su volumen, observamos en la figura, donde se representa la distribución de los volúmenes de BFPTA antes y después de aplicar el factor de aminoramiento de la biomasa por distancia al desembosque, que las parcelas con menor cantidad de biomasa -en torno o por debajo de la media (≈ 131 t)- son las que en términos relativos se han visto más afectadas, mientras que el porcentaje de las celdas que contenían un volumen mayor de biomasa han aumentado. Esto parece responder a la lógica de que las masas más productivas son las que poseen una mejor infraestructura de accesos y están por lo tanto mejor acondicionadas para la implementación de los distintos sistemas logísticos de explotación de los recursos forestales, algo entendible en una región con la tradición

de explotación forestal, especialmente maderera, como es Bizkaia.

9.2.3. Limitantes y condicionantes jurídico-administrativos:

La existencia de espacios protegidos y de normas sectoriales en materia de montes establecen a menudo limitaciones a la explotación o condicionan prácticas forestales en las masas de especies potencialmente aprovechables. En el caso de Bizkaia son varios los ámbitos en los que se produce este tipo de limitaciones.

En las zonas en los que el aprovechamiento forestal no está permitido o, estando tolerado se perdería tal consideración en el momento de una corta final, se ha optado por suprimir dichas parcelas del cálculo de la BFPA.

Resulta más difícil valorar y ponderar la biomasa que resulta accesible en aquellas zonas condicionadas por categorías de ordenación en las que el aprovechamiento forestal no está prohibido, pero sí condicionado. Un ejemplo de este tipo de situaciones la podemos encontrar en el "Artículo 4.4.3.21.- *Uso forestal*" del Plan Rector de Uso y Gestión de la Reserva de la Biosfera de Urdaibai:

"7. Las actividades forestales garantizarán la conservación de los recursos edáficos e hídricos. Para ello, se limitarán al mínimo preciso el uso extensivo del fuego, los desbroces extensivos y los movimientos de tierras y se propiciará la plantación de carácter manual sobre la mecanizada y los métodos progresivos de corta, como entresacas, cortas por bosquetes o aclareos sucesivos, sobre la corta a hecho."

(...)iv. El acceso a las plantaciones forestales productivas, así como la extracción del material procedente de las cortas se realizará, preferentemente, a través de servidumbres existentes. En ningún caso se realizará el arrastre de troncos a través de masas forestales autóctonas (robledal, sauceda, aliseda, encinar...).

v. Tras una corta no se destocará, salvo que se realice con fines de restauración ambiental. Queda prohibida la utilización de bulldozer y las técnicas de laboreo que impliquen grandes movimientos de tierra o mezclas de horizontes del suelo.

(DECRETO 139/2016, de 27 de septiembre, por el que se aprueba el Plan Rector de Uso y Gestión de la Reserva de la Biosfera de Urdaibai, 2016).

En este tipo de zonas se ha optado por una minoración de la biomasa accesible en un 50%. Se entiende que no se puede descartar su aprovechamiento como fuente de energía renovable, pero el aumento de los costes de explotación que vienen de la mano del empleo de los sistemas logísticos necesarios para cumplir con la normativa puede desincentivar a los productores a movilizar este recurso en forma de biomasa de árbol completo.

No podemos descartar, sin embargo, que se produzcan ciertas "laxitudes" en la aplicación de la normativa, como pusieron de manifiesto (Ainz Ibarrondo y González Amuchastegui, 2008) al detectar 1.108 expedientes mediante los que el Servicio de Montes de la DFB autorizaba el correspondiente aprovechamiento: "corta a hecho" o "clara" entre 1993 y 2003 en los municipios de Muxika y Mendata, ambos en la cabecera del río Oka. Zonas en las que según el Art. 93 del Capítulo VII de las Normas de ordenación directa para

las Áreas de Protección, Interés Agrario y Forestales del PRUG de 1993: quedaban totalmente prohibidas las cortas a hecho o matarrasa en superficies continuas superiores a 5 Has. o en terrenos con riesgo evidente de erosión...” (Ainz Ibarrondo y González Amuchastegui, 2008).

Se han incluido dentro de estos factores los siguientes Espacios naturales y zonas del PTS Agroforestal:

- Reserva de la Biosfera de Urdaibai
- Parque Natural del Gorbeia
- Parque Natural de Urkiola
- Zona de Especial Conservación de Armañon
- Zona de Especial Conservación de Ordunte
- Áreas Erosionables del Plan Territorial Sectorial Agroforestal

En la **Reserva de la Biosfera de Urdaibai** a partir de las determinaciones del PRUG respecto a los usos y aprovechamientos forestales de especies de repoblación de crecimiento rápido se han considerado las siguientes zonas limitantes y condicionantes (tabla 52 y fig. 58)

Tabla 51 Condicionantes y limitantes en la Reserva de la Biosfera de Urdaibai

CÓDIGO	LIMITANTES
B1	Área de Protección de la Ría
B2	Área de Protección del Litoral
B3.2	Zona de ecosistemas íntegros o poco degradados de del Área de Protección de Encinares Cantábricos
B3.4	Zona de praderas y cultivos del Área de Protección de Encinares Cantábricos
B4.1	Zonas de Protección de la Red Fluvial
B6	Área de Bosques Autóctonos

N1.1	Zonas intermareales o supramareales constituidas por fangos con o sin vegetación y zonas de marisma
N1.2	Zonas intermareales o supramareales constituidas por fangos o zonas de marisma aisladas del sistema de circulación hídrica mediante el empleo de lezones munas o muros de contención
N1.3	Zonas supramareales aisladas mediante el empleo de lezones munas o muros de contención del sistema de circulación hídrica con ocupación urbana
N1.4	Zonas intermareales o supramareales de arenas
N1.5	Zonas submareales constituidas por el cauce de la Ría.
N2	Área del Litoral
N3.1	Zona de interés naturalístico especial
N3.2	Zona de ecosistemas íntegros o poco degradados
N3.4	Zona de praderas y cultivos
N3.5	Zona de equipaminetos. Kantera Gorria
N4	Área de la Red Fluvial de Urdaibai
N5.1	Zonas de Patrimonio Histórico-Arquitectónico
N5.2	Zonas de Patrimonio arqueológico
OPUM	Ordenación sujeta al planeamiento urbanístico
T1.A1	Zonas de Alto valor agrológico
T3.NR	Zona de Núcleo Rural
T4.ECR	Zonas públicas o privadas con o sin asentamientos constructivos precisas para el equipamiento comunitario que han de disponerse en suelo no urbanizable.
T4.IS	Zonas destinadas a soportar las infraestructuras y servicios de la comunidad
CÓDIGO	CONDICIONANTES
B3.3	Zona de ecosistemas degradados y repoblaciones forestales del Área de Protección de Encinares Cantábricos
B4.2	Zonas de Protección de cuencas vertientes a la Red Fluvial
B5	Área de Protección Paisajística Territorios de Alta Vulnerabilidad Visual y Escenográfica
B7	Área de Montes declarados como de Utilidad Pública (MUP)
N3.3	Zona de ecosistemas degradados y repoblaciones forestales
T1.PRT	Zonas de Paisaje Rural de Transición
T2.F1	Zonas forestales con pendientes entre el treinta y el cuarenta y cinco por ciento (30-45%)
T2.F2	Zonas forestales con pendientes entre el cuarenta y cinco y el sesenta por ciento (45-60%)

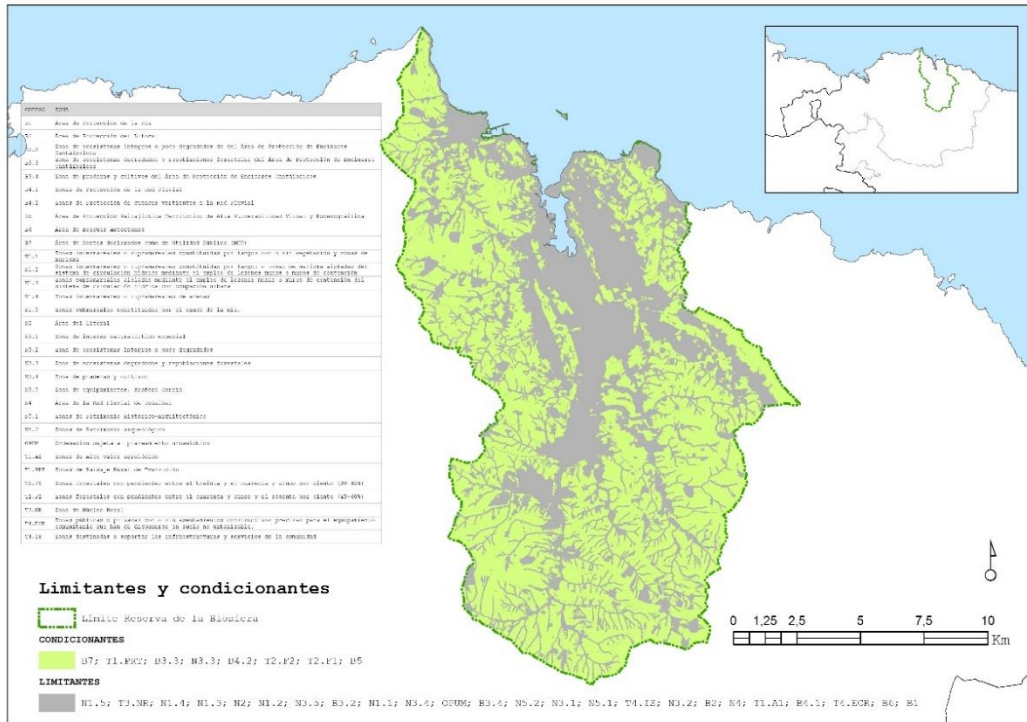


Figura 58 Limitantes y condicionantes de la Reserva de la Biosfera de Urdaibai (Elaboración propia)

Para el caso de los **Parques Naturales de Gorbeia y de Urkiola**, ambos han visto cómo eran recientemente designados Zonas de Especial Conservación por sendos decretos del Departamento de Medio Ambiente y Política Territorial del Gobierno Vasco, pero ninguno de los dos tiene aprobados definitivamente los correspondientes Planes Rectores de Uso y Gestión, por lo que aún se encuentran en vigor los aprobados en 1998 y 2006 respectivamente. Las determinaciones de estos planes respecto al aprovechamiento forestal de especies de ciclo corto, como es el caso del *P. radiata*, permiten considerar

los limitantes y los condicionantes en las zonas que se recogen en la tabla 53 y la figura 59

Tabla 52 Condicionantes y limitantes en los Parques Naturales de Urkiola y Gorbeia

ESPACIO NATURAL	LIMITANTES	CONDICIONANTES
Parque Natural de Gorbeia	Zona de reserva	Zona de Progresión Ecológica
	Zona de Conservación Activa I y II	Zona de potenciación ganadera-forestal
	Zona de Protección	
Parque Natural del Urkiola	Zona de especial protección	Zona de restauración ganadera
	Zona de acogida	Zona forestal-ganadera
		Zona periférica de protección

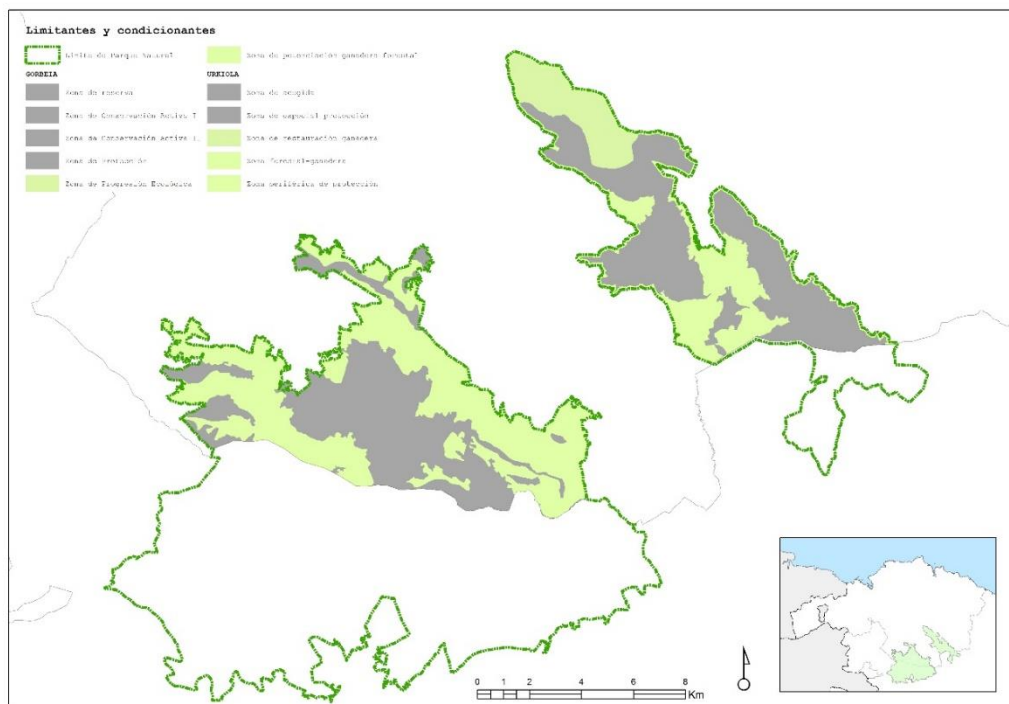


Figura 59 Limitantes y condicionantes de los Parques Naturales de Urkiola y Gorbeia (Elaboración propia)

En el caso **Armañón y Ordunte**, recientemente designados Zonas de Especial Conservación (marzo de 2015 y noviembre de 2014 respectivamente) tan sólo la ZEC de Armañón tiene aprobado el Plan Rector de Uso y Gestión. Ambas zonas reúnen características comunes en cuanto a la problemática ambiental y las necesidades de conservación y regeneración. Ambas zonas comparten igualmente, como hecho diferencial respecto de buena parte del territorio forestal de Bizkaia, especialmente para el caso del pino insigne, una buena proporción de espacios de titularidad pública (Montes de Utilidad Pública), que en muchos municipios del entorno cuentan con Planes de Gestión de dichos Montes. Buena parte de ellos se encuentran en régimen de consorcio con distintas instituciones, tanto públicas como privadas.

Las determinaciones del PRUG de Armañón, y los Planes de Gestión Forestal de los Montes de Utilidad Pública respecto al aprovechamiento forestal de especies de ciclo corto, como es el caso del *P. radiata*, permiten considerar las siguientes limitaciones y condiciones en las zonas que se recogen en la tabla 54 y la figura 60:

Tabla 53 Condicionantes y limitantes en las Zonas de Especial Conservación de Armañón y Ordunte

ZEC	LIMITANTES	CONDICIONANTES
Armañón	Zona de aprovechamiento ganadero extensivo	Zona de restauración ecológica
	Zona de protección estricta	
	Zona de evolución natural	
	Zona de uso público	
Ordunte	Zona de aprovechamiento ganadero extensivo	Zonas de evolución natural
	Zona de protección estricta	Zonas de restauración ecológica

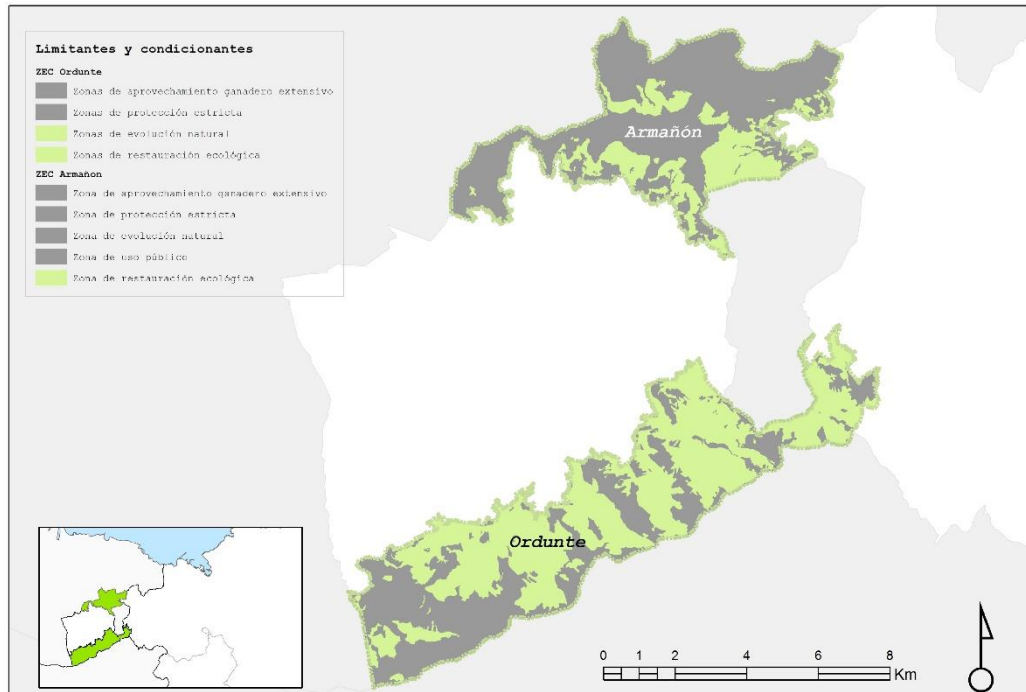


Figura 60 Mapa de limitantes y condicionantes de las Zonas de Especial Conservación de Ordunte y Armañón (Elaboración propia)

Por último, se han incluido, por un lado, como condicionantes, y por lo tanto minorarán la cantidad de BFPA, las categorías de ordenación del PTS Agroforestal vasco que consideran como admisible la práctica forestal, pero las supeditan a análisis de las afecciones generadas y a la incorporación de medidas correctoras. Por otro lado, se ha incluido como limitantes el condicionamiento superpuesto "Áreas erosionables". El art. 56 establece que:

"(...) se mantendrá, con criterio protector-restaurador, la cubierta arbórea y arbustiva cuando ésta exista en

la actualidad, o se introducirá y extenderá la misma en el caso de contar con cubierta vegetal rala, como elemento fundamental de protección frente a los fenómenos erosivos. Se evitarán aquellas actividades que afecten a la estabilidad del suelo, extremando el cuidado de las prácticas agroforestales necesarias en estas zonas. 3. Las actuaciones en estas zonas requerirán de la elaboración de Planes de Gestión Hidrológico- Forestal.”

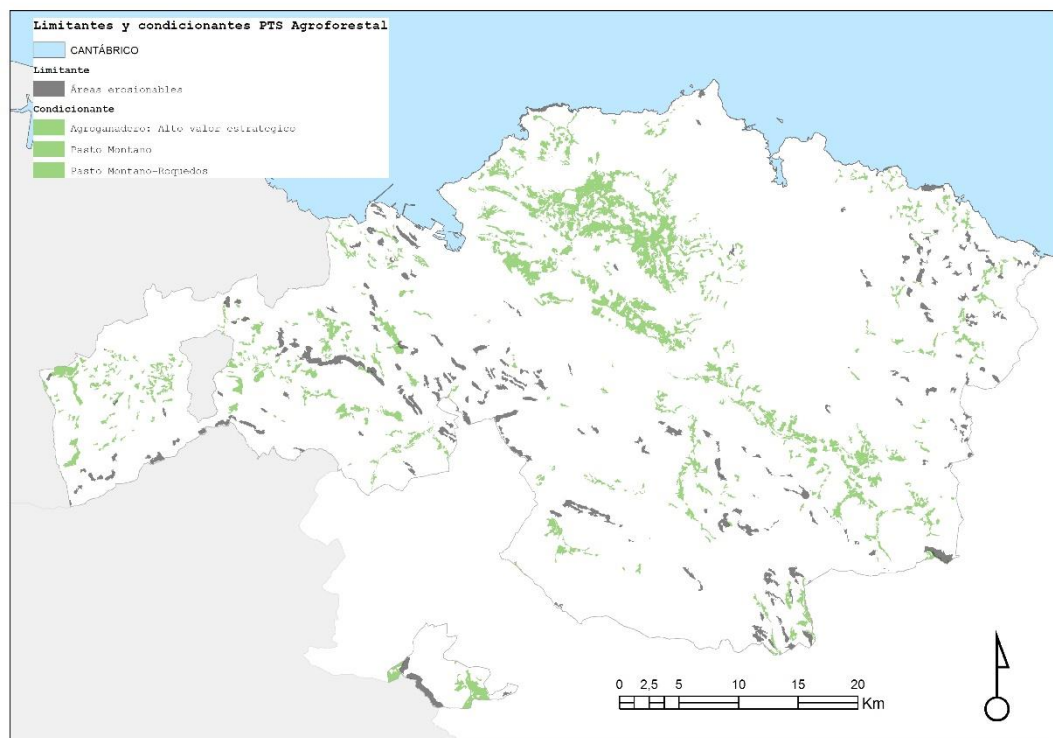


Figura 61 Mapa de limitantes y condicionantes del Plan Territorial Sectorial Agroforestal (Elaboración propia)

9.2.4. Coeficiente de recogida.

En los procesos de recogida de la biomasa se deben dejar restos en el terreno por razones ecológicas además de la posibilidad de no ser recogido el 100% del material (mayor dificultad por pendiente, etcétera). El coeficiente de recogida supone pues, el porcentaje final de restos que se extrae de las masas tras cada tratamiento. En el estudio Técnico de Evaluación de Energía de la Biomasa se fija este coeficiente en un 65% (Cabrera et al., 2011). De esta manera, a falta de estudios que determinen con más exactitud este parámetro y su relación con, p.ej. las pendientes, se ha optado por minorar la cantidad de biomasa potencial accesible resultante de aplicar los limitantes y condicionantes anteriores en un 45% para cada tesela.

9.3. Biomasa Forestal Accesible. Cartografía e inventario a escala provincial, comarcal y municipal.

Siguiendo a (Cabrera et al., 2011) hemos definido la biomasa disponible como aquella biomasa accesible que no tiene un uso maderero, es decir, que no entra en competencia con las cortas madereras que se llevan a cabo en la actualidad.

El cálculo de esta biomasa se ha llevado a cabo para cada tesela a partir de la biomasa accesible minorada por un coeficiente de cortas actuales de madera (CCA). Este sería el resultante del cociente entre las cortas actuales de madera y la disponibilidad de madera estimado por el IFN. Este coeficiente, que puede variar en función de la fracción de biomasa que sea considerada, será igual a 0 para el caso de que se estime el aprovechamiento del árbol

completo. Esto supondría que, en aquellas superficies donde actualmente se llevan a cabo cortas madereras únicamente se obtendrían restos. Por tanto, es necesario definir dos ecuaciones de cálculo en función de si se desea considerar un escenario con biomasa procedente de restos (en cuyo caso se considera la superficie donde se llevan a cabo las cortas madereras de las cuales se extraen los restos) o de árboles completos (caso en el que se considera la superficie sobre la que no se llevan a cabo aprovechamientos maderos, de modo que no entren en competencia) (Cabrera et al., 2011). En el caso del *P. radiata*, (Cabrera et al., 2011) fijan el CCA en la Cornisa Cantábrica en un 65% para fracciones de biomasa de restos y pies de diámetro inferior a 7 cm.

De esta manera, se han generado dos escenarios diferentes:

- Uno de aprovechamiento de la biomasa de restos compuesta por rabezones, ramas y ramillas procedentes de los distintos tratamientos y árboles completos con diámetro inferior a 7 cm, básicamente producto de clareos, en el cual la biomasa accesible estimada en los procedimientos anteriores se ha minorado en un 35%.
- Un segundo escenario de aprovechamiento del árbol completo para usos energéticos en el que no se ha aplicado ninguna minoración a la BFPA.

En la tabla 55 se recoge los datos de volumen de Biomasa Forestal Potencial Accesible de restos a las escalas provincial, comarcal y municipal. Junto a este dato se incluyen otros que se consideran pueden ayudar a valorar mejor estas magnitudes: el dato de tonelada equivalente de petróleo (tep) y la superficie susceptible de ser explotada para fines energéticos a partir de la biomasa de restos.

Por poner un ejemplo que sirva para ilustrar la capacidad energética de esta biomasa señalar que cada tep equivale a aproximadamente 11 MWh, y que, según datos publicados por Red Eléctrica Española, en 2011 la producción de todas las centrales nucleares españolas fue de 57.731.000 MWh, diez millones de MWh menos que el potencial energético que se podría aprovechar de la biomasa de restos de *P. radiata* en Bizkaia.

A nivel comarcal se observa que excepto el Gran Bilbao y la comarca de Plentzia-Mungia, el resto de comarcas tiene un potencial similar. Destaca la comarca de Arratia-Nerviión con casi un 23% del total de BFPA de restos de toda la provincia. Precisamente en esta comarca podemos encontrar algunos de los proyectos energéticos más importantes de la provincia, en el municipio de Orozko, donde existen dos instalaciones: una que abastece de calor al Polideportivo y a la escuela Municipal con una potencia de 400 kW y una segunda de "Distric Heating" que abastece de calor a más de 430 usuarios repartidos en 34 bloques de viviendas con una potencia de 1.400 kW.

Tabla 54 Inventario de la Biomasa Forestal Potencial Disponible a escala provincial, comarcal y municipal

	BPPA restos		BPPA árbol completo	
	BFPD (t)	Tep	BFPD (t)	Tep
DATOS POR PROVINCIA				
BIZKAIA	13.714.947	6.123.761	21.099.920	9.421.170
DATOS POR COMARCA				
DURANGUESADO	2.607.511	1.164.261	4.011.555	1.791.170
ENCARTACIONES	2.546.791	1.137.149	3.918.141	1.749.460
GERNIKA-BERMEO	1.497.845	668.792	2.304.378	1.028.911
GRAN BILBAO	901.308	402.437	1.386.628	619.133
MARKINA-ONDARROA	2.562.916	1.144.349	3.942.949	1.760.537
PLENTZIA-MUNGIA	445.411	198.877	685.248	305.965
ARRATIA-NERVION	3.153.164	1.407.896	4.851.021	2.165.994
DATOS POR MUNICIPIO				
ABADIÑO	171.945	76.774	264.531	118.114
ABANTO Y CIÉRVANA	12.145	5.423	18.685	8.343
AJANGIZ	32.325	14.433	49.731	22.205
ALONSOTEGI	64.186	28.659	98.748	44.091
AMOREBIETA-ETXANO	368.400	164.492	566.770	253.064
AMOROTO	170.444	76.104	262.222	117.083
ARAKALDO	27.504	12.280	42.313	18.893
ARANTZAZU	31.603	14.111	48.620	21.709
AREATZA	55.702	24.871	85.696	38.263
ARRANKUDIAGA	230.385	102.867	354.438	158.257
ARRATZU	40.054	17.884	61.622	27.514
ARRIETA	152.198	67.957	234.150	104.549
ARRIGORRIAGA	101.215	45.193	155.715	69.527
ARTEA	163.911	73.187	252.170	112.595
ARTZENTALES	181.085	80.855	278.593	124.393
ATXONDO	202.112	90.244	310.942	138.836
AULESTI	338.146	150.983	520.225	232.282
BAKIO	5.311	2.372	8.171	3.648
BALMASEDA	330.189	147.430	507.983	226.816
BARAKALDO	36.405	16.255	56.008	25.008
BARRIKA	265	118	407	182
BASAURI	341	152	525	235
BEDIA	184.732	82.483	284.203	126.898
BERANGO	578	258	890	397
BERMEO	74.740	33.372	114.985	51.341
BERRIATUA	236.949	105.799	364.538	162.767
BERRIZ	280.144	125.085	430.990	192.438
BILBAO	39.279	17.538	60.429	26.982
BUSTURIA	53.786	24.016	82.748	36.947
DERIO	16.061	7.171	24.709	11.033
DIMA	386.380	172.520	594.431	265.415

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportados. Propuesta metodológica

DURANGO	47.585	21.247	73.208	32.687
EA	95.597	42.684	147.073	65.668
ELANTXOBE	2.606	1.164	4.009	1.790
ELORRIO	467.727	208.842	719.581	321.295
ERANDIO	12.410	5.541	19.092	8.525
EREÑO	81.614	36.441	125.560	56.063
ERMUA	51.949	23.195	79.921	35.685
ERRIGOITI	150.303	67.110	231.235	103.247
ETXEBARRI	10.267	4.584	15.795	7.052
ETXEBARRIA	261.820	116.903	402.801	179.852
FORUA	27.113	12.106	41.713	18.625
FRUIZ	34.337	15.331	52.826	23.587
GALDAKAO	169.267	75.578	260.411	116.274
GALDAMES	204.082	91.123	313.972	140.189
GAMIZ-FIKA	74.004	33.043	113.852	50.835
GARAI	80.565	35.973	123.946	55.342
GATIKA	23.399	10.447	35.998	16.073
GAUTEGIZ ARTEAGA	19.127	8.540	29.426	13.139
GERNIKA-LUMO	29.303	13.084	45.081	20.129
GETXO	66	29	101	45
GIZABURUAGA	71.117	31.754	109.411	48.852
GORDEXOLA	540.223	241.211	831.112	371.094
GORLIZ	225	101	347	155
GÚÑES	280.570	125.275	431.646	192.731
IBARRANGELU	29.409	13.131	45.244	20.202
IGORRE	197.761	88.301	304.247	135.847
ISPASTER	262.227	117.085	403.426	180.131
IURRETA	132.211	59.032	203.401	90.819
IZURTZA	49.327	22.024	75.887	33.884
VALLE DE CARRANZA	262.213	117.079	403.404	180.121
KORTEZUBI	37.990	16.962	58.445	26.096
LANESTOSA	353	158	544	243
LARRABETZU	183.867	82.097	282.873	126.304
LAUKIZ	18.033	8.052	27.742	12.387
LEIOA	160	71	246	110
LEKEITIO	389	174	598	267
LEMOA	131.972	58.926	203.034	90.655
LEMOIZ	4.380	1.956	6.739	3.009
LEZAMA	54.012	24.117	83.096	37.102
LOIU	10.835	4.838	16.669	7.443
MALLABIA	256.284	114.432	394.284	176.049
MAÑARIA	51.055	22.796	78.546	35.071
MARKINA-XEMEIN	589.859	263.373	907.475	405.190
MARURI-JATABE	24.408	10.898	37.551	16.766
MENDATA	158.200	70.637	243.385	108.672
MENDEXA	32.666	14.585	50.255	22.439
MEÑAKA	46.747	20.873	71.918	32.112
MORGA	172.750	77.133	265.769	118.667
MUNDAKA	3.538	1.580	5.444	2.431
MUNGIA	59.520	26.576	91.569	40.886
MUNITIBAR-ARBATZEGI GERRIKAITZ	282.826	126.283	435.117	194.281
MURUETA	12.854	5.739	19.775	8.830
MUSKIZ	37.386	16.693	57.516	25.681
MUXIKA	342.974	153.139	527.653	235.598
NABARNIZ	130.011	58.050	200.017	89.308
ONDARROA	35.033	15.642	53.897	24.065
OROZKO	701.270	313.119	1.078.877	481.722
ORTUELLA	9.601	4.287	14.771	6.595
OTXANDIO	1.435	641	2.207	986
PLENTZIA	165	74	255	114
PORTUGALETE	0	0	0	0
SANTURTZI	1.191	532	1.832	818
SESTAO	0	0	0	0

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

SONDIKA	267	119	411	184
SOPELANA	402	179	618	276
SOPUERTA	286.206	127.792	440.318	196.603
SUKARRIETA	3.551	1.586	5.463	2.439
TRUCIOS-TURTZIOZ	95.920	42.829	147.569	65.890
UBIDE	2.016	900	3.102	1.385
UGAO-MIRABALLES	35.828	15.997	55.120	24.611
URDULIZ	2.018	901	3.105	1.386
URDUÑA-ORDUÑA	28.576	12.759	43.963	19.629
VALLE DE TRÁPAGA- TRAPAGARAN	13.609	6.076	20.936	9.348
ZALDIBAR	131.503	58.716	202.312	90.333
ZALLA	365.951	163.398	563.001	251.381
ZAMUDIO	35.945	16.050	55.301	24.692
ZARATAMO	90.154	40.254	138.698	61.929
ZEANURI	494.266	220.691	760.409	339.525
ZEBERIO	796.528	355.652	1.225.428	547.157
ZIERBENA	2.062	921	3.172	1.416
ZIORTZA-BOLIBAR	281.440	125.664	432.985	193.329

10. CONCLUSIONES

En primer lugar, creemos que hay que destacar que se haya podido confirmar la hipótesis inicial de que era posible desarrollar un modelo de predicción de la biomasa aérea para la especie *P. radiata* para Bizkaia mediante el análisis y la modelización de datos LiDAR aerotransportados y su posterior validación.

De los modelos alométricos existentes a nivel nacional y regional para la estimación de la biomasa aérea del *P. radiata* a nivel de árbol a partir de variables dasométricas como el Dn y la altura, el que mejores resultados ha proporcionado ha sido el de (Canga et al., 2013).

El modelo resultante es capaz de explicar la variabilidad de la variable dependiente en un 81%, y además permite realizar estimaciones de la misma con un margen de error de en torno al 20%. Estos resultados son coherentes con los obtenidos por otros autores aplicando metodologías similares (Anderson y Breidenbach, 2007; Dong et al., 2003; García, Riaño, Chuvieco y Danson, 2010; Gregoire et al., 2011; Næsset, 2007, 2011; Næsset y Gobakken, 2008).

El modelo MARS (Multivariate Adaptive Regression Splines), de (Friedman, 1991) ha proporcionado los mejores resultados a nivel de ajuste y de exactitud, demostrando ser una excelente alternativa a los modelos de Regresión Múltiple tradicionales. El tratamiento de las Funciones de Base como variables independientes permite además realizar la inferencia estadística necesaria para poder validar los

supuestos de los modelos de regresión clásicos y de esta manera asegurar la robustez de las predicciones, sí como de la correcta interpretación de los coeficientes al mantener bajo control la colinealidad de las variables.

El error inducido por el método de posicionamiento de las parcelas durante la fase de campo del IFN4 ha podido ser controlado y minimizado al verificarse la homogeneidad del entorno de la muestra en casi todas las parcelas, quedando aquellas que presentaban valores anómalos descartadas de la calibración final de los modelos.

Se ha desarrollado una propuesta metodológica que permite superar una de las principales barreras identificadas para la utilización de la biomasa forestal como fuente de energía: el desconocimiento de la cantidad disponible de este recurso en una masa forestal.

Se han podido establecer distintos niveles de idoneidad para la explotación de la biomasa forestal de *P. radiata* existente con vistas a su valorización energética. Esto, unido a una gestión responsable de las masas, puede resultar de gran ayuda no solo a la hora de contribuir a al impulso de las energías renovables. Otros procesos en los que podría resultar beneficiosa la movilización de este recurso podrían ser por ejemplo: el mantenimiento de las masas forestales con un menor combustible potencial de biomasa que contribuiría en la prevención de incendios, el impulso de una industria en torno a las energías limpias y renovables, además de potenciar otros sectores generadores de riqueza y empleo local, tanto en las labores

forestales como en la puesta en marcha y mantenimiento de las instalaciones energéticas que podrían sacar provecho de las mismas. Además, su explotación directa por las entidades locales en proyectos de generación de calor/frío permitirían alcanzar una mayor cota de autoabastecimiento energético y el consiguiente abaratamiento de los costes de mantenimiento de determinados equipamientos locales.

Sin embargo, no podemos olvidar las contradicciones y debates que las plantaciones forestales de especies de crecimiento rápido como el *P. radiata* o el *Eucalyptus globulus* (también con gran presencia en Bizkaia) generan en la sociedad y en el mundo académico (Ainz Ibarrondo y González Amuchastegui, 2008; Edeso, Merino, González y Marauri, 1998; Edeso, Merino, González, Ouro y Marauri, 2000), sobre todo ligadas a su manejo.

Por este motivo, este estudio, de carácter metodológico, no pretende ser un alegato a la perpetuación de las masas forestales de especies de crecimiento rápido; más bien al contrario, creemos que puede resultar una buena oportunidad para que la biomasa acumulada durante los años en los que la rentabilidad de la madera pueda tener otra alternativa y que, respetando las buenas prácticas en la gestión forestal, especialmente en el momento de la corta final (evitando las cortas a echo y minimizando la construcción de pistas forestales) y mediante el impulso de políticas públicas, se tienda a una progresiva sustitución de este tipo de masas por otras más adaptadas al territorio y que también resultarían aprovechables desde el punto de vista energético.

Estimación de la biomasa potencial disponible de *Pinus radiata* para usos energéticos en Bizkaia a partir de datos LiDAR aerotransportado. Propuesta metodológica

Referencias

- Aiken, L. S., West, S. G. y Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*: Sage.
- Ainz Ibarrondo, M. y González Amuchastegui, M. J. (2008). Gestión de montes en la reserva de la biosfera de Urdaibai: Una oportunidad perdida. (Sin especificar),
- Aldred, A. H., Bonnor, G. M. y others. (1985). *Application of airborne lasers to forest surveys* (Vol. 51).
- Andersen, H.-e., McGaughey, R. J. y Reutebuch, S. E. (2005). Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment*, 94(4), 441-449. <https://doi.org/10.1016/j.rse.2004.10.013>
- Andersen, H.-e., Strunk, J., Temesgen, H., Atwood, D. y Winterberger, K. (2014). Using multilevel remote sensing and ground data to estimate forest biomass resources in remote regions: A case study in the boreal forests of interior Alaska. *Canadian Journal of Remote Sensing*, 37(6), 596-611. <https://doi.org/10.5589/m12-003>
- Anderson, H. E. y Breidenbach, J. (2007). Statistical properties of mean stand biomass estimators in a LIDAR-based double sampling forest survey design.
- Arp, H., Griesbach, J. C. y Burns. (1982). Mapping in Tropical Forests: A New Approach Using the Laser APR. *Photogrammetric Engineering & Remote Sensing*, 48(1), 48-91.
- Aseguinolaza, C., Gómez, D., Lizaur, X., Montserrat, G., Morante, G., Salaverria y Uribe-Echebarria, P. M. (1989). Vegetación de la comunidad autónoma del País Vasco. *Gobierno Vasco. Vitoria-Gasteiz*,
- Asner, G. P., Mascaró, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., ... van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, 168(4), 1147-1160. <https://doi.org/10.1007/s00442-011-2165-z>
- Austin, M. P. y Smith, T. M. (1989). A new model for the continuum concept. *Vegetatio*, 83(1-2), 35-47. <https://doi.org/10.1007/BF00031679>
- Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, 33(B4/1; PART 4), 111-118.
- Baccini, A., Laporte, N., Goetz, S., Sun, M., Walker, W., Kellndorfer, J. y Houghton, R. A. (2009). Pan-Tropical Forest Carbon Mapped with Satellite and Field Observations. *WHRC, COP15*,
- Barron, M. (2014). LOOCV: Boston College Department of Economics.
- Baskerville, G. L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research*, 2(1), 49-53.
- Bortolot, Z. J. y Wynne, R. H. (2005). Estimating forest biomass using small footprint LiDAR data: An individual tree-based approach that incorporates training data. *ISPRS Journal of Photogrammetry and*

- Remote Sensing*, 59(6), 342-360. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0924271605000559>
- Brandtberg, T. (1999). Automatic individual tree based analysis of high spatial resolution aerial images on naturally regenerated boreal forests. *Canadian Journal of Forest Research*, 29(10), 1464-1478.
- Bravo, F., del Río, R., Montero, G., Ordoñez, C. y Cañellas, I. (2002). El diseño de las parcelas del Inventario Forestal Nacional y la estimación de variables dasométricas. *El Inventario Forestal Nacional. Elemento clave para la Gestión Forestal Sostenible Palencia*, 19-35.
- Breidenbach, J. y Kublin, E. (2009). Estimating timber volume using airborne laser scanning data based on Bayesian methods. En *Proceedings of the IUFRO Division 4 conference extending forest inventories over space and time, Quebec City, Canada* (pp. 19-22).
- Breidenbach, J., Gläser, C. y Schmidt, M. (2008). Estimation of diameter distributions by means of airborne laser scanner data. *Canadian Journal of Forest Research*, 38(6), 1611-1620.
- Breidenbach, J., Næsset, E., Lien, V., Gobakken, T. y Solberg, S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment*, 114(4), 911-924. <https://doi.org/10.1016/j.rse.2009.12.004>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140. <https://doi.org/10.1007/BF00058655>
- Brown, S. (1997). *Estimating biomass and biomass change of tropical forests: A primer* (Vol. 134): Food & Agriculture Org.
- Cabrera, M., Vera, A., Cornejo, J. M., Ordas, I., Tolosana, E., Ambrosio, Y., ... Lafarga, A. (2011). *Evaluación del potencial de energía de la biomasa. Estudio técnico PER 2011-2020*: IDAE. <http://idae.electura.es/libros/516/> [Último acceso: Noviembre 2012].
- Canga, E., Diéguez-Aranda, U., Elias, A. K. y Cámara, A. (2013). Above-ground biomass equations for *Pinus radiata* D. Don in Asturias. *Forest Systems*, 22(3), 408. <https://doi.org/10.5424/fs/2013223-04143>
- Cantero, A. (2014). El futuro del pino radiata en el País Vasco. *Sustrai: revista agropesquera*, (102), 52-55.
- Chuvieco, E. (2008). *Teledetección ambiental: La observación de la Tierra desde el espacio* (3a ed., act). Madrid: Ariel.
- Clarke, B., Fokoue, E. y Zhang, H. H. (2009). *Principles and theory for data mining and machine learning*: Springer Science & Business Media.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* Lawrence Earlbaum Associates. Hillsdale, NJ, 20-26.
- Collins, C. a., Parker, R. C. y Evans, D. L. (2004). Using multispectral imagery and multi-return LIDAR to estimate tree stand attributes in a southern bottomland hardwood forest. *Proceedings ASPRS 2004 Annual Conference, Denver, USA*, (May), 12. Recuperado de

<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:USIN+G+MULTISPECTRAL+IMAGERY+AND+MULTI-RETURN+LIDAR+TO+ESTIMATE+TREE+AND+STAND+ATTRIBUTES+IN+A+SOUTHERN+BOTTOMLAND#0>

- Condés, S. y Riaño, D. (2005). El uso del escáner láser aerotransportado para la estimación de la biomasa foliar del " *Pinus sylvestris*" L. en Canencia (Madrid). *Cuadernos de la Sociedad Española de Ciencias Forestales*, (19), 63-70.
- Cuasante, D. e. a. (2009). Estimación de recursos forestales con tecnología LiDAR aerotransportada. Aplicación práctica en varios montes de la Provincia de Burgos. En *Congresos Forestales*.
- Diéguez Aranda, U., Barrio Anta, M., Castedo Dorado, F., Ruiz González, A. D., Alvarez Taboada, M. F., Alvarez González, J. G. y Rojo Alboreca, A. (2003). Dendrometría.
- Decreto foral 23/1990, de 8 de mayo, sobre Normas Provisionales de Evaluación de Impacto Ambiental por actuaciones agropecuarias y forestales que afectan al medio natural del Territorio Histórico de Bizkaia. BOB, 118 de 5 de junio de 1990., DIPUTACIÓN FORAL DE BIZKAIA 1990.
- Decreto Foral 101/1994 de 20 de diciembre, por el que se determinan las distancias entre plantaciones forestales y fincas colindantes. BOB, 4 de 5 de enero de 1995, págs. 145-147., DIPUTACIÓN FORAL DE BIZKAIA 1994.
- Norma Foral 3/1994, de 2 de Junio de Montes y Administración de Espacios Naturales Protegidos. BOB, 123 de 28 de junio de 1994, págs. 7839-7873., DIPUTACIÓN FORAL DE BIZKAIA 1994.
- Decreto Foral 52/95 de 30 de mayo, sobre Autorización de Repoblación Forestal en montes públicos no catalogados y particulares no protectores. BOB, 115 de 16 de junio de 1995, págs. 8677-8679., DIPUTACIÓN FORAL DE BIZKAIA 1995.
- Domínguez Bravo, J. (2002). *Los sistemas de información geográfica en la planificación e integración de energías renovables. Colección Documentos CIEMAT*. Madrid: Ciemat.
- Dong, J., Kaufmann, R. K., Myneni, R. B., Tucker, C. J., Kauppi, P. E., Liski, J., ... Hughes, M. K. (2003). Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sensing of Environment*, 84(3), 393-410. [https://doi.org/10.1016/S0034-4257\(02\)00130-X](https://doi.org/10.1016/S0034-4257(02)00130-X)
- Ebdon, D. (1982). *Estadística para geógrafos*. Vilassar de Mar: Oikos-Tau.
- Edeso, J. M., Merino, A., González, M. J. y Marauri, P. (1998). Manejo de explotaciones forestales y pérdida de suelo en zonas de elevada pendiente del País Vasco. *Cuaternario y Geomorfología*, 12(1-2), 105-116.
- Edeso, J. M., Merino, A., González, M. J., Ouro, G. y Marauri, P. (2000). Pérdidas de suelo en laderas sometidas a distintos manejos forestales. *Urdaibaiko Biosfera Erreiserbari aplikatutako ikerketa. Garapen eusgarriari buruzko Urdaibaiko V. jardunaldiak*, 203-207.
- Ene, L., Næsset, E. y Gobakken, T. (2012). Single tree detection in heterogeneous boreal forests using airborne laser scanning and area-

- based stem number estimates. *International Journal of Remote Sensing*, 33(16), 5171-5193.
<https://doi.org/10.1080/01431161.2012.657363>
- Ene, L. T., Næsset, E., Gobakken, T., Gregoire, T. G., Ståhl, G. y Nelson, R. (2012). Assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach. (*Sin especificar*), 123, 579-592.
- Esteban Pascual, Luis S. (2004). *Evaluación de los recursos de biomasa forestal en la provincia de Soria: Análisis de alternativas para su aprovechamiento energético*. Colección Documentos CIEMAT. Madrid: Ciemat.
- Etzeberria Murgiondo, J. (2007). *Regresión múltiple (2a. ed.)*. Cuadernos de estadística: Vol. 4. Madrid: La Muralla.
- Fatoyinbo, T. E. y Simard, M. (2013). Height and biomass of mangroves in Africa from ICESat/GLAS and SRTM. *International Journal of Remote Sensing*, 34(2), 668-681.
- Fatoyinbo T. (2014). *Remote Sensing of Biomass: Principles and Applications (Vol. 1)*: Intech.
- Fernández, J. (2003). *Energía de a biomasa*: Haya Comunicación.
- Friedman, J. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, 19(1), 1-67. <https://doi.org/10.2307/2241837>
- García, D. (2011). *Estimación de variables de interés forestal basada en datos LiDAR en el monte numero 117 del C.U.P. Término municipal de Cuenca*.
- García, M., Riaño, D., Chuvieco, E. y Danson, F. M. (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment*, 114(4), 816-830.
<https://doi.org/10.1016/j.rse.2009.11.021>
- Gatziolis, D. y Andersen, H.-E. (2008). *A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest*. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.
- Gleason, C. J. y Im, J. (2012). Forest biomass estimation from airborne LiDAR data using machine learning approaches. *Remote Sensing of Environment*, 125, 80-91. <https://doi.org/10.1016/j.rse.2012.07.006>
- Gobakken, T. y Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38(5), 1095-1109.
<https://doi.org/10.1139/X07-219>
- Decreto 177/2014, de 16 de septiembre, por el que se aprueba definitivamente el Plan Territorial Sectorial Agroforestal de la Comunidad Autónoma del País Vasco, Gobierno Vasco 2014.
- DECRETO 139/2016, de 27 de septiembre, por el que se aprueba el Plan Rector de Uso y Gestión de la Reserva de la Biosfera de Urdaibai, Gobierno Vasco 2016.
- González, M. J. y Serrano, E. (1996). El relieve. En Ostoa (Ed.), *El clima de Euskal Herria: Vol. 3. Geografía de Euskal Herria* (pp. 30-239). Lasarte-Oria: Ostoa.

- Gonzalez-Ferreiro, E., Dieguez-Aranda, U. y Miranda, D. (2012). Estimation of stand variables in *Pinus radiata* D. Don plantations using different LiDAR pulse densities. *Forestry*, 85(2), 281-292. <https://doi.org/10.1093/forestry/cps002>
- Gregoire, T. G., Ståhl, G., Næsset, E., Gobakken, T., Nelson, R. y Holm, S. (2011). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Canadian Journal of Forest Research*, 41(1), 83-95. <https://doi.org/10.1139/X10-195>
- Gurrutxaga San Vicente, M. y IKT S.A. (2005). *Red de Corredores Ecológicos de la Comunidad Autónoma de Euskadi*.
- Hair, J. F. y Gómez Suárez, M. (2010). *Análisis multivariante* (5a. ed., última reimp). Madrid [etc.]: Prentice-Hall.
- Hall, S. A., Burke, I. C., Box, D. O., Kaufmann, M. R. y Stoker, J. M. (2005). Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, 208(1-3), 189-209. <https://doi.org/10.1016/j.foreco.2004.12.001>
- Haralick, R. M., Sternberg, S. R. y Zhuang, X. (1987). Image analysis using mathematical morphology. *IEEE transactions on pattern analysis and machine intelligence*, (4), 532-550.
- Hauglin, M., Dibdiakova, J., Gobakken, T. y Næsset, E. (2013). Estimating single-tree branch biomass of Norway spruce by airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 79, 147-156. <https://doi.org/10.1016/j.isprsjprs.2013.02.013>
- Hauglin, M., Gobakken, T., Lien, V., Bollandsås, O. M. y Næsset, E. (2012). Estimating potential logging residues in a boreal forest by airborne laser scanning. *Biomass and Bioenergy*, 36, 356-365. <https://doi.org/10.1016/j.biombioe.2011.11.004>
- Heath, L. S., Hansen, M., Smith, J. E., Miles, P. D. y Smith, B. W. (2009). Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach.
- Heinzel, J. N., Weinacker, H. y Koch, B. (2011). Prior-knowledge-based single-tree extraction. *International Journal of Remote Sensing*, 32(17), 4999-5020. <https://doi.org/10.1080/01431161.2010.494633>
- Hoerl, A. E. y Kennard, R. W. (1976). Ridge regression iterative estimation of the biasing parameter. *Communications in Statistics-Theory and Methods*, 5(1), 77-88.
- Hollaus, M., Wagner, W., Schadauer, K., Maier, B. y Gabler, K. (2009). Growing stock estimation for alpine forests in Austria: A robust lidar-based approach. *Canadian Journal of Forest Research*, 39(7), 1387-1400. <https://doi.org/10.1139/X09-042>
- Holmgren, J. y Jonsson, T. (2004). Large scale airborne laser scanning of forest resources in Sweden. *International Archives of photogrammetry, remote sensing and spatial information sciences*, 36(Part 8), W2.
- Houghton, J. T., Meira Filho, L. G., Callander, B. A., Harris, N., Kattenberg, A. y Maskell, K. Y. Ding, DJ Griggs, M. Nogueir, PJ van

- der Linden, and D. Xiaosu, Eds., 2001: *Climate Change 2001: The Scientific Basis*: Cambridge University Press.
- Houghton, J. T. (1996). *Climate change 1995: The science of climate change: contribution of working group I to the second assessment report of the Intergovernmental Panel on Climate Change* (Vol. 2): Cambridge University Press.
- Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E. y Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Earth Observations for Terrestrial Biodiversity and Ecosystems Special Issue*, 112(5), 2232-2245. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0034425708000138>
- Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X. y Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29(July 2014), 1339-1366. <https://doi.org/10.1080/01431160701736489>
- Hyyppä, J., Kelle, O., Lehtikoinen, M. y Inkinen, M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, 39(5), 969-975. <https://doi.org/10.1109/36.921414>
- Hyyppä, J., Hyyppä, H., Yu, X., Kaartinen, H., Kukko, H. y Holopainen, M. (2009). Forest inventory using small-footprint airborne lidar. *Topographic laser ranging and scanning: principles and processing*. CRC Press, Taylor & Francis London, 335-370.
- Ibáñez, J. J., Vayreda, J. y Gracia, C. (2002). Metodología complementaria al Inventario Forestal Nacional en Catalunya. _____. *El Inventario Forestal Nacional: elemento clave para la Gestión Forestal Sostenible*. Fundação General de la Universidad de Valladolid, 67-77.
- Ibáñez Martínez, E. (2008). *Validación de Modelos Digitales del Terreno de precisión a partir de datos Láser Escáner Aerotransportado; Aplicación a la Marisma del Parque Nacional de Doñana*: Universitat Politècnica de Catalunya.
- ICONA. (1990). *Segundo Inventario Forestal Nacional: Explicaciones y métodos. 1986-1995*. Madrid.
- Izenman, A. J. (2013). *Modern multivariate statistical techniques: Regression, classification, and manifold learning* (Corrected 2nd printing). *Springer texts in statistics*. New York, [London]: Springer.
- Jaccard, J. y Turrisi, R. (2003). *Interaction effects in multiple regression*: Sage.
- Jaccard, J., Wan, C. K. y Turrisi, R. (1990). The detection and interpretation of interaction effects between continuous variables in multiple regression. *Multivariate Behavioral Research*, 25(4), 467-478.
- Jarabo, F., Fernández, J., Trujillo, D., Elortegui, N. y Pérez, C. (1999). La energía de la biomasa. *Energías alternativas renovables*,

- Jenkins, J. C., Chojnacky, D. C., Heath, L. S. y Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, 49(1), 12-35.
- Jensen, J. L. R., Humes, K. S., Conner, T., Williams, C. J. y DeGroot, J. (2006). Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar. *Canadian Journal of Forest Research*, 36(5), 1129-1138. <https://doi.org/10.1139/x06-007>
- Johansen, K., Phinn, S. y Witte, C. (2010). Mapping of riparian zone attributes using discrete return LiDAR, QuickBird and SPOT-5 imagery: Assessing accuracy and costs. *Remote Sensing of Environment*, 114(11), 2679-2691. <https://doi.org/10.1016/j.rse.2010.06.004>
- Junttila, V., Maltamo, M. y Kauranne, T. (2008). Sparse Bayesian estimation of forest stand characteristics from airborne laser scanning. *Forest Science*, 54(5), 543-552.
- Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., ... others. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. *Remote Sensing*, 4(4), 950-974.
- Kankare, V., Rätty, M., Yu, X., Holopainen, M., Vastaranta, M., Kantola, T., ... Viitala, R. (2013). Single tree biomass modelling using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 85, 66-73. <https://doi.org/10.1016/j.isprsjprs.2013.08.008>
- Kankare, V., Vastaranta, M., Holopainen, M., Rätty, M., Yu, X., Hyyppä, J., ... Viitala, R. (2013). Retrieval of Forest Aboveground Biomass and Stem Volume with Airborne Scanning LiDAR. *Remote Sensing*, 5(5), 2257-2274. <https://doi.org/10.3390/rs5052257>
- Knörr, W. (2008). Ecopassenger-Environmental Methodology and Data. Final Report. *IFEU-Institut für Energie-und Umweltforschung, Heidelberg*,
- Koch, B., Dees, M., van Brusselen, J., Eriksson, L., Fransson, J., Gallaun, H., ... others. (2008). Forestry applications. *Advances in photogrammetry, remote sensing and spatial information sciences*, 439-465.
- Kollmann, F. (1959). *Tecnología de la madera y sus aplicaciones*.
- Korpela, I., Dahlin, B., Schäfer, H., Bruun, E., Haapaniemi, F., Honkasalo, J., ... others (2007). Single-tree forest inventory using lidar and aerial images for 3D treetop positioning, species recognition, height and crown width estimation. En *Proceedings of ISPRS workshop on laser scanning* (pp. 227-233).
- Kraus, K. y Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(4), 193-203.
- Lahivaara, T., Seppanen, A., Kaipio, J. P., Vauhkonen, J., Korhonen, L., Tokola, T. y Maltamo, M. (2014). Bayesian Approach to Tree Detection Based on Airborne Laser Scanning Data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(5), 2690-2699. <https://doi.org/10.1109/TGRS.2013.2264548>

- Latifi, H., Nothdurft, A. y Koch, B. (2010). Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: Application of multiple optical/LiDAR-derived predictors. *Forestry*, 83(4), 395-407. <https://doi.org/10.1093/forestry/cpq022>
- Latifi, H., Fassnacht, F. E., Hartig, F., Berger, C., Hernández, J., Corvalán, P. y Koch, B. (2015). Stratified aboveground forest biomass estimation by remote sensing data. *International Journal of Applied Earth Observation and Geoinformation*, 38, 229-241. <https://doi.org/10.1016/j.jag.2015.01.016>
- Lefsky, M. A., Cohen, W. B. y Spies, T. A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, 31(1), 78-87. <https://doi.org/10.1139/x00-142>
- Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, 37(15), n/a-n/a. <https://doi.org/10.1029/2010GL043622>
- Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., ... Oliveira, R. de. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, 32(22), n/a-n/a. <https://doi.org/10.1029/2005GL023971>
- Lim, K., Treitz, P., Wulder, M., St-Onge, B. y Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in physical geography*, 27(1), 88-106.
- Lim, K. S. y Treitz, P. M. (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research*, 19(6), 558-570. <https://doi.org/10.1080/02827580410019490>
- Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J. y Olsson, H. (2010). Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *International Journal of Remote Sensing*, 31(5), 1175-1192. <https://doi.org/10.1080/01431160903380649>
- Lomas, J. M., Urbano, C., Merino, J. M. y Camarero, L. (2001). *Valorización de la Biomasa en el País Vasco*: Bilbao: Ente Vasco de la Energía.
- Maclean, G. A. y Krabill, W. B. (1986). Gross-merchantable timber volume estimation using an airborne LIDAR system. *Canadian Journal of Remote Sensing*, 12(1), 7-18.
- Maclean, G. A. y Martin, G. L. (1984). Merchantable timber volume estimation using cross-sectional photogrammetric and densitometric methods. *Canadian Journal of Forest Research*, 14(6), 803-810.
- Magnussen, S. y Boudewyn, P. (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research*, 28(7), 1016-1031. <https://doi.org/10.1139/x98-078>
- Magnussen, S., Naesset, E. y Gobakken, T. (2013). Prediction of tree-size distributions and inventory variables from cumulants of canopy

- height distributions. *Forestry*, 86(5), 583-595.
<https://doi.org/10.1093/forestry/cpt022>
- Mäkinen, A., Korpela, I., Tokola, T. y Kangas, A. (2006). Effects of imaging conditions on crown diameter measurements from high-resolution aerial images. *Canadian Journal of Forest Research*, 36(5), 1206-1217. <https://doi.org/10.1139/x06-011>
- Maltamo, M. (2004). Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment*, 90(3), 319-330. <https://doi.org/10.1016/j.rse.2004.01.006>
- Maltamo, M., Malinen, J., Packalén, P., Suvanto, A. y Kangas, J. (2006). Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research*, 36(2), 426-436. <https://doi.org/10.1139/x05-246>
- Maltamo, M., Packalén, P., Kallio, E., Kangas, J., Uuttera, J., Heikkilä, J. y others (2011). Airborne laser scanning based stand level management inventory in Finland. En *Proceedings of SilviLaser 2011, 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems, University of Tasmania, Australia, 16-20 October 2011* (pp. 1-10).
- Maltamo, M., Packalén, P., Peuhkurinen, J., Suvanto, A., Pesonen, A. y Hyyppä, J. (2007). Experiences and possibilities of ALS based forest inventory in Finland. En *Proceedings of the ISPRS Workshop on Laser Scanning* (pp. 270-279).
- Matti Maltamo, Erik Næsset y Jari Vauhkonen. (2014). *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies* (Vol. 1): Springer.
- McGaughey, B. (2016). *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*. Seattle.
- McRoberts, R. E., Cohen, W. B., Naesset, E., Stehman, S. V. y Tomppo, E. O. (2010). Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scandinavian Journal of Forest Research*, 25(4), 340-367.
- Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L., Hendrix, C. J. y others. (2000). Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering and Remote Sensing*, 66(11), 1367-1372.
- Meaza, G. (1996). El entorno natural biótico. En Ostoa (Ed.), *El clima de Euskal Herria: Vol. 3. Geografía de Euskal Herria* (pp. 8-192). Lasarte-Oria: Ostoa.
- Ministerio de Agricultura, Alimentación y Medio Ambiente. (2013). *Cuarto Inventario Forestal Nacional - COMUNIDAD DEL PAÍS VASCO / EUSKADI*. Madrid.
- Ministerio de Medio Ambiente. (2014). *Instrucciones para el apeo de las parcelas de campo del 4º Inventario Forestal Nacional*.
- Mitchard, E. T. A., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N. L. y Brown, S. (2013). Uncertainty in the spatial distribution of tropical forest biomass: A comparison of pan-tropical maps. *Carbon balance and management*, 8(1), 10.

- Montero González, G., Ruiz-Peinado, R. y Muñoz, M. (2006). *Producción de biomasa y fijación de CO2 por los bosques españoles. Monografías INIA. Serie Forestal: n. 13.* [Madrid]: Instituto Nacional de Investigación y Técnica Agraria y Alimentaria.
- Muukkonen, P. y Heiskanen, J. (2007). Biomass estimation over a large area based on standwise forest inventory data and ASTER and MODIS satellite data: A possibility to verify carbon inventories. *Remote Sensing of Environment*, 107(4), 617-624. <https://doi.org/10.1016/j.rse.2006.10.011>
- Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(2), 49-56. [https://doi.org/10.1016/S0924-2716\(97\)83000-6](https://doi.org/10.1016/S0924-2716(97)83000-6)
- Næsset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, 61(2), 246-253.
- Næsset, E. (2002). Determination of Mean Tree Height of Forest Stands by Digital Photogrammetry. *Scandinavian Journal of Forest Research*, 17(5), 446-459. <https://doi.org/10.1080/028275802320435469>
- Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research*, 19(2), 164-179. <https://doi.org/10.1080/02827580310019257>
- Næsset, E. (2007). Airborne laser scanning as a method in operational forest inventory: Status of accuracy assessments accomplished in Scandinavia. *Scandinavian Journal of Forest Research*, 22(5), 433-442.
- Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 113(1), 148-159. <https://doi.org/10.1016/j.rse.2008.09.001>
- Næsset, E. (2011). Estimating above-ground biomass in young forests with airborne laser scanning. *International Journal of Remote Sensing*, 32(2), 473-501. <https://doi.org/10.1080/01431160903474970>
- Næsset, E. y Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, 112(6), 3079-3090. <https://doi.org/10.1016/j.rse.2008.03.004>
- NASA. (2017). *NASA Carbon Monitoring System Initiative*. Recuperado de <https://carbon.nasa.gov/>
- Nelson, R. (2013). How did we get here? An early history of forestry lidar 1. *Canadian Journal of Remote Sensing*, 39(s1), S6-S17.
- Nelson, R., Krabill, W. y MacLean, G. (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sensing of Environment*, 15(3), 201-212.
- Nelson, R., Krabill, W. y Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, 24(2), 247-267. [https://doi.org/10.1016/0034-4257\(88\)90028-4](https://doi.org/10.1016/0034-4257(88)90028-4)

- Nelson, R., Parker, G. y Hom, M. (2003). A portable airborne laser system for forest inventory. *Photogrammetric Engineering & Remote Sensing*, 69(3), 267-273.
- Nord-Larsen, T. y Riis-Nielsen, T. (2010). Developing an airborne laser scanning dominant height model from a countrywide scanning survey and national forest inventory data. *Scandinavian Journal of Forest Research*, 25(3), 262-272.
- Packalén, P. y Maltamo, M. (2006). Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. *Forest Science*, 52(6), 611-622.
- Packalén, P., Maltamo, M. y Tokola, T. (2008). Detailed assessment using remote sensing techniques. En *Designing green landscapes* (pp. 53-77). Springer.
- Palomares, O. S. y Serrano, F. S. (2000). *Mapa de la productividad potencial forestal de España: Cartografía digital*.
- Parresol, B. R. (1999). Assessing tree and stand biomass: a review with examples and critical comparisons. *Forest Science*, 45(4), 573-593.
- Peña, D. (2002). *Análisis de datos multivariantes* (Vol. 24): McGraw-Hill Madrid.
- Pérez López, C. (2009). *Técnicas de análisis de datos con SPSS 15*. Madrid [etc.]: Pearson Educación.
- Persson, A., Holmgren, J. y Söderman, U. (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, 68(9), 925-932.
- Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J. y Packalén, P. (2007). Preharvest measurement of marked stands using airborne laser scanning. *Forest Science*, 53(6), 653-661.
- Peuhkurinen, J., Mehtätalo, L. y Maltamo, M. (2011). Comparing individual tree detection and the area-based statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. *Canadian Journal of Forest Research*, 41(3), 583-598. <https://doi.org/10.1139/X10-223>
- Picard, N., Saint-André, L. y Henry, M. (2012). Manual for building tree volume and biomass allometric equations: From field measurement to prediction. (*Sin especificar*),
- Popescu, S. C., Wynne, R. H. y Nelson, R. F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. *Canadian Journal of Remote Sensing*, 29(5), 564-577.
- Popescu, S. C., Zhao, K., Neuenschwander, A. y Lin, C. (2011). Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sensing of Environment*, 115(11), 2786-2797. <https://doi.org/10.1016/j.rse.2011.01.026>
- Puig Sales, R. (D.L. 1985). *La biomasa como energía renovable: Recuperación de recursos de los residuos*. [Madrid]: Eresa.
- Raile, G. K. y Station, N. C. F. E. (1982). *Estimating stump volume*. [Saint Paul, Minn.]: U.S. Dept. of Agriculture, Forest Service, North Central Forest Experiment Station.

- Räty, M., Kankare, V., Yu, X., Holopainen, M., Vastaranta, M., Kantola, T., ... Viitala, R. (2011). Tree biomass estimation using ALS features. En *Proceedings of Silvilaser, the 11th International Conference on ALS Applications for Assessing Forest Ecosystems, Hobart, Australia* (pp. 16-20).
- Reitberger, J., Schnörr, C., Krzystek, P. y Stilla, U. (2009). 3D segmentation of single trees exploiting full waveform LIDAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), 561-574. <https://doi.org/10.1016/j.isprsjprs.2009.04.002>
- Rempel RC, P. A. K. An information note on an airborne laser terrain profiler for micro-relief studies. En *Proceedings of the 3rd symposium on remote sensing of environment* (pp. 321-337).
- Reutebuch, S. E., Schreuder, G. F. y Andersen, H.-e. (2001). Automated individual tree measurement through morphological analysis of a LIDAR-based canopy surface model. *Proc. International Precision Forestry Symposium, Seattle, WA, USA*, 11-21.
- Richardson, J., Björheden, R., Hakkila, P., Lowe, A. T. y Smith, C. T. (2006). *Bioenergy from sustainable forestry: Guiding principles and practice* (Vol. 71): Springer Science & Business Media.
- Rosette, J. A. B., North, P. R. J. y Suárez, J. C. (2008). Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. *International Journal of Remote Sensing*, 29(5), 1475-1493. <https://doi.org/10.1080/01431160701736380>
- Rosette, J., Cook, B., Suárez, J., North, P., Nelson, R. y Los, S. (2012). *Lidar remote sensing for biomass assessment*: INTECH Open Access Publisher.
- Ruiz Urrestarazu, E. (1982). *La transición climática del Cantábrico Oriental al valle medio del Ebro*. VITORIA-GASTEIZ: Diputación Foral de Alava.
- Ruiz Urrestarazu, E. y Galdos Urrutia, R. (1989). El clima de la vertiente atlántica del País Vasco. AA. VV.: *lbaiak eta Haranak: el agua, el río y los espacios agrícola, industrial y urbano*. Ed. Etor, Donostia-San Sebastián, 45-72.
- Ruiz Urrestarazu, E. y Galdos Urrutia, R. (2008). *Geografía del País Vasco*: Editorial NEREA.
- Salas, C., Ene, L., Gregoire, T. G., Næsset, E. y Gobakken, T. (2010). Modelling tree diameter from airborne laser scanning derived variables: A comparison of spatial statistical models. *Remote Sensing of Environment*, 114(6), 1277-1285. <https://doi.org/10.1016/j.rse.2010.01.020>
- Sanz, F. y Piñeiro, G. (2003). Aprovechamiento de la biomasa forestal producida por la cadena monte-industria. Parte I: Situación actual y evaluación de sistemas de tratamiento. *Revista CIS-Madeira*, 10, 6-37.
- Schreier, H., Lougheed, J., Tucker, C. y Leckie, D. (1985). Automated measurements of terrain reflection and height variations using an airborne infrared laser system. *International Journal of Remote Sensing*, 6(1), 101-113. <https://doi.org/10.1080/01431168508948427>

- Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D. y DiMarzio, J. P. (2005). Overview of the ICESat mission. *Geophysical Research Letters*, 32(21).
- Sithole, G. y Vosselman, G. (2004). Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(1-2), 85-101. <https://doi.org/10.1016/j.isprsjprs.2004.05.004>
- Solomon, N., Oatley, G. y McGarry, K. (2007). A dynamic method for the evaluation and comparison of imputation techniques.
- Somogyi, Z., Teobaldelli, M., Federici, S., Matteucci, G., Pagliari, V., Grassi, G. y Seufert, G. (2008). Allometric biomass and carbon factors database. *iForest-Biogeosciences and Forestry*, 1(3), 107-113.
- Sorin C Popescu. (2007). Estimating biomass of individual pine trees using airborne lidar, 31, 646-655. <https://doi.org/10.1016/j.biombioe.2007.06.022>
- Splithoff, H. (2010). *Power generation from solid fuels*: Springer Science & Business Media.
- Sprugel, D. G. (1983). Correcting for bias in log-transformed allometric equations. *Ecology*, 64(1), 209-210.
- Strunk, J., Temesgen, H., Andersen, H.-e., Flewelling, J. P. y Madsen, L. (2014). Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing*, 38(5), 644-654. <https://doi.org/10.5589/m12-052>
- Suárez, J. C., Ontiveros, C., Smith, S. y Snape, S. (2005). Use of airborne LiDAR and aerial photography in the estimation of individual tree heights in forestry. *Computers & Geosciences*, 31(2), 253-262. <https://doi.org/10.1016/j.cageo.2004.09.015>
- SUN, G., RANSON, K., Kimes, D., BLAIR, J. y KOVACS, K. (2008). Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sensing of Environment*, 112(1), 107-117. <https://doi.org/10.1016/j.rse.2006.09.036>
- Swatantran, A., Dubayah, R., Roberts, D., Hofton, M. y Blair, J. B. (2011). Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. (*Sin especificar*), 115(11), 2917-2930.
- Tang, S., Dong, P. y Buckles, B. P. (2013). Three-dimensional surface reconstruction of tree canopy from lidar point clouds using a region-based level set method. *International Journal of Remote Sensing*, 34(4), 1373-1385. <https://doi.org/10.1080/01431161.2012.720046>
- Thomas, V., Treitz, P., McCaughey, J. H. y Morrison, I. (2006). Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: An examination of scanning density. *Canadian Journal of Forest Research*, 36(1), 34-47. <https://doi.org/10.1139/x05-230>
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288.

- Torre, L. (2016). *Diseño y contraste de nuevos modelos de estimación del potencial energético de biomasa forestal en el Territorio de Bizkaia mediante técnicas de análisis estadístico espacial usando herramientas GIS con datos LiDAR* (Tesis). Universidad de Cantabria, Cantabria.
- Torre Tojal, L. (2016). *Diseño y contraste de nuevos modelos de estimación del potencial energético de biomasa forestal en el Territorio de Bizkaia mediante técnicas de análisis estadístico espacial usando herramientas GIS con datos LiDAR*. Universidad de Cantabria.
- Tracsatec. *El 3er Inventario Forestal Nacional. Descripción de los códigos de la base de descripción de los códigos de la base de datos de SIG*.
- Uriarte, A. (1996). El Clima de Euskal Herria. En Ostoa (Ed.), *El clima de Euskal Herria: Vol. 3. Geografía de Euskal Herria*. Lasarte-Oria: Ostoa.
- Usos del suelo. Banco de datos del Mapa Forestal. (2006).
- Valbuena-Rabadán, M. A. (2014). *Determinación de variables forestales de masa y de árboles individuales mediante delineación de copas a partir de datos LIDAR Aerotransportado. Aplicación a las masas de Pinus sylvestris L. en Álava* (Tesis). Universidad de La Rioja, Logroño.
- Valbuena-Rabadán, M.-Á., Santamaría-Peña, J. y Sanz-Adán, F. (2016). Estimation of diameter and height of individual trees for *Pinus sylvestris* L. based on the individualising of crowns using airborne LiDAR and the National Forestry Inventory data. *Forest Systems*, 25(1), 46. <https://doi.org/10.5424/fs/2016251-05790>
- Valentine, H. T., Tritton, L. M. y Furnival, G. M. (1984). Subsampling trees for biomass, volume, or mineral content. *Forest Science*, 30(3), 673-681.
- Vauhkonen, J., Mehtätalo, L. y Packalén, P. (2011). Combining tree height samples produced by airborne laser scanning and stand management records to estimate plot volume in Eucalyptus plantations. *Canadian Journal of Forest Research*, 41(8), 1649-1658. <https://doi.org/10.1139/x11-083>
- VERICAT GRAU, P. y PIQUÉ NICOLAU, M. (Eds.) 2013. *Factores ambientales que afectan al crecimiento de las cuatro principales especies de coníferas en Cataluña*.
- Vosselman, G. (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, 33(B3/2; PART 3), 935-942.
- Woods, M., Pitt, D., Penner, M., Lim, K., Nesbitt, D., Etheridge, D. y Treitz, P. (2011). Operational implementation of a LiDAR inventory in Boreal Ontario. *The Forestry Chronicle*, 87(4), 512-528.
- Wulder, M. A., White, J. C., Alvarez, F., Han, T., Rogan, J. y Hawkes, B. (2009). Characterizing boreal forest wildfire with multi-temporal Landsat and LIDAR data. *Remote Sensing of Environment*, 113(7), 1540-1555. <https://doi.org/10.1016/j.rse.2009.03.004>

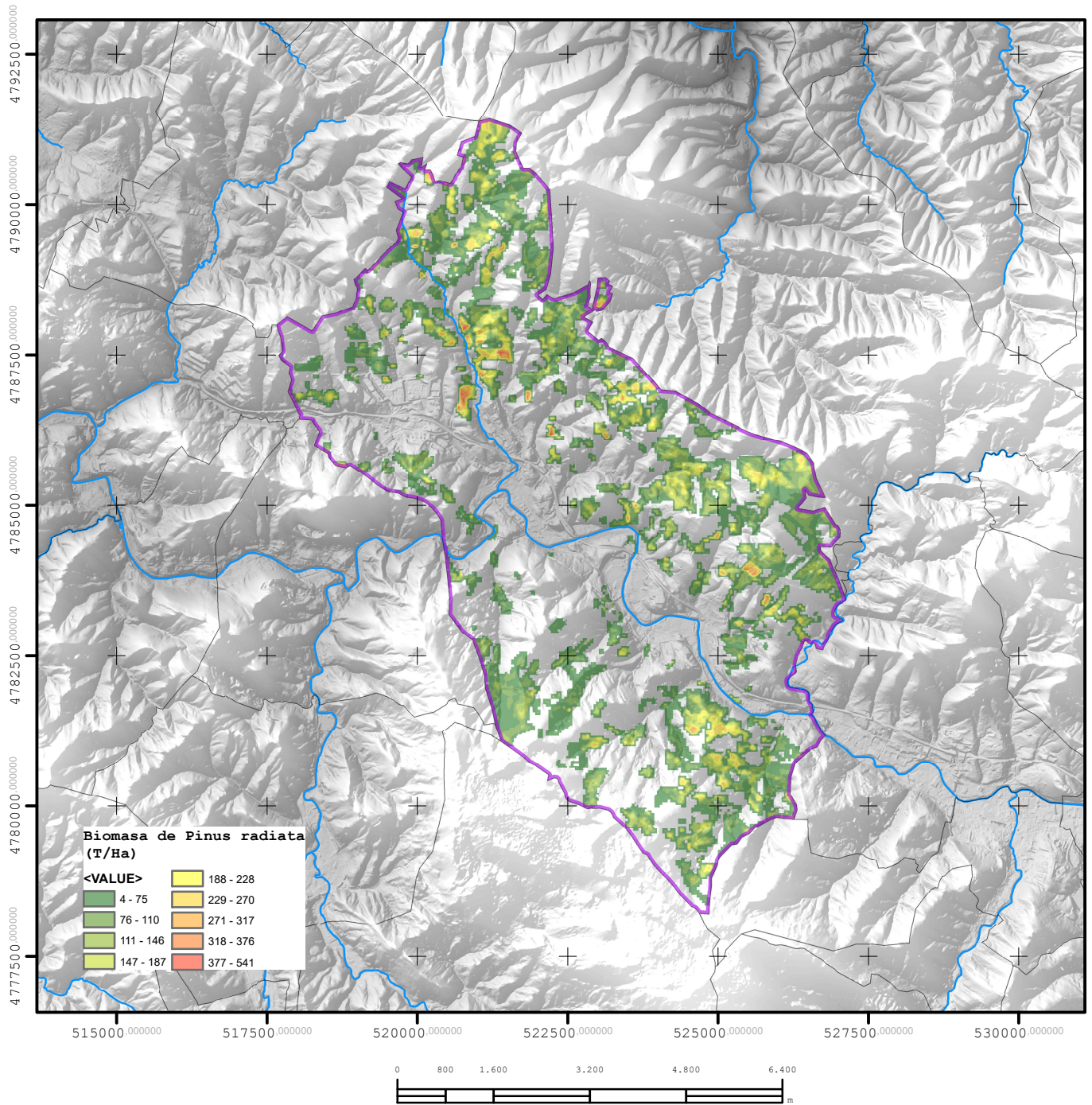
- Yee, T. W. y Mitchell, N. D. (1991). Generalized additive models in plant ecology. *Journal of Vegetation Science*, 2(5), 587-602. <https://doi.org/10.2307/3236170>
- Zavitkovski, J. (1976). Ground Vegetation Biomass, Production, and Efficiency of Energy Utilization in Some Northern Wisconsin Forest Ecosystems. *Ecology*, 57(4), 694-706. <https://doi.org/10.2307/1936183>
- Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J. y Zhang, C. (2003). A progressive morphological filter for removing nonground measurements from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(4), 872-882. <https://doi.org/10.1109/TGRS.2003.810682>
- Zhao, F., Guo, Q. y Kelly, M. (2012). Allometric equation choice impacts lidar-based forest biomass estimates: A case study from the Sierra National Forest, CA. *Agricultural and Forest Meteorology*, 165, 64-72. <https://doi.org/10.1016/j.agrformet.2012.05.019>
- Zhao, K., Popescu, S., Meng, X., Pang, Y. y Agca, M. (2011). Characterizing forest canopy structure with lidar composite metrics and machine learning. *Remote Sensing of Environment*, 115(8), 1978-1996. <https://doi.org/10.1016/j.rse.2011.04.001>
- Zhao, K., Popescu, S. y Nelson, R. (2009a). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment*, 113(1), 182-196. <https://doi.org/10.1016/j.rse.2008.09.009>
- Zhao, K., Popescu, S. y Nelson, R. (2009b). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. *Remote Sensing of Environment*, 113(1), 182. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0034425708002770>
- Zianis, D., Muukkonen, P., Mäkipää, R., Mencuccini, M. y others. (2005). *Biomass and stem volume equations for tree species in Europe*: FI.

ANEXO CARTOGRÁFICO

- 1. MUNICIPIOS⁸**
 - a. AMOREBIETA-ETXANO
 - b. AMOROTO
 - c. BALMASEDA
 - d. BERRIATUA
 - e. ELORRIO
 - f. ETXEBARRIA
 - g. BERMEO
 - h. ZIORTZA-BOLIBAR
- 2. BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL**
 - a. ARRATIA-NERVIÓN
 - b. ARATIA-NERVIÓN (ORDUÑA)
 - c. DURANGUESADO
 - d. ENCARTACIONES
 - e. GUERNIKA-BERMEO
 - f. GRAN BILBAO
 - g. MARKINA-ONDARROA
 - h. PLENTZIA-MUNGUÍA
- 3. MAPAS DE INCERTIDUMBRE**

⁸ Se adjuntan a modo de ejemplo 8 mapas correspondientes a municipios. La totalidad de los mapas se puede consultar en el CD adjunto "ANEXO CARTOGRÁFICO".

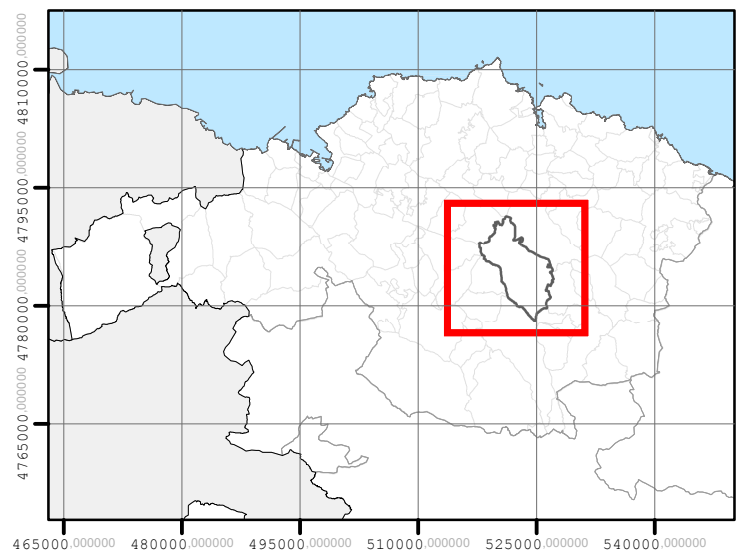
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



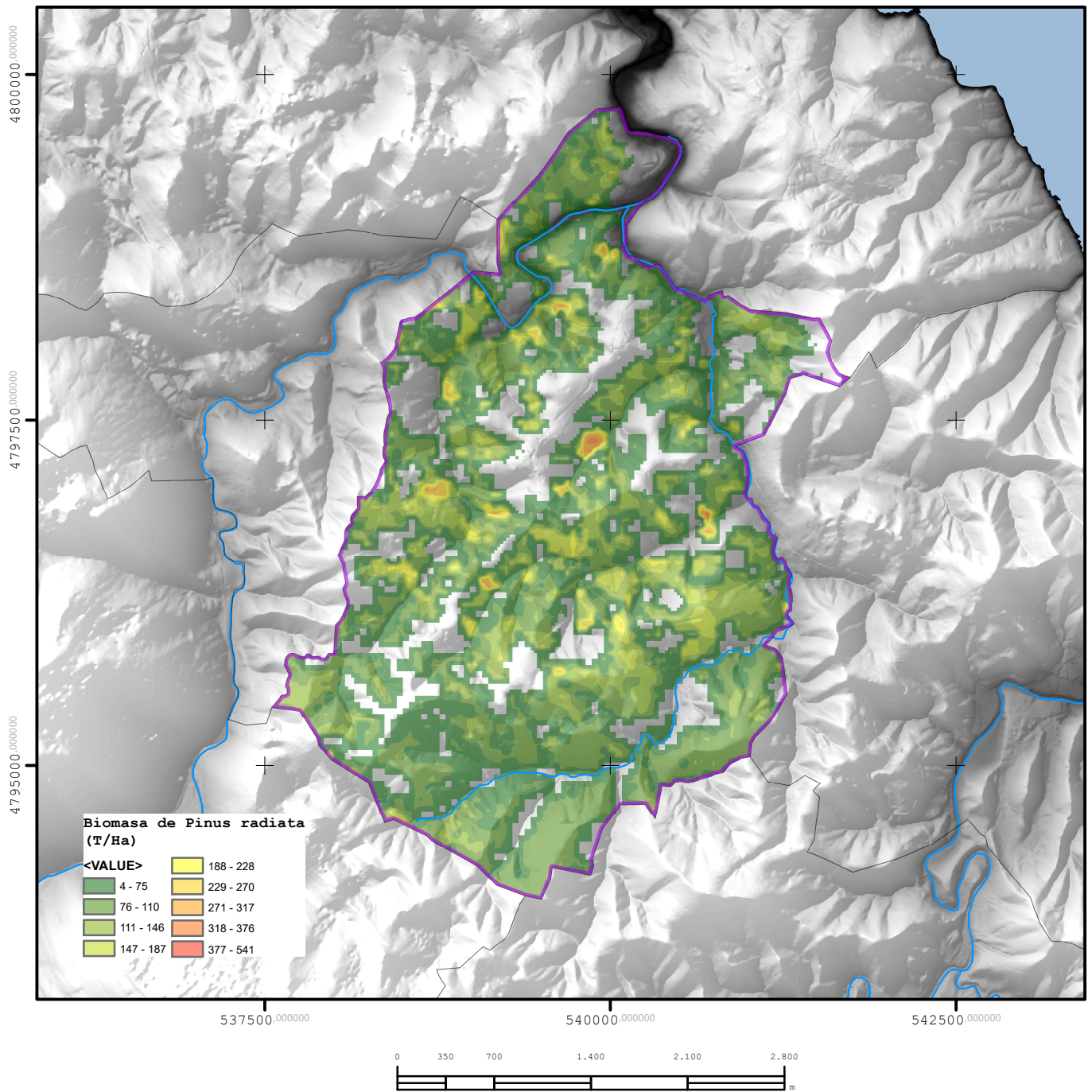
AMOREBIETA-ETXANO

Superficie *Pinus radiata* 5879,85 ha
 Ocupación 25,5%
 Biomasa potencial total 174.563 t/ha

Tamaño del pixel = 0,196 ha



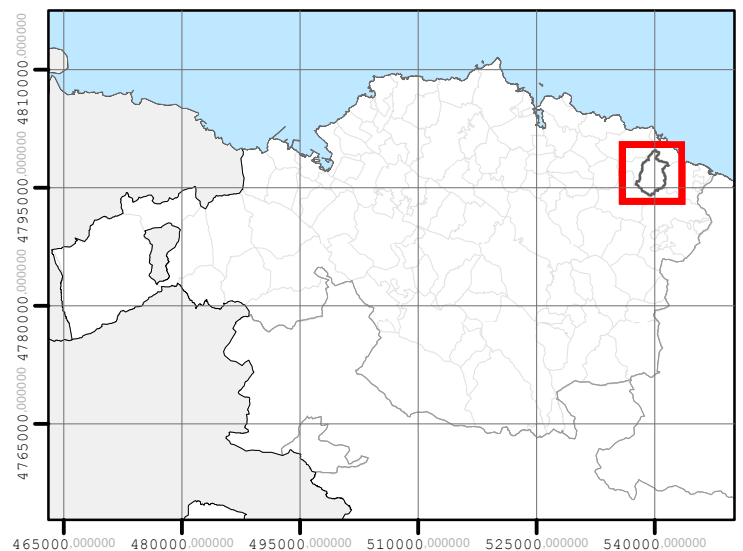
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



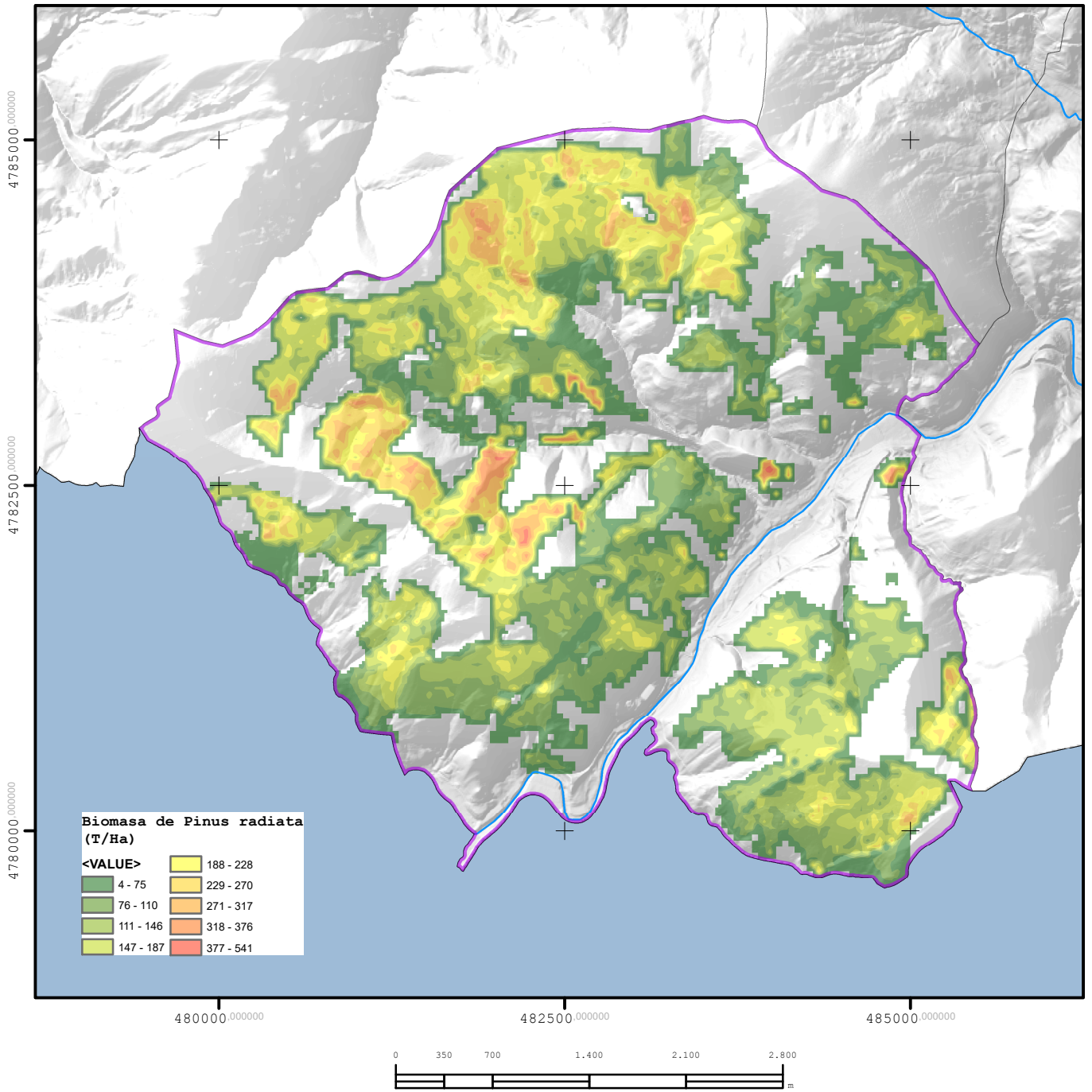
AMOROTO

Superficie *Pinus radiata* 1305,9 ha
 Ocupación 55,9%
 Biomasa potencial total 78.961,2 t/ha

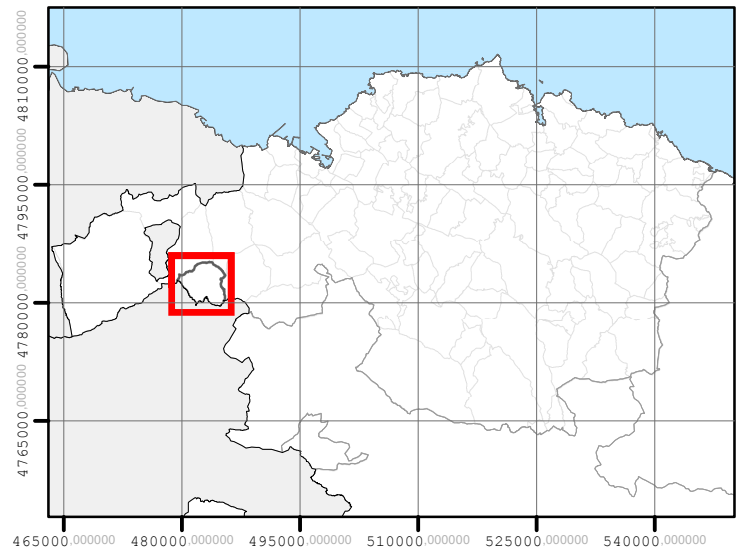
Tamaño del pixel = 0,196 ha



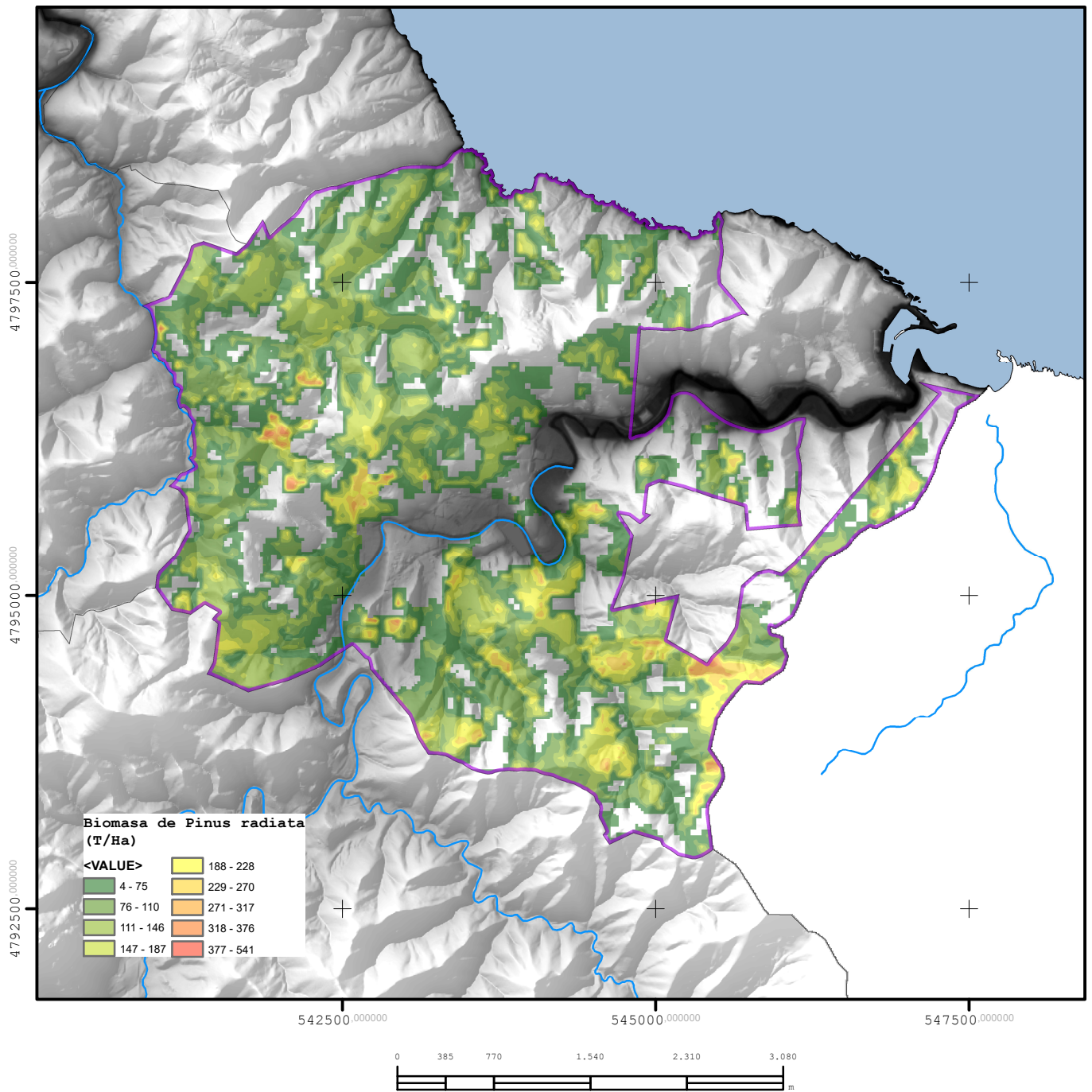
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



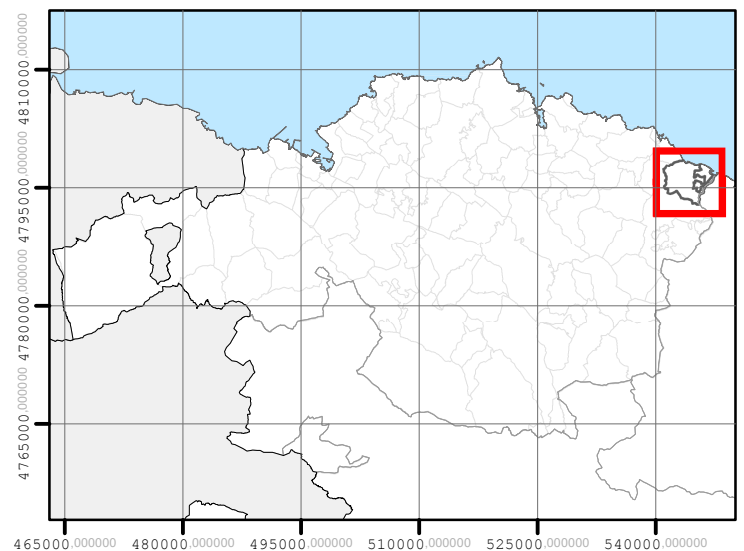
BALMASEDA	
Superficie <i>Pinus radiata</i>	2235,45 ha
Ocupación	47,9%
Biomasa potencial total	163.123 t/ha
Tamaño del pixel = 0,196 ha	



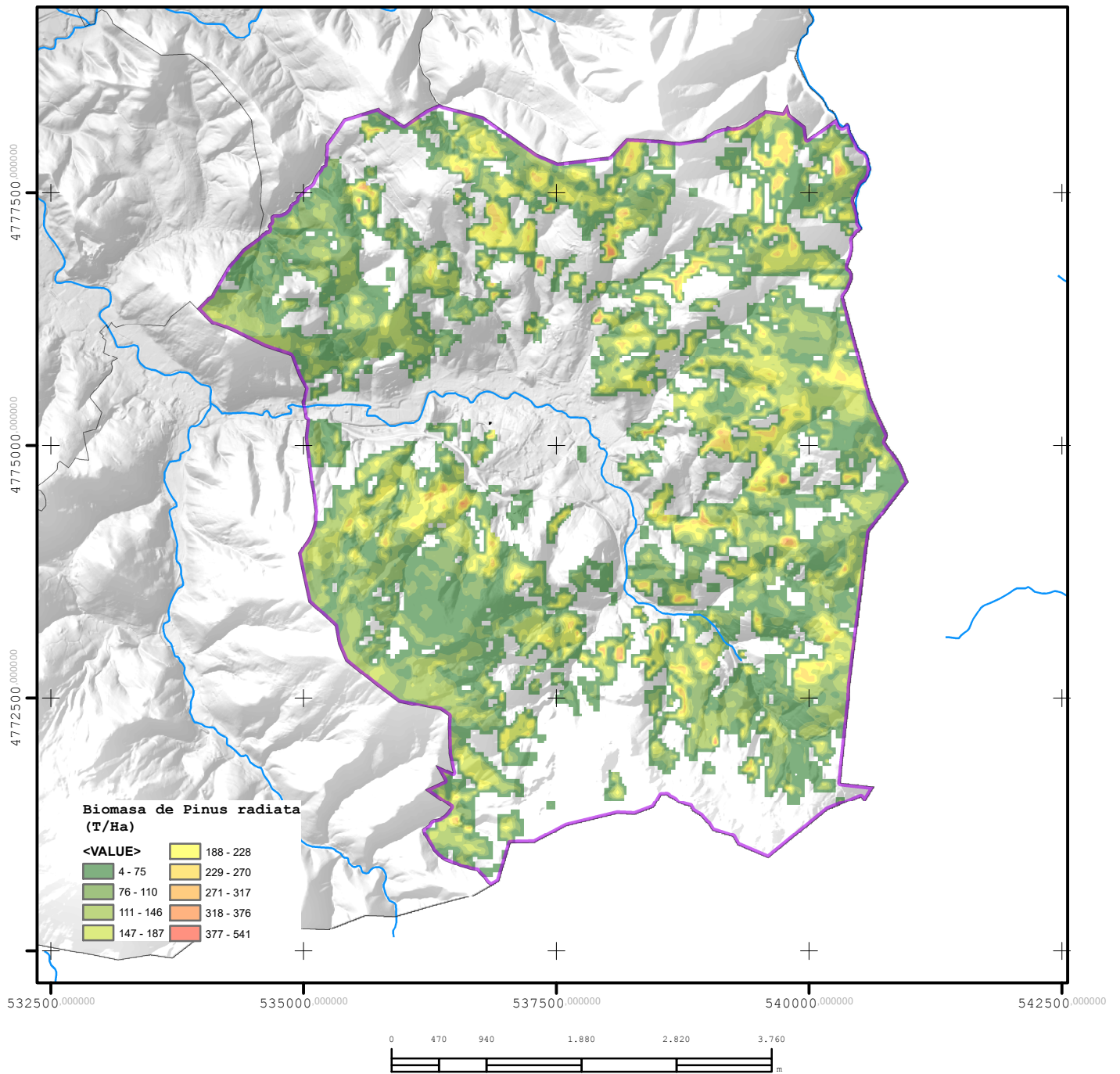
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



BERRIATUA	
Superficie <i>Pinus radiata</i>	1907,58 ha
Ocupación	48,2%
Biomasa potencial total	115.652 t/ha
Tamaño del pixel = 0,196 ha	



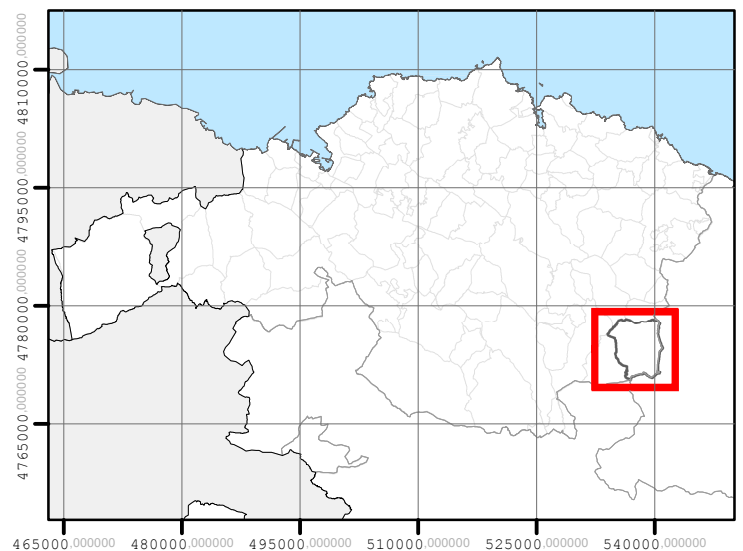
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



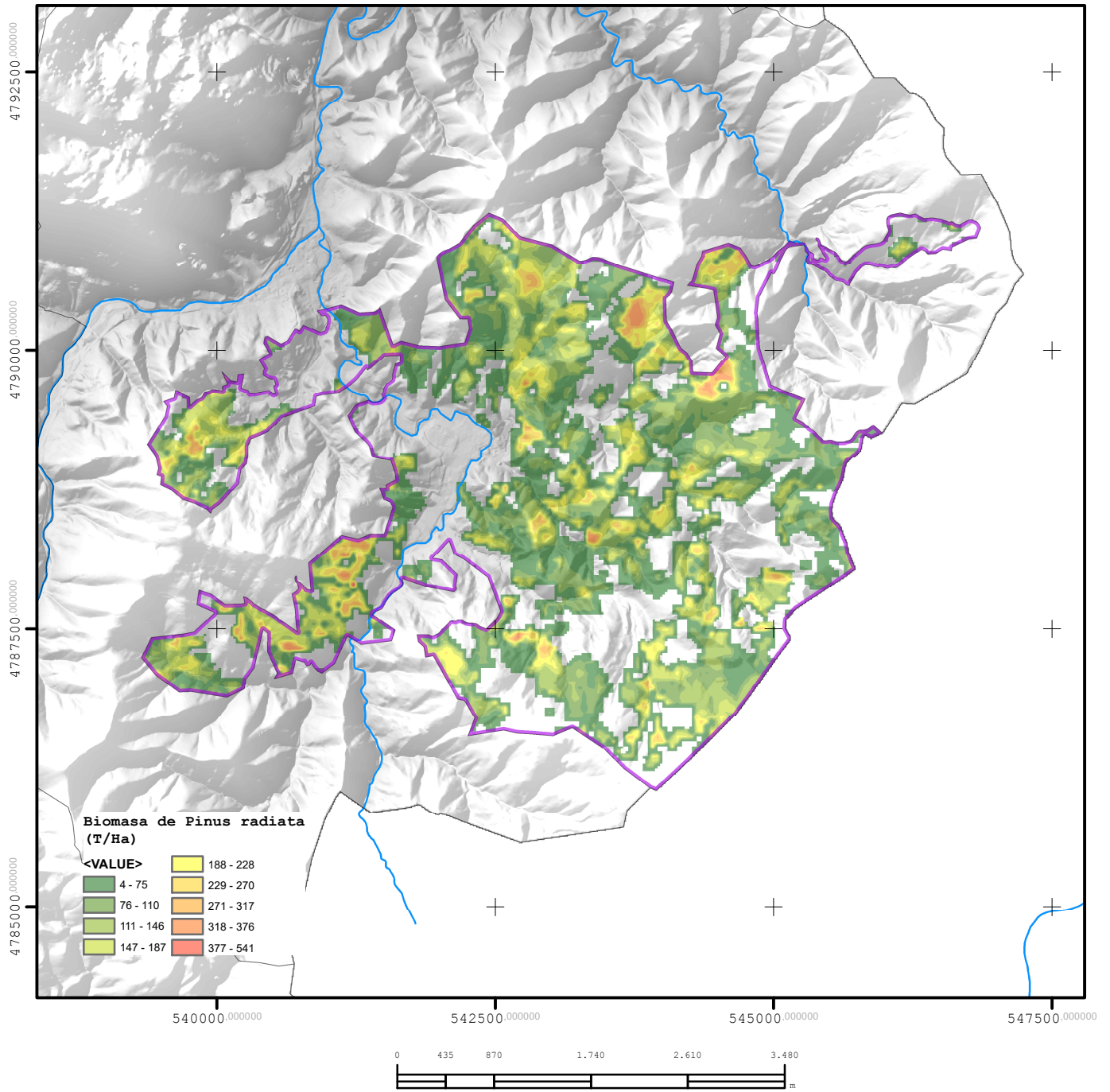
ELORRIO

Superficie *Pinus radiata* 3694,32 ha
 Ocupación 45,1%
 Biomasa potencial total 210.894 t/ha

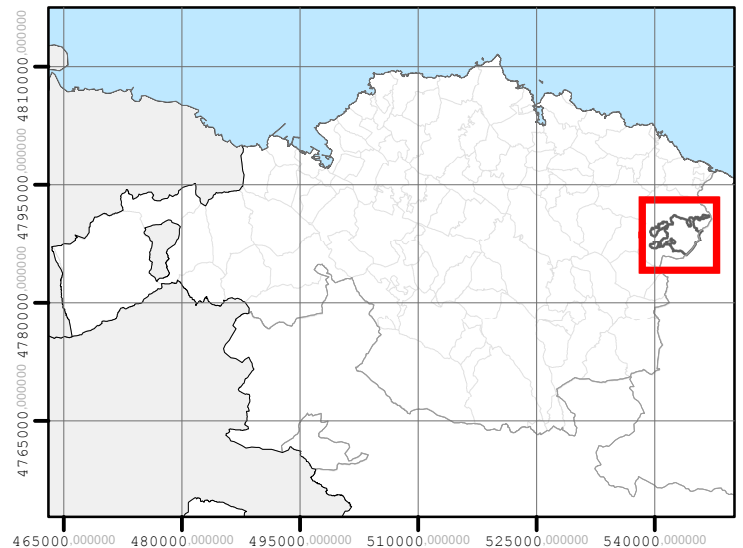
Tamaño del pixel = 0,196 ha



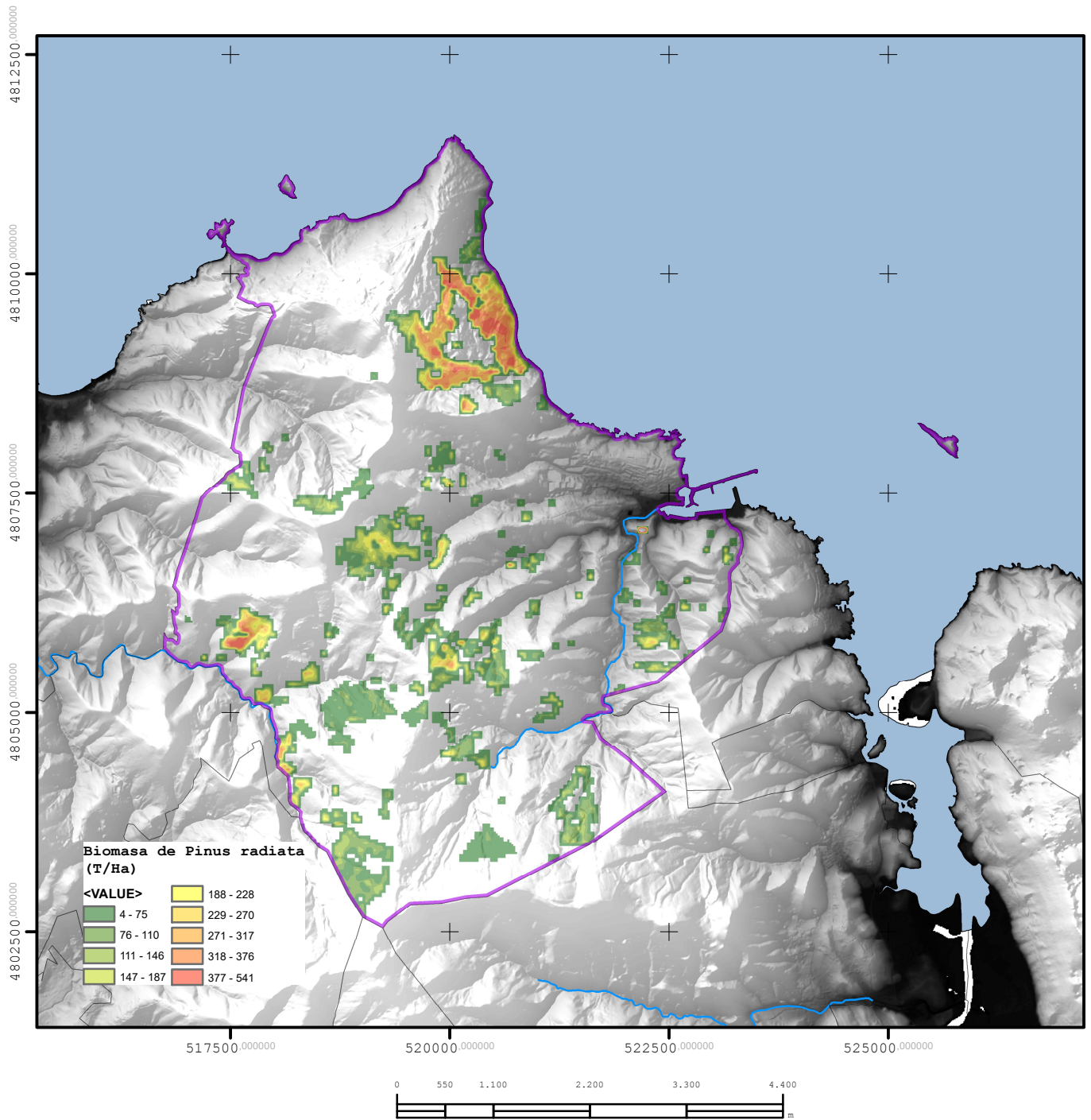
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



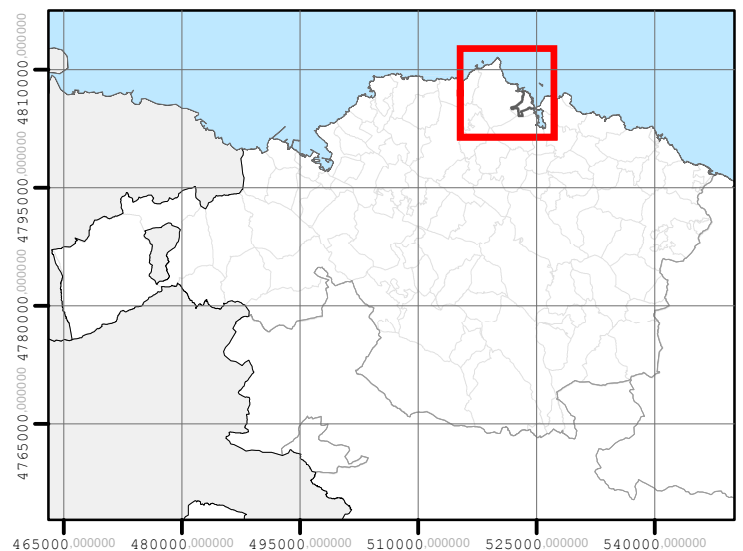
ETXEBARRIA	
Superficie <i>Pinus radiata</i>	1799,05 ha
Ocupación	49,8%
Biomasa potencial total	117.874 t/ha
Tamaño del pixel = 0,196 ha	



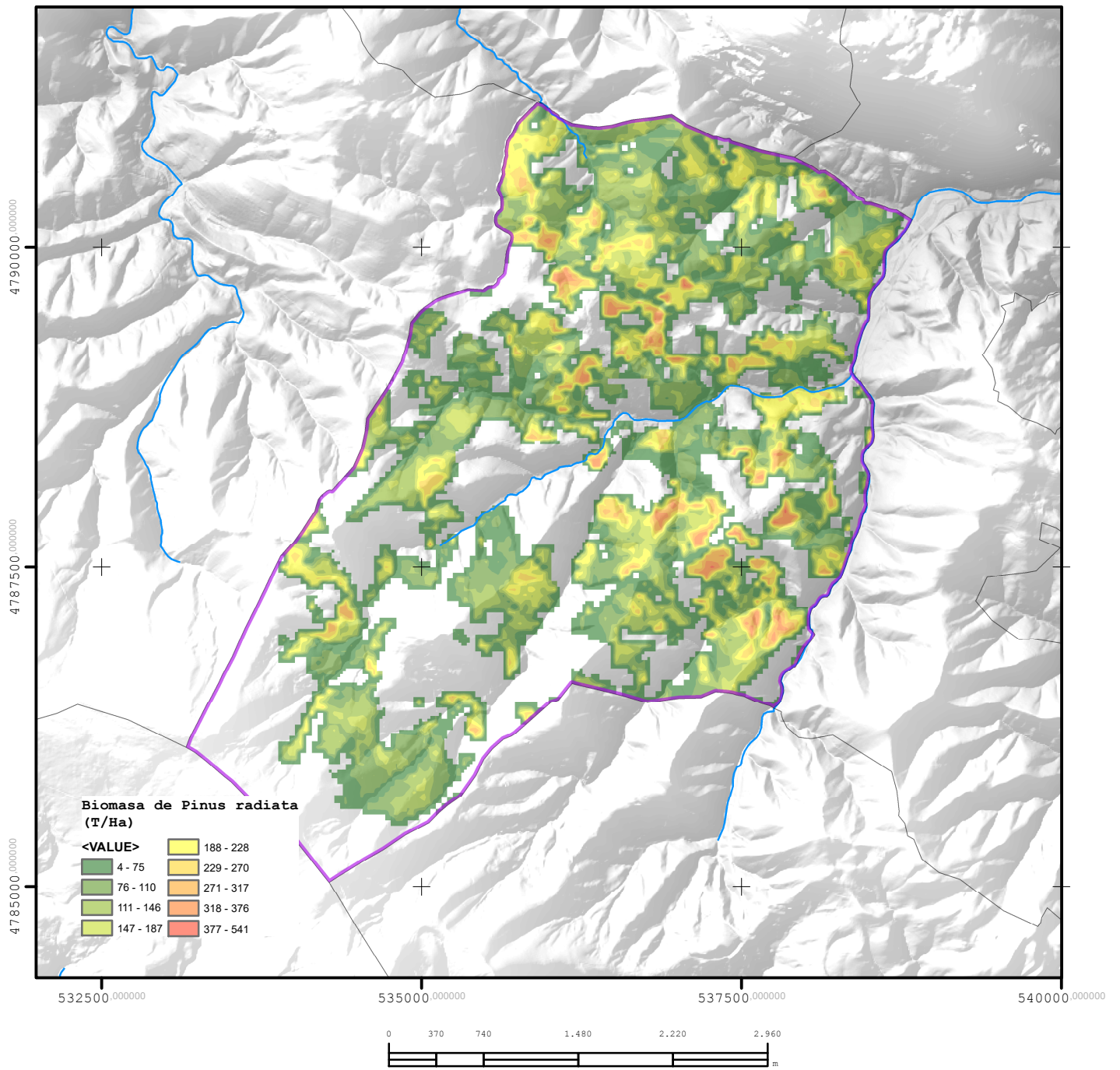
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



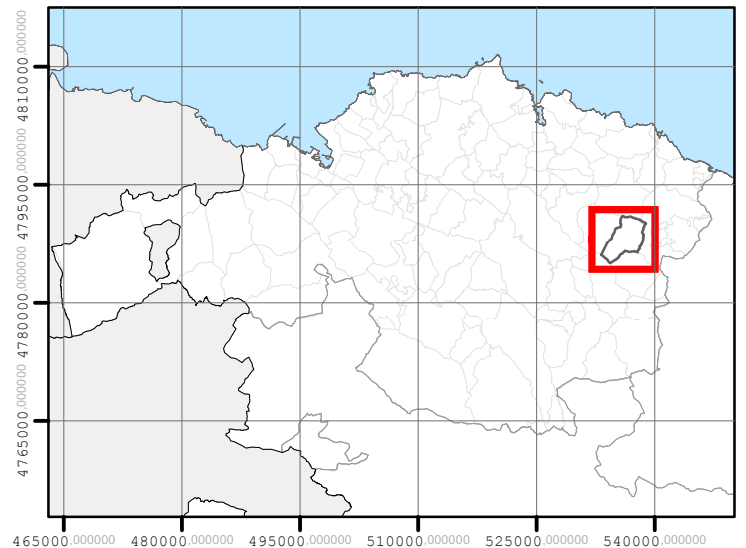
BERMEO	
Superficie <i>Pinus radiata</i>	3367,86 ha
Ocupación	12,2%
Biomasa potencial total	67.820,9 t/ha
Tamaño del pixel = 0,196 ha	



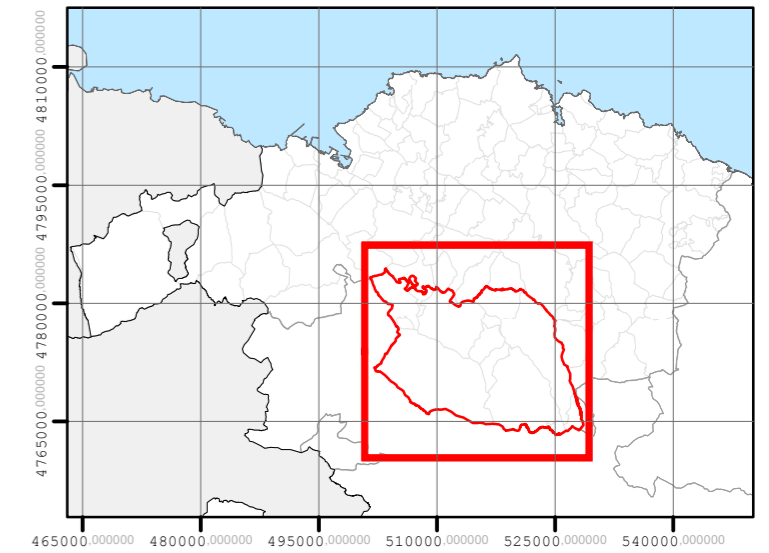
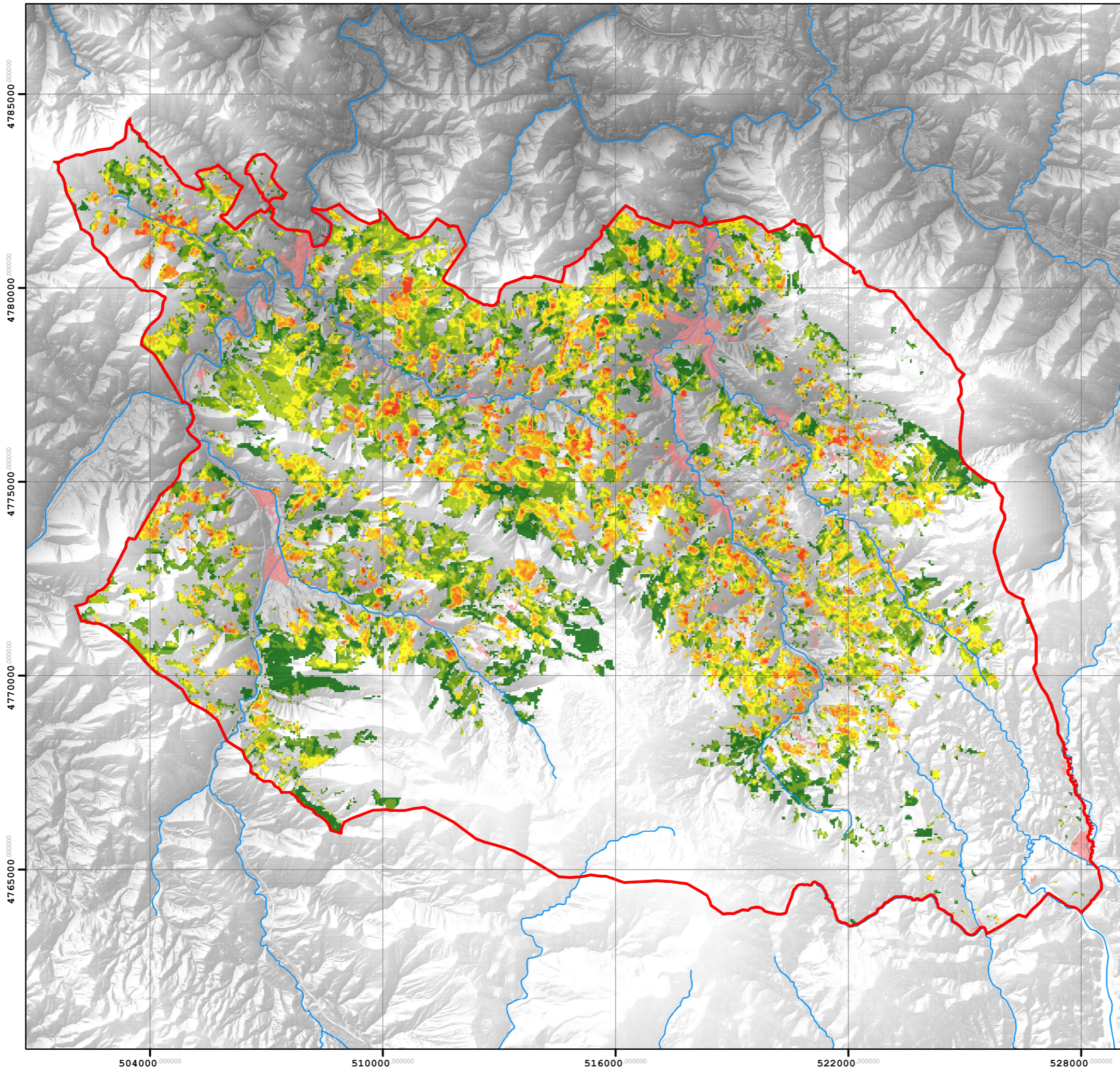
BIOMASA POTENCIAL DE *Pinus radiata* LATIZAL/FUSTAL (t/ha)



ZIORTZA-BOLIBAR	
Superficie <i>Pinus radiata</i>	1884,04 ha
Ocupación	45,2%
Biomasa potencial total	125.480 t/ha
Tamaño del pixel = 0,196 ha	



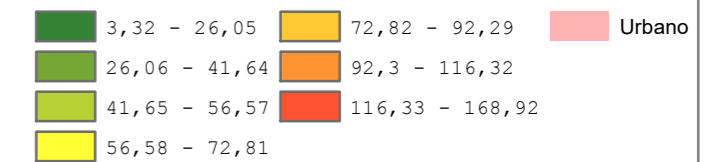
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



ARRATIA-NERVION

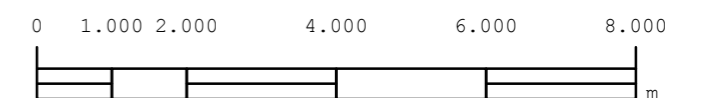
Superficie *Pinus radiata* 11.899,8 ha
 Porcentaje de ocupación 32,9%
 Biomasa potencial total 1.477.305 t/ha

Biomasa de *P. radiata* (t)

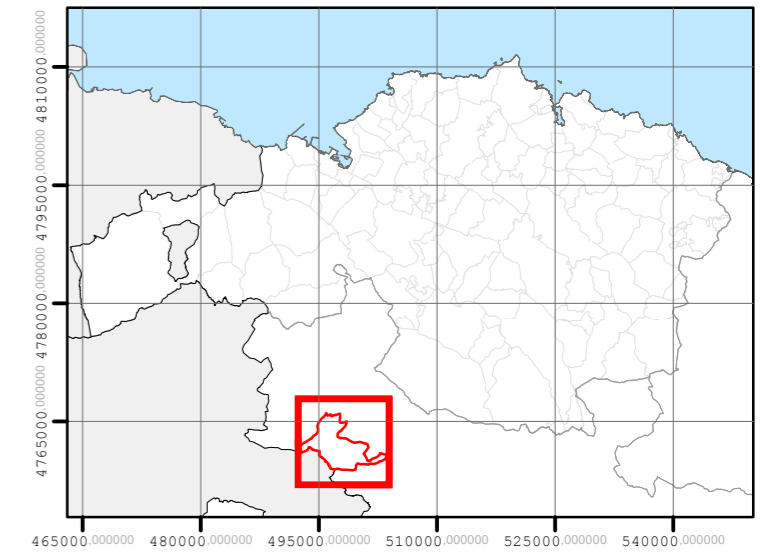
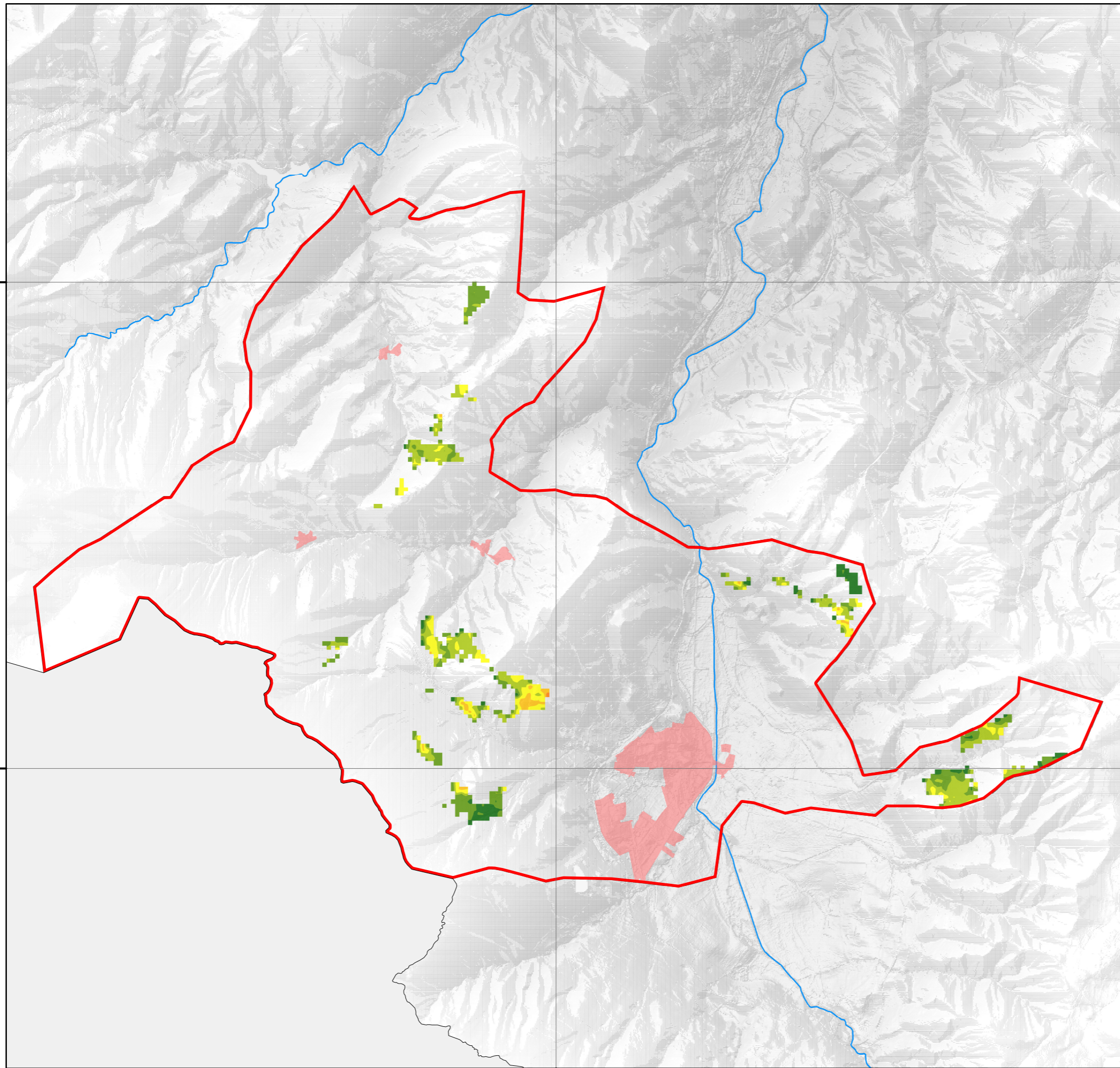


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



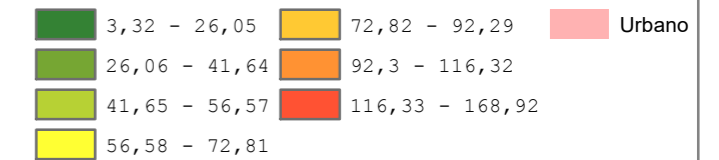
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



ARRATIA-NERVION (ORDUÑA)

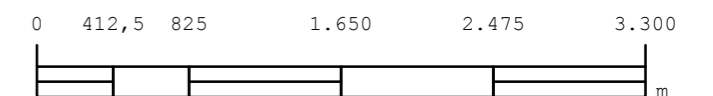
Superficie *Pinus radiata* 130,7 ha
 Porcentaje de ocupación 3,9%
 Biomasa potencial total 13.308,22 t/ha

Biomasa de *P. radiata* (t)

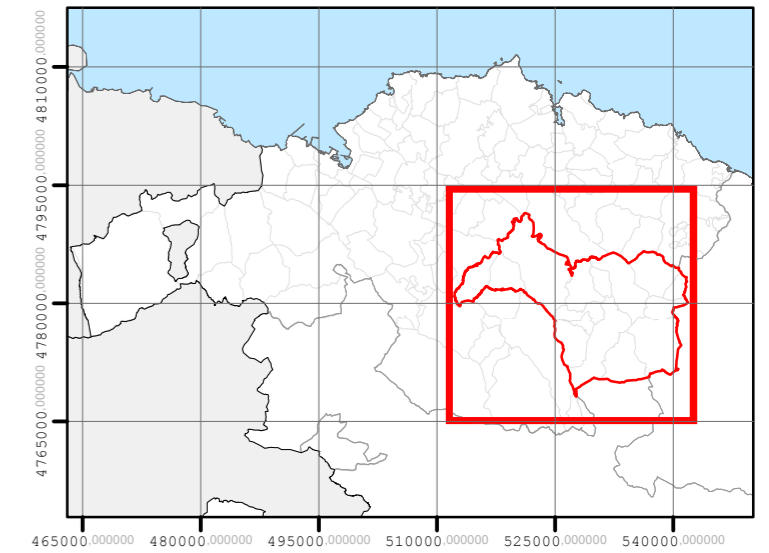
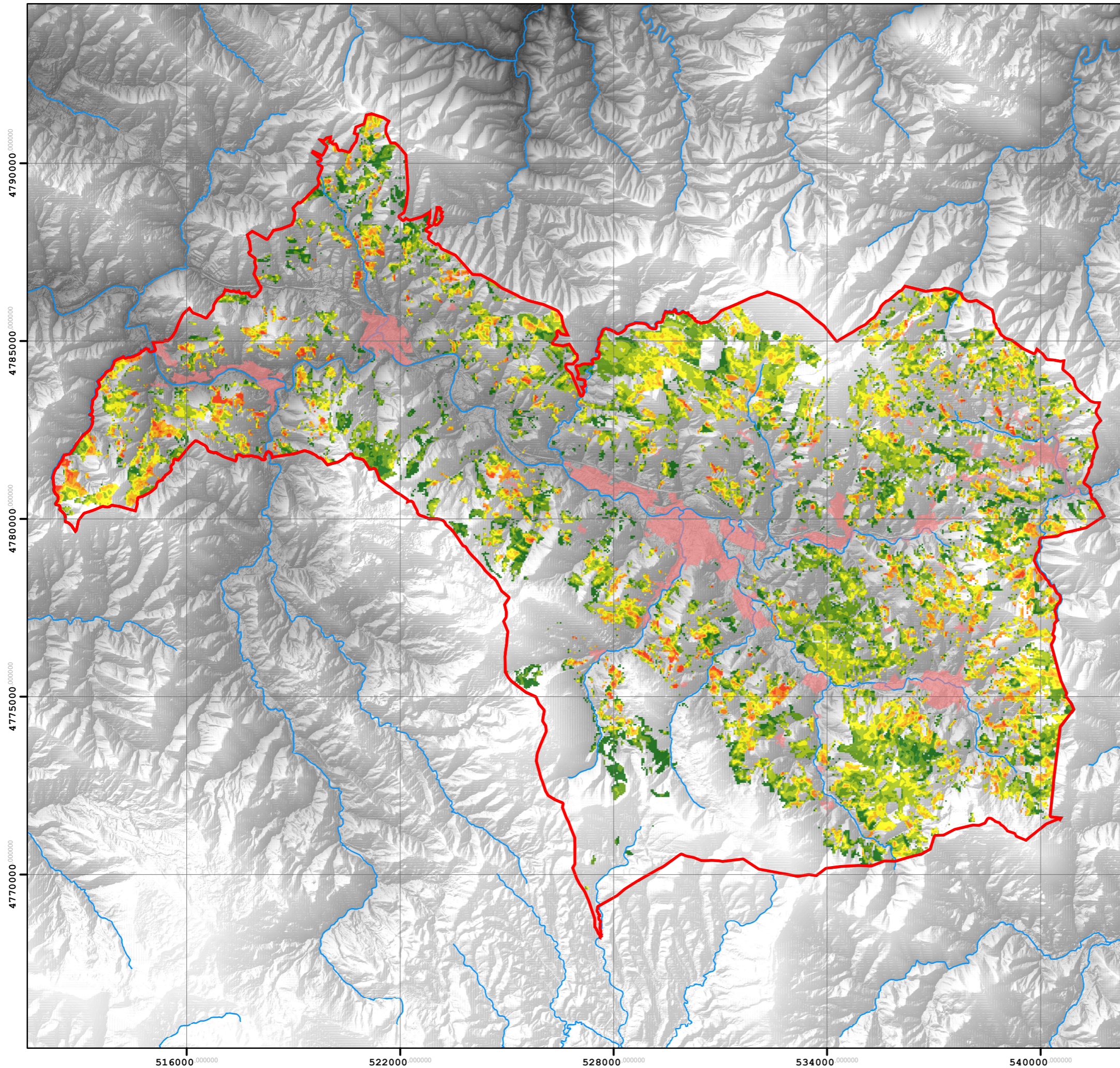


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



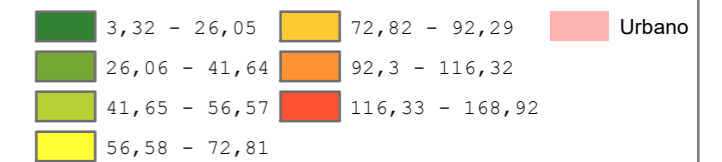
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



DURANGUESADO

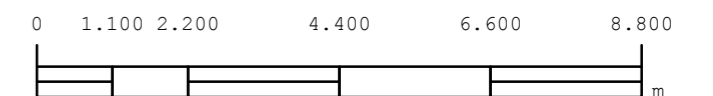
Superficie *Pinus radiata* 9.889,7 ha
 Porcentaje de ocupación 31,3%
 Biomasa potencial total 1.224.292 t/ha

Biomasa de *P. radiata* (t)

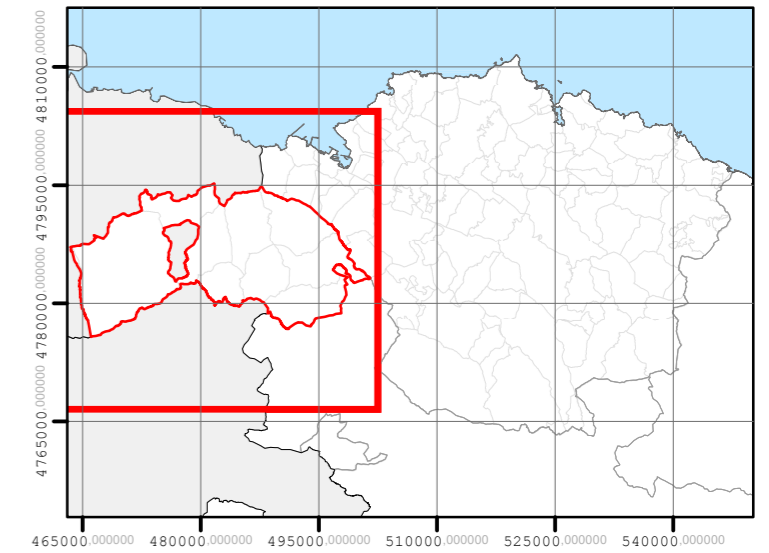
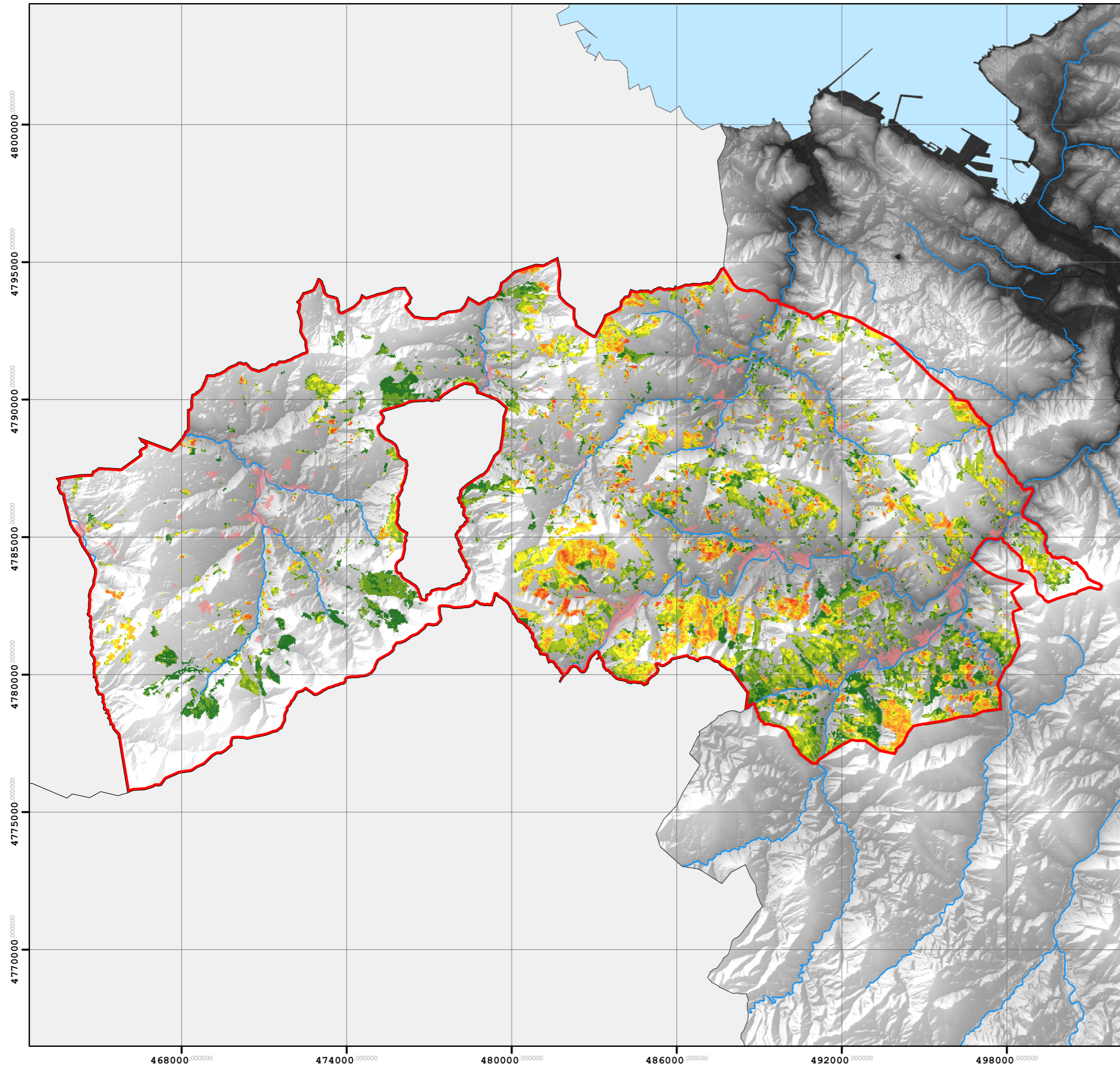


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000.0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



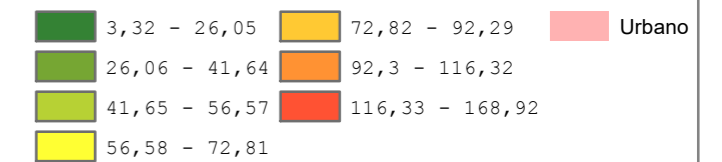
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



ENCARTACIONES

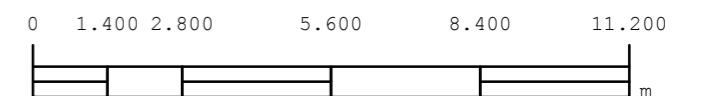
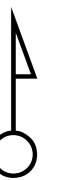
Superficie *Pinus radiata* 9.774,1 ha
 Porcentaje de ocupación 22,8%
 Biomasa potencial total 1.280.983 t/ha

Biomasa de *P. radiata* (t)

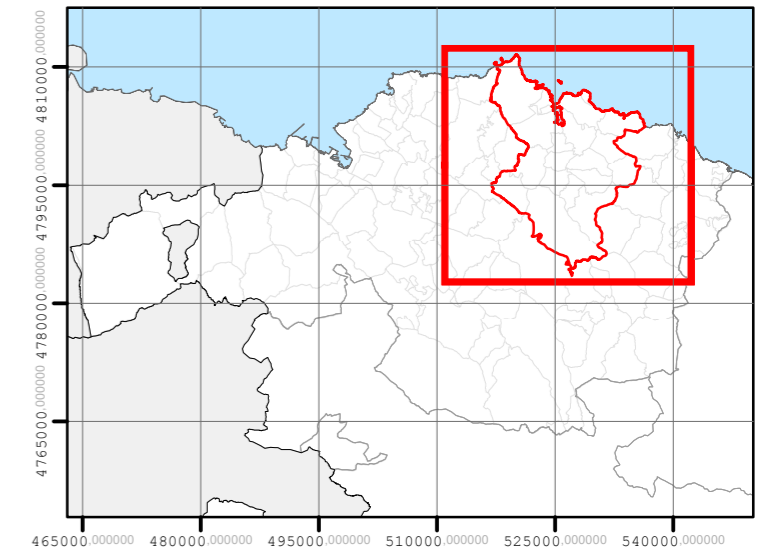
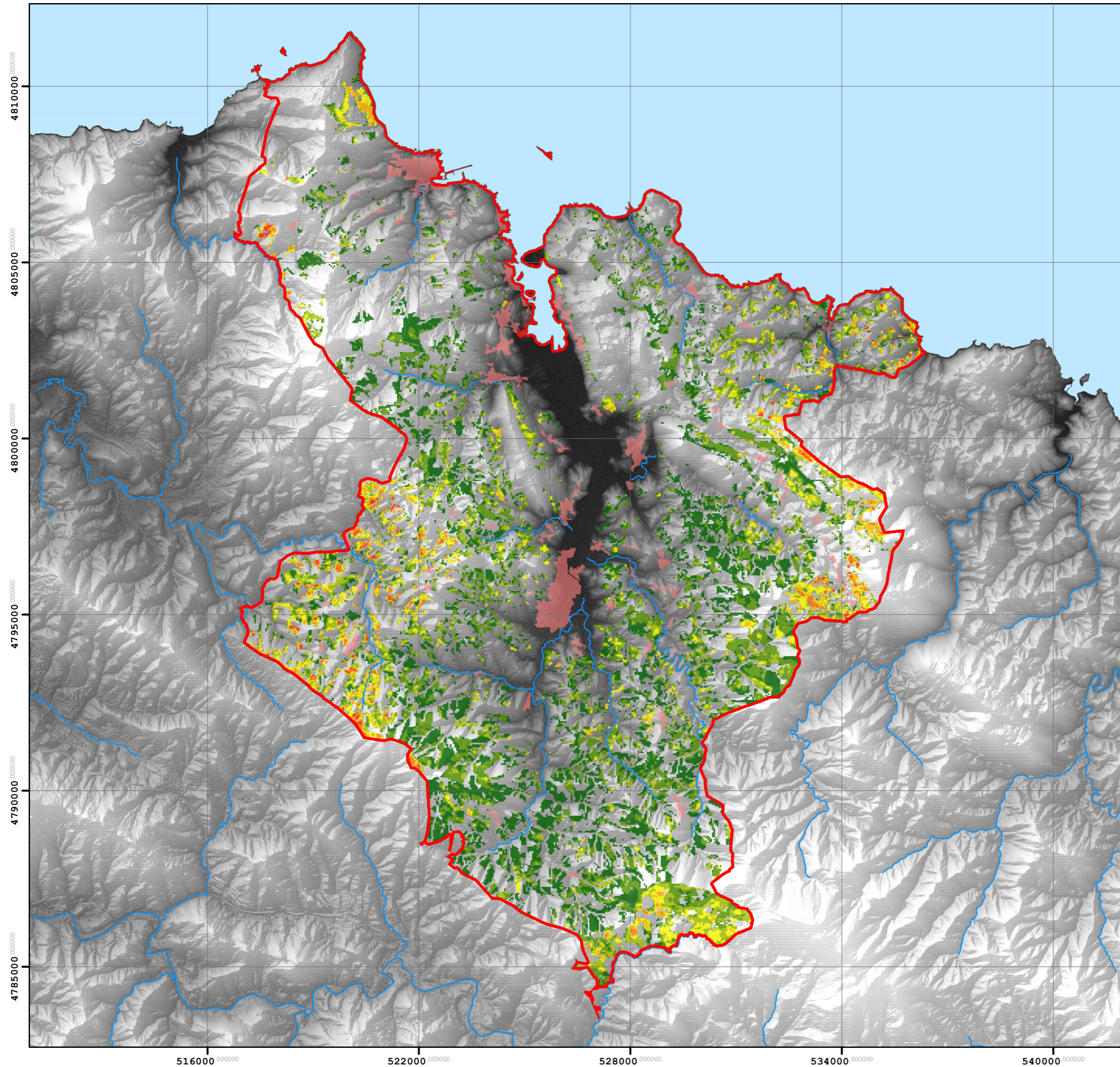


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



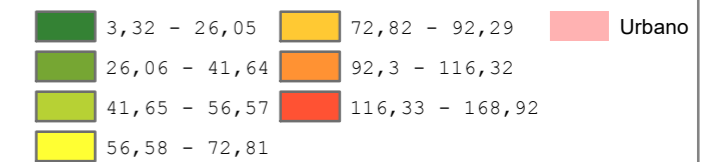
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



GERNIKA-BERMEO

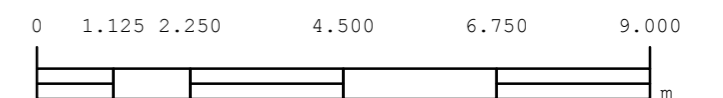
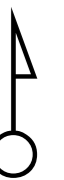
Superficie *Pinus radiata* 9.293,2 ha
 Porcentaje de ocupación 33,4%
 Biomasa potencial total 1.244.535 t/ha

Biomasa de *P. radiata* (t)

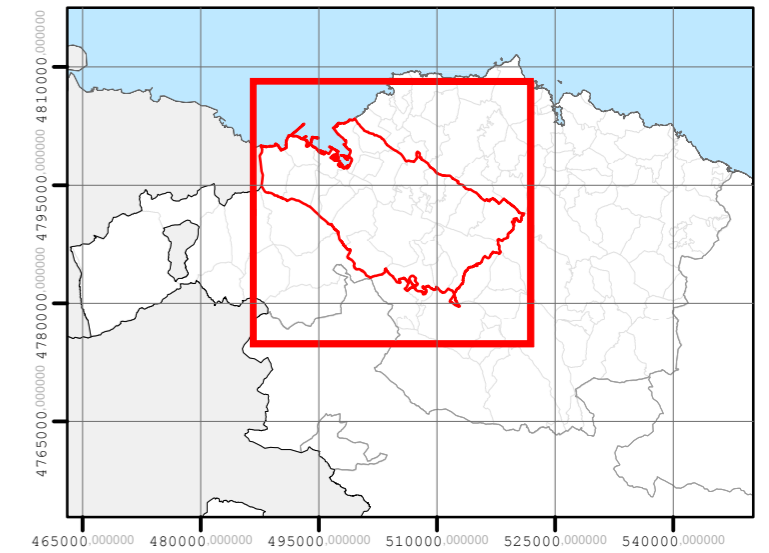
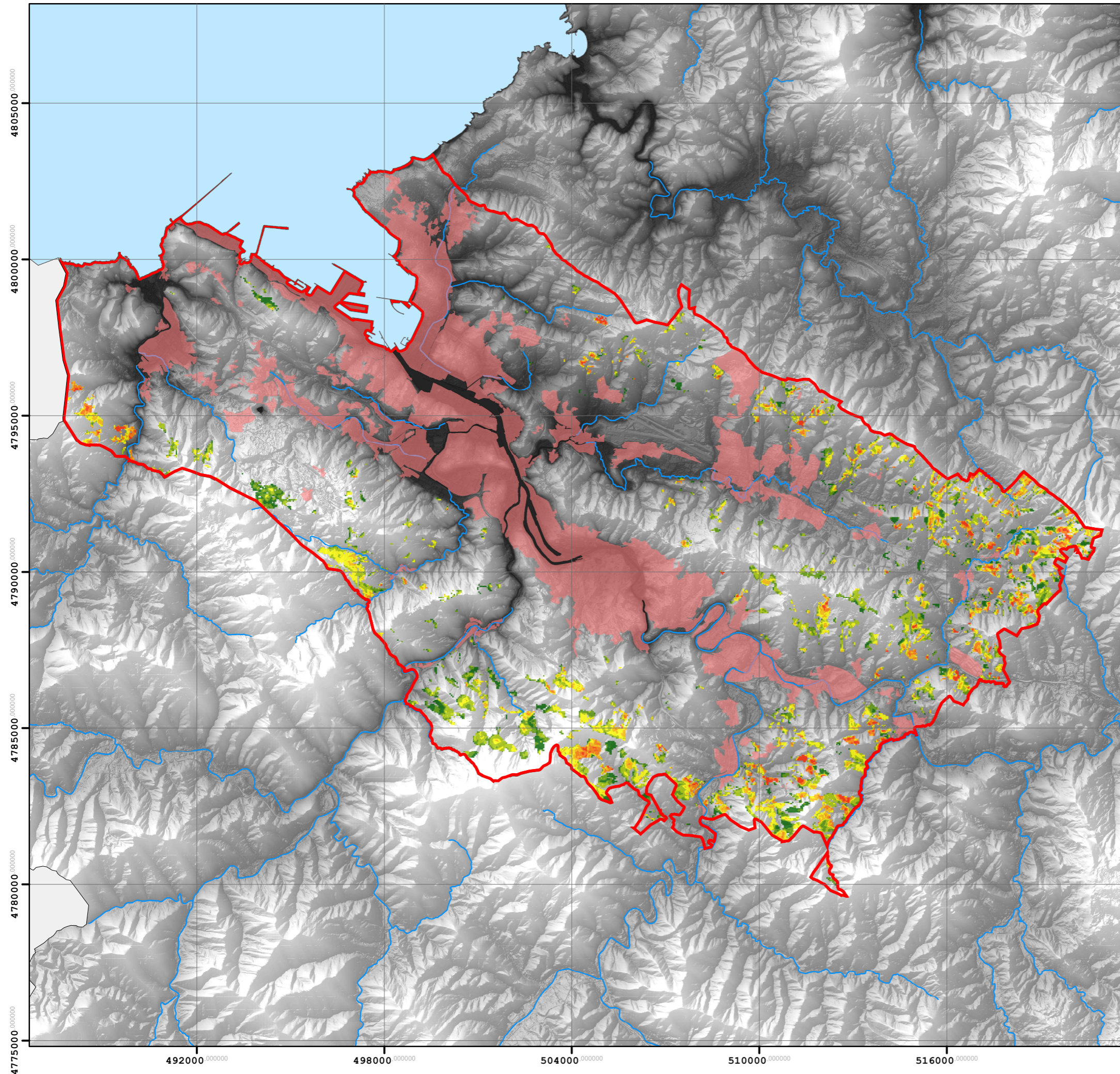


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



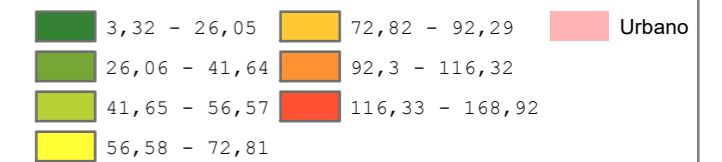
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



GRAN BILBAO

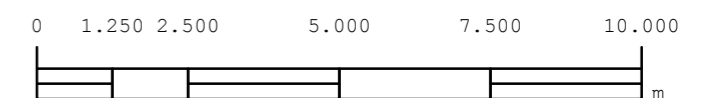
Superficie *Pinus radiata* 3.139,2 ha
 Porcentaje de ocupación 8,4%
 Biomasa potencial total 430.946,8 t/ha

Biomasa de *P. radiata* (t)

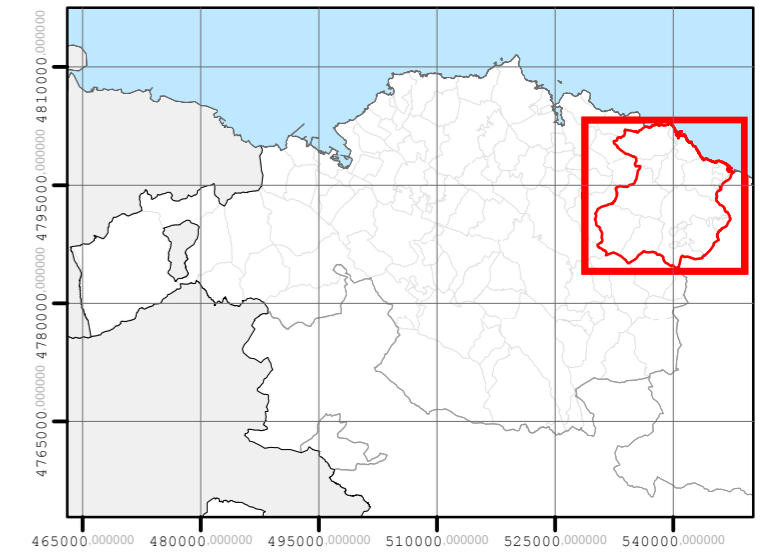
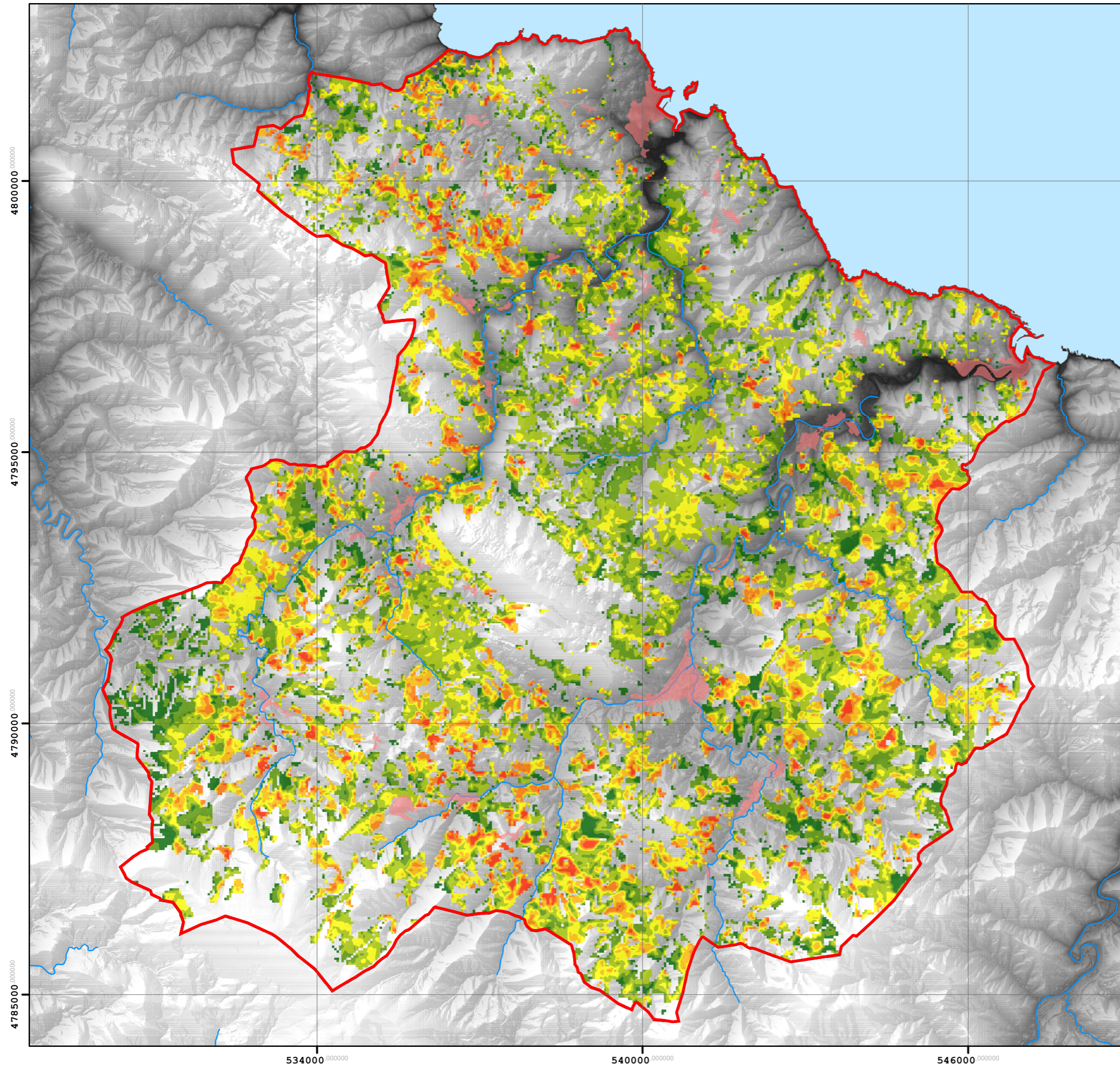


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



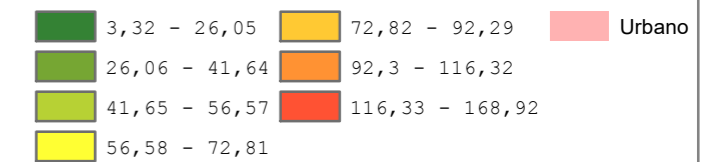
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



MARKINA-ONDARROA

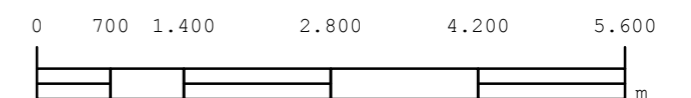
Superficie *Pinus radiata* 9.300,2 ha
 Porcentaje de ocupación 45,2%
 Biomasa potencial total 1.201.139 t/ha

Biomasa de *P. radiata* (t)

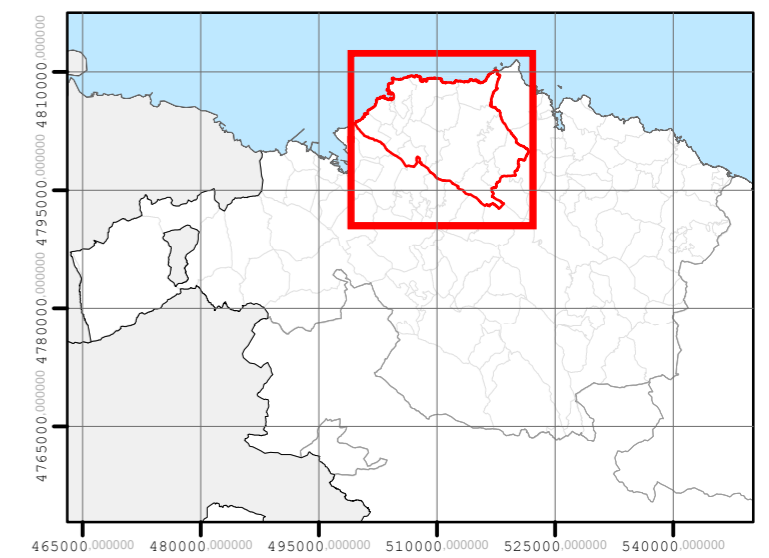
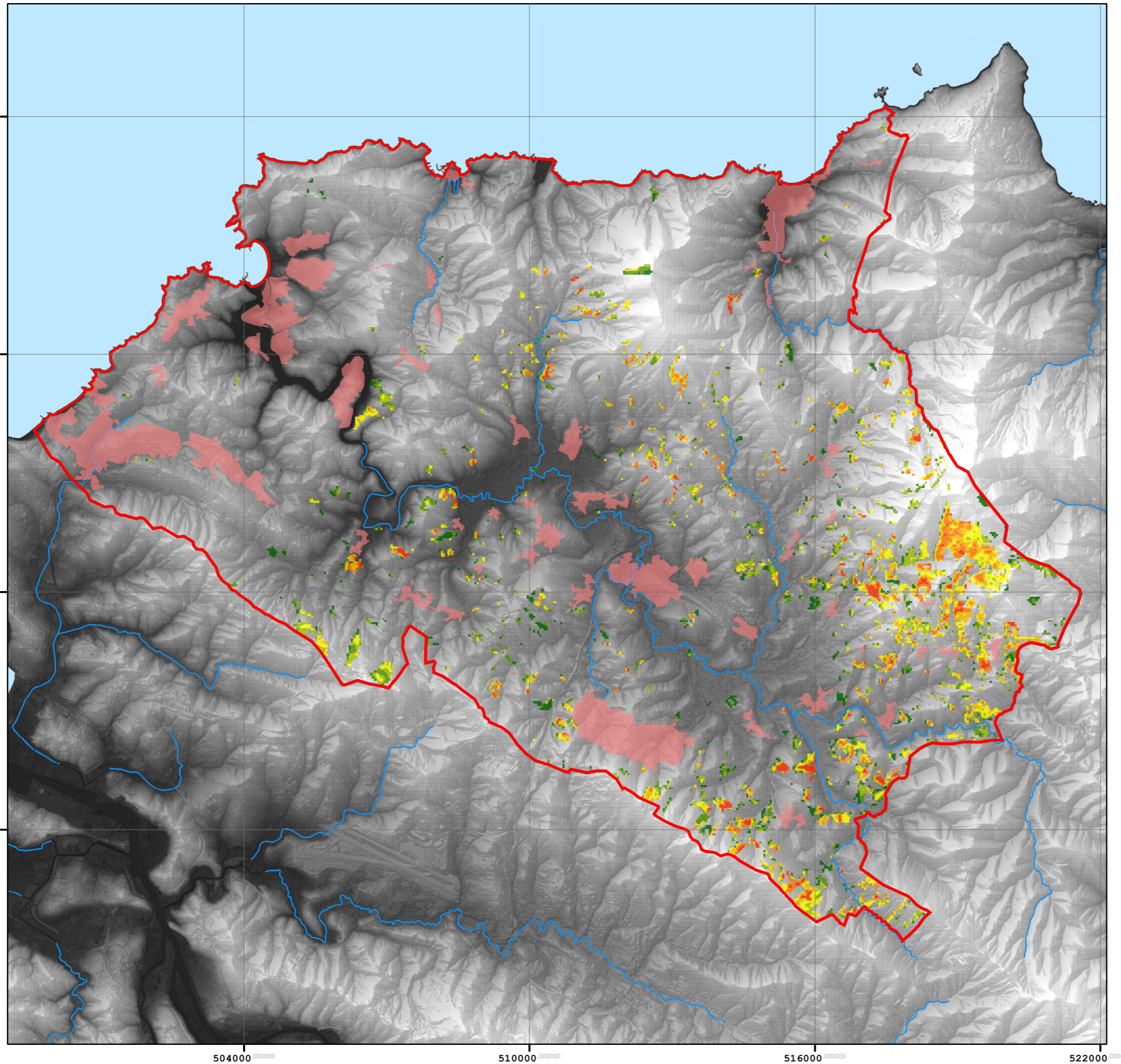


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



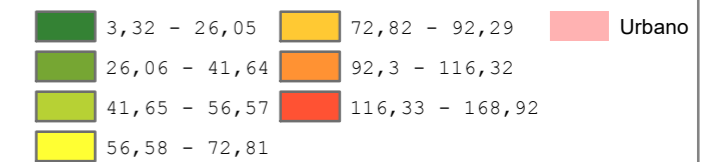
BIOMASA FORESTAL POTENCIAL DISPONIBLE DE RESTOS DE *P. radiata* EN MASAS DE LATIZAL/FUSTAL COMARCAS



PLENTZIA-MUNGIA

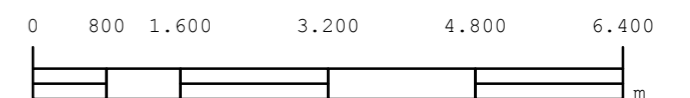
Superficie *Pinus radiata* 1.386,4 ha
 Porcentaje de ocupación 6,7%
 Biomasa potencial total 222.213,3 t/ha

Biomasa de *P. radiata* (t)

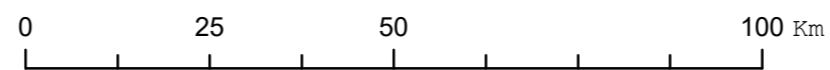
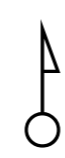
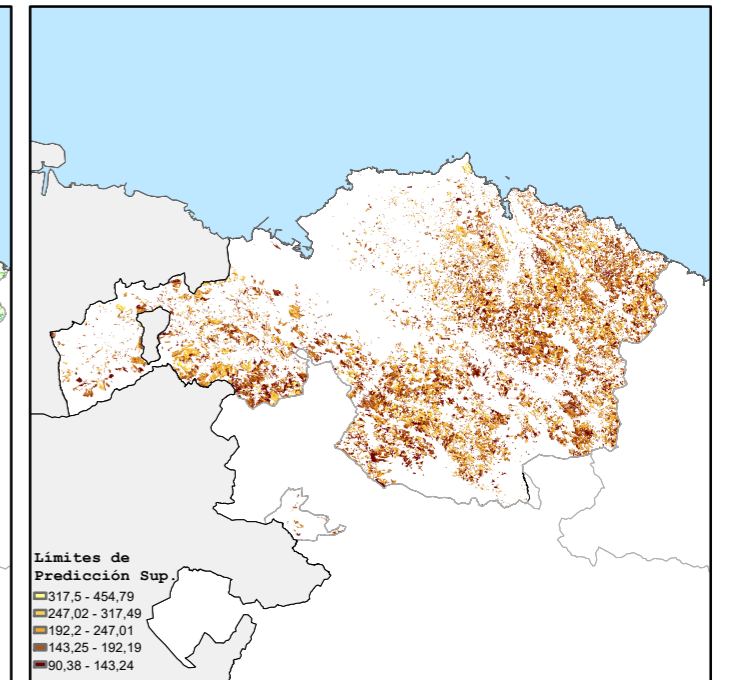
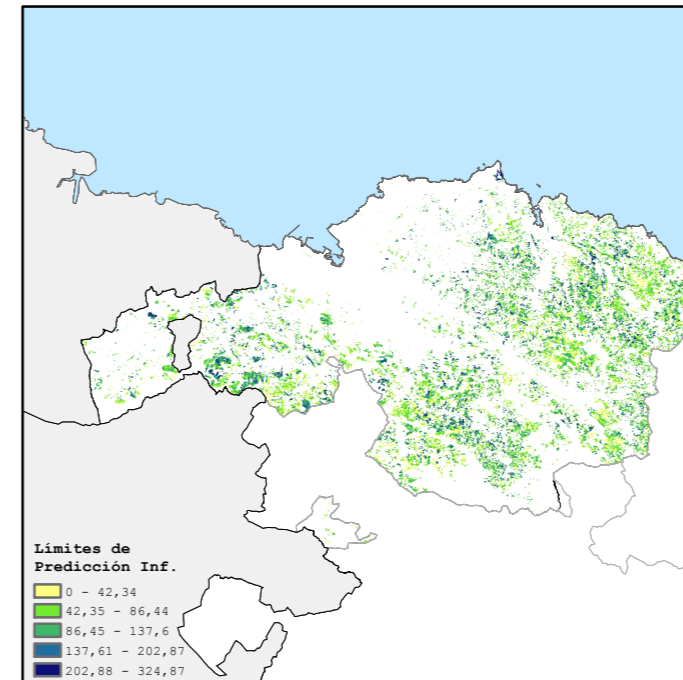
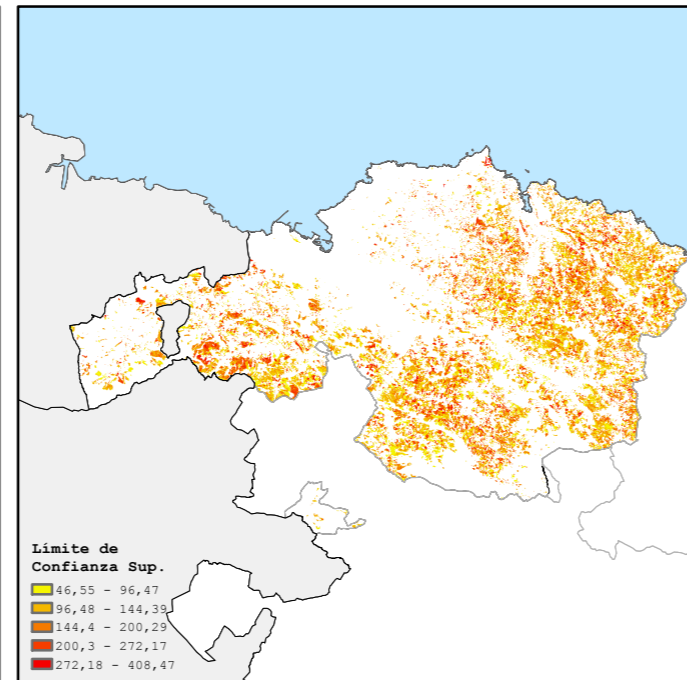
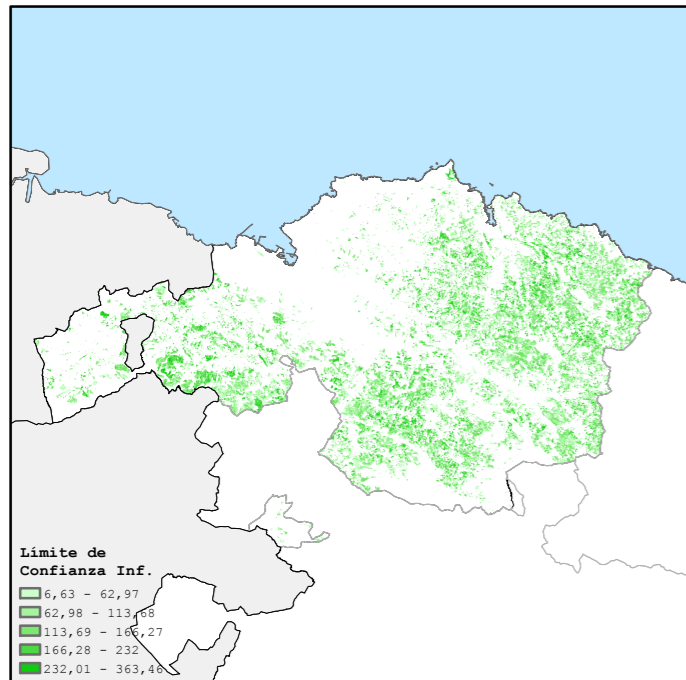
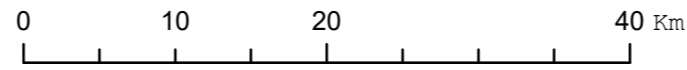
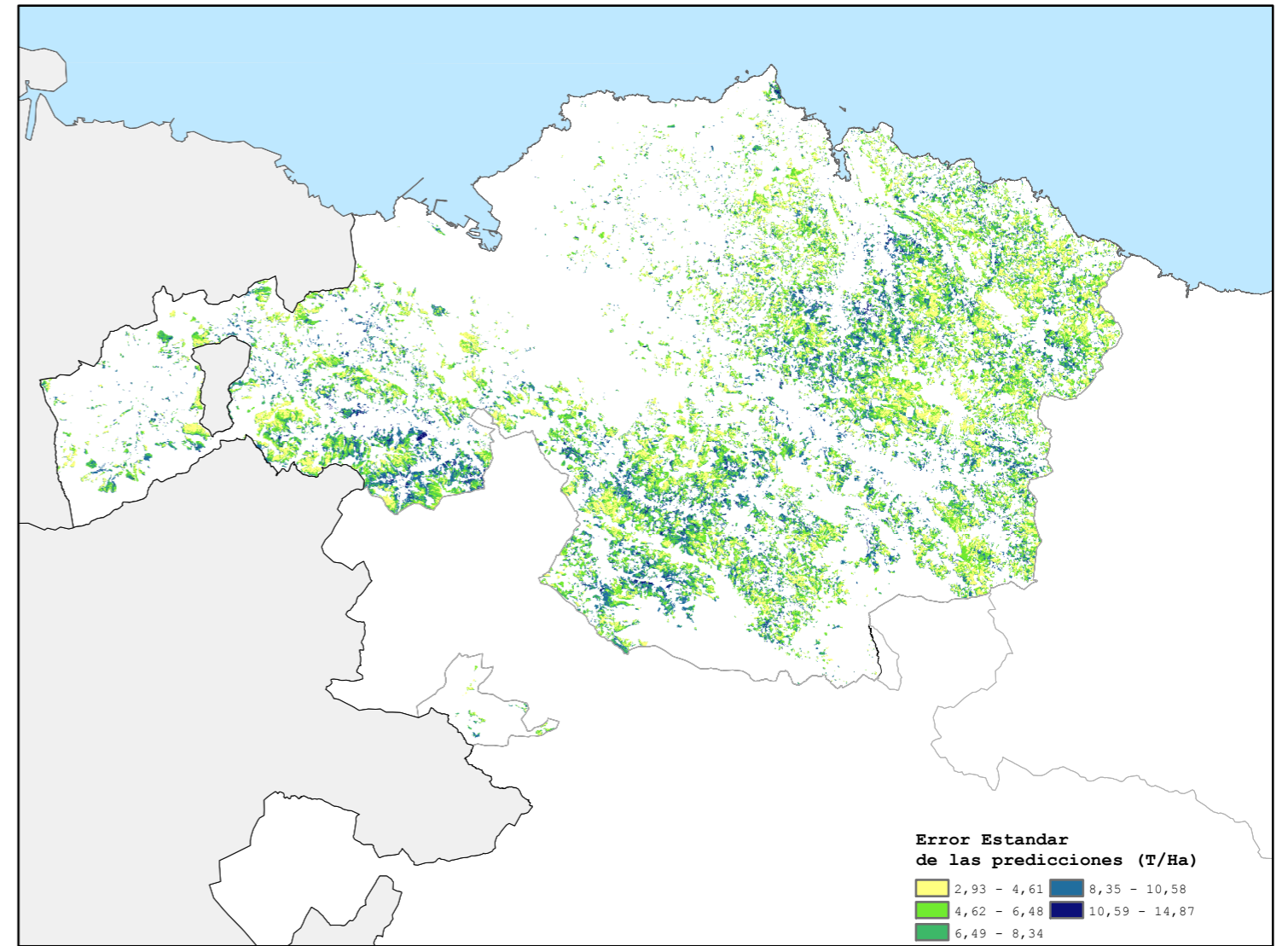
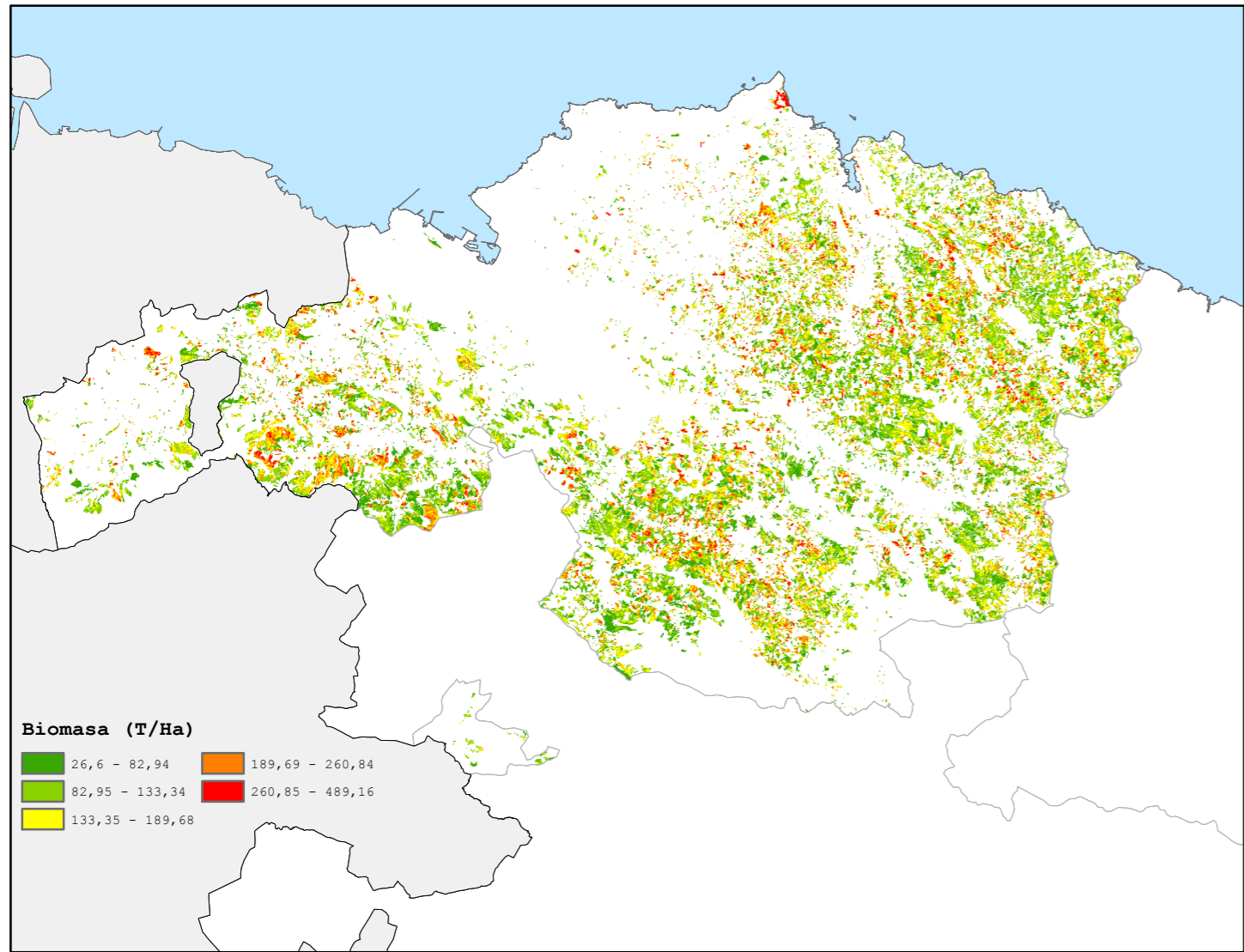


Tamaño del pixel = 0,196 ha

Coordinate System: ETRS 1989 UTM Zone 30N
 Projection: Transverse Mercator
 Datum: ETRS 1989
 False Easting: 500.000,0000
 False Northing: 0,0000
 Central Meridian: -3,0000
 Scale Factor: 0,9996
 Latitude Of Origin: 0,0000
 Units: Meter



MAPAS DE INCERTIDUMBRE



ANEXO II
PARCELAS CON DATOS ANÓMALOS

