



8º CONGRESO FORESTAL ESPAÑOL

La **Ciencia forestal** y su contribución a
los **Objetivos de Desarrollo Sostenible**



8CFE

Edita: Sociedad Española de Ciencias Forestales

Cataluña | Catalunya - 27 junio | juny - 1 julio | juliol 2022

ISBN 978-84-941695-6-4

© Sociedad Española de Ciencias Forestales

Organiza



Evaluación y validación de los modelos de crecimiento forestal IBEROPT e IBEROPS para su implementación en SIMANFOR

VÁZQUEZ-VELOSO, A.^{1,2}, PANDO, V.^{1,3}, ORDÓÑEZ, A.C.^{1,2}, BRAVO, F.^{1,2}

¹ Instituto Universitario de Investigación en Gestión Forestal Sostenible (iuFOR) Universidad de Valladolid – INIA

² Departamento de Producción Vegetal y Recursos Forestales, ETS de Ingenierías Agrarias de Palencia, Universidad de Valladolid

³ Departamento de Estadística e Investigación Operativa. ETS de Ingenierías Agrarias de Palencia, Universidad de Valladolid

Resumen

La modelización forestal es una herramienta que nos permite predecir la evolución de masas forestales y explicar su estado bajo determinadas condiciones, respaldando su gestión. Ecuaciones para el cálculo de variables de interés forestal a nivel de árbol, clases de tamaño y masa han sido desarrolladas, cuya precisión en las predicciones es comúnmente evaluada de manera individual. Para su empleo, se desarrollaron modelos que agrupan ecuaciones y simuladores que facilitan su utilización, centralizando la información y facilitando su empleo. No obstante, pese a que la validación de ecuaciones individuales es un campo conocido, la validación conjunta del conglomerado de ecuaciones y reglas que componen los modelos es un terreno poco explorado. En este trabajo se dan los primeros pasos para el desarrollo de una metodología de análisis que permita la validación de modelos. Para ello, se tratan de validar los modelos IBEROPT e IBEROPS desarrollados para *Pinus pinaster mesogeensis* y *Pinus sylvestris*, respectivamente, utilizando los datos del Inventario Forestal Nacional (IFN) y la plataforma SIMANFOR para su empleo. Los resultados obtenidos demuestran que la metodología desarrollada es adecuada para mejorar la precisión de las ecuaciones de crecimiento, aunque tiene carencias en las ecuaciones de supervivencia y masa incorporada.

Palabras clave

Desarrollo de metodología, supervivencia, masa incorporada, modelización, selvicultura.

1. Introducción

Cuantificar y clasificar han sido trabajos que nos han llevado a simplificar la realidad para poder entenderla y manejarla, con el objetivo de lograr nuestros propósitos. En el sector forestal, este proceso se realiza a través de la modelización, que supone crear una representación abstracta de la realidad forestal en la que simular procesos y predecir sus resultados (SECF, 2021). En España se han hecho grandes esfuerzos por avanzar en el campo de la modelización, desarrollando ecuaciones para el cálculo del crecimiento y producción de las especies forestales más atractivas desde el punto de vista ecológico y comercial (BRAVO et al., 2011). Además, también se han desarrollado plataformas que permiten el empleo de estas ecuaciones para el desarrollo de modelos conjuntos, como es el caso de SIMANFOR (BRAVO et al., 2010; BRAVO et al., 2012; SIMANFOR, 2021), con una interfaz que facilita el manejo de esta información. Por otro lado, y no menos importante, desde el sector público se ha realizado una gran apuesta para el desarrollo de una gran base de datos, el Inventario Forestal Nacional (IFN) (ALBERDI et al., 2016; BRAVO et al., 2002), la cual permite tanto el desarrollo de modelos como la predicción del crecimiento y producción de nuestros montes, información de gran importancia para el avance del sector forestal. Pese a que hasta ahora los esfuerzos estaban focalizados en el desarrollo de modelos, el avance en este campo y la mejora de las herramientas ha provocado que los usuarios se pregunten sobre la precisión de sus predicciones. Este campo novedoso se conoce como la evaluación y validación de modelos, y pese a que algunos autores ya han tratado sobre ello (PRETZSCH, 2009; WEISKITTEL et al., 2011), lo cierto es que

todavía está poco desarrollado. Mientras que la precisión de las predicciones de cada ecuación suele ser evaluada por sus desarrolladores antes de publicarlas, esto no es una práctica común en la validación de modelos, que implica la validación conjunta de todas las ecuaciones desarrolladas para una determinada especie y lugar. En este campo novedoso existe una importante falta de conocimiento acerca de cómo realizar la validación de modelos forestales, incluyendo sus tres partes fundamentales: crecimiento, supervivencia y masa incorporada. El propósito de este estudio es sentar las bases para el desarrollo de una primera metodología que permita la validación de modelos forestales. Para ello, hemos tomado los modelos IBERO^{PT} (LIZARRALDE et al., 2010a) e IBERO^{PS} (LIZARRALDE et al., 2010b), modelos desarrollados para *Pinus pinaster* y *Pinus sylvestris*, como objeto de evaluación. Su empleo se ha realizado a través de la plataforma SIMANFOR en una versión alojada en el supercomputador SCAYLE (SCAYLE, 2019), y empleando los datos de IFN2 e IFN3 como base para la simulación y comparación de sus resultados, respectivamente. A través de este trabajo pretendemos analizar la precisión de los modelos IBERO^{PT} e IBERO^{PS}, calibrar y validar su funcionamiento, y desarrollar una metodología que permita sistematizar la validación de modelos para su posterior empleo.

2. Objetivos

Los objetivos principales de este estudio son:

- Desarrollar una metodología que permita la validación conjunta de modelos, incluyendo sus ecuaciones de crecimiento en diámetro y altura, supervivencia y masa incorporada
- Analizar la precisión de las predicciones de los modelos IBERO^{PT} e IBERO^{PS}, corrigiendo y validando sus ecuaciones, o determinando en su defecto cuáles han de ser analizadas bajo otra metodología

3. Metodología

Datos

Se han utilizado los datos de parcelas forestales permanentes de las ediciones 2 y 3 del Inventario Forestal Nacional Español (IFN) (ALBERDI et al., 2016; BRAVO et al., 2002), utilizando los datos del IFN2 como base para la simulación (datos predichos) y el IFN3 como datos reales para comprobar la precisión de las mismas (datos observados). El IFN es una base de datos diferente a la utilizada para el desarrollo de los modelos, lo cual es fundamental para un proceso de evaluación (VANCLAY, 1994). Tras una preparación inicial y una adaptación de la base de datos a las necesidades del simulador mediante el empleo de R (versión 3.6.3; R CORE TEAM, 2019), se han filtrado aquellas parcelas que cumplen los requisitos de los modelos IBERO^{PT} e IBERO^{PS}, objeto de validación:

- parcela medida en ambos inventarios y sobre la que no hubo alteraciones destacables, como cortas o alta mortalidad debido a incendios forestales u otros factores
- conocer la edad promedio de la masa, debido a que es una variable necesaria en las ecuaciones de crecimiento
- tener como especie principal de la parcela *Pinus pinaster* o *Pinus sylvestris*, dependiendo del modelo
- estar presentes las parcelas en el rango geográfico para el que fueron diseñados los modelos

Tras aplicar dichos filtros a los datos iniciales, hemos encontrado un total de 49 parcelas con 575 árboles para *P. pinaster*, y 136 parcelas con 2.810 árboles para *P. sylvestris*, respectivamente, que cumplen estos criterios. Las parcelas del IFN3 se han modificado para que el tiempo transcurrido entre las mediciones del IFN2 e IFN3 sea de 10 años, recalculando los crecimientos en altura y diámetro de los árboles del IFN3 de manera lineal (asumiendo crecimientos iguales todos los años), a la nueva edad de la parcela (10 años después de ser medida en el IFN2), y permitiendo así hacer

coincidir en el tiempo los datos predichos (mediante dos simulaciones consecutivas de 5 años del IFN2) con los observados (datos del IFN3) para su comparación. Esta modificación es necesaria debido al diseño original de los modelos, dado que sus ecuaciones de crecimiento fueron diseñadas para períodos de tiempo de 5 años, lo que imposibilita su aplicación para rangos temporales diferentes.

Por otro lado, se han utilizado datos de precipitación ($\text{mm}\cdot\text{año}^{-1}$) y temperatura media anual ($^{\circ}\text{C}$) obtenidos de la Agencia Estatal de Meteorología (AEMET, 2021) para evaluar una posible mejora en la aplicación de la ecuación de supervivencia. Estos datos fueron obtenidos de las estaciones meteorológicas más cercanas a las parcelas de estudio para los años en los que se realizó la medición de la parcela, detallando su forma de empleo en apartados posteriores.

Simulaciones de crecimiento

Para las simulaciones del crecimiento se ha utilizado la plataforma SIMANFOR, empleando los modelos IBEROPT e IBEROPS^{PS}. Las simulaciones han sido ejecutadas en una versión del simulador alojada en el centro de Supercomputación de Castilla y León (SCAYLE, 2019). El empleo de este servicio nos permite tener una mayor capacidad de cómputo, traducida en una mayor rapidez de cálculo. Para cada uno de los casos de estudio planteados para evaluar las ecuaciones que componen cada modelo, ha sido necesario modificar los modelos y realizar una nueva simulación, siguiendo el diagrama de flujo general mostrado en la Figura 1.

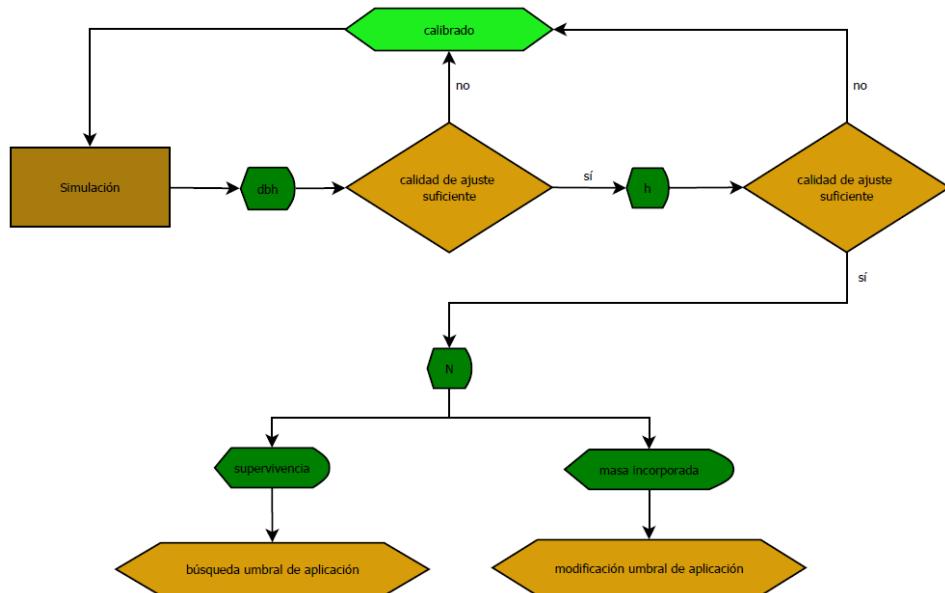


Figura 1. Procedimiento empleado para la validación de los modelos IBEROPT e IBEROPS^{PS}.

En cada simulación se proyecta el crecimiento de las parcelas del IFN2 10 años (divididas en dos proyecciones de 5 años cada una, debido al diseño de los modelos originales). Para los datos del IFN3 se ha realizado un recálculo de las parcelas, consistente en un proceso de inicialización en SIMANFOR sin proyecciones en el tiempo ni cortas, que permite disponer de todas las variables de estudio calculadas por el mismo modelo y simulador.

Tras realizar las simulaciones, los datos de salida han sido procesados con R (R CORE TEAM, 2019), dotándolos de una mejor estructura de cara a su análisis. Finalmente se obtienen para cada caso de estudio dos archivos diferentes con información de árboles y parcelas, respectivamente, englobando los datos predichos (IFN2 + simulaciones) y observados (IFN3) de cada uno de ellos, con los que se realizan los análisis estadísticos.

Modelos empleados

Los modelos seleccionados para su calibración fueron IBERO^{PT} e IBERO^{PS}, siendo ambos modelos de crecimiento de árbol individual independientes de la distancia, programados en SIMANFOR y adaptados a las necesidades de este estudio para cada una de las ecuaciones a analizar, tal y como se describe en el apartado siguiente. Estos modelos fueron desarrollados para el Sistema Ibérico Meridional (IBERO^{PT}), comprendiendo las provincias de Soria, Guadalajara, Cuenca y Teruel; y para los Sistemas Ibérico y Central (IBERO^{PS}), englobando las provincias de Ávila, Burgos, Segovia y Soria. Aunque ambos modelos están formados por una mayor cantidad de ecuaciones para el cálculo de variables de interés forestal (copa, volumen, biomasa...), todas ellas dependen de las variables fundamentales de árbol: d (diámetro normal, medido a la altura del pecho, 1,30 m, en cm), h (altura del árbol, en m) y expan (factor de expansión del árbol, que expresa la cantidad de árboles equivalentes al árbol sujeto a nivel de hectárea). De este modo, nuestro estudio se ha centrado en el análisis de las ecuaciones de crecimiento en diámetro y altura (LIZARRALDE, 2008), comparando directamente los valores predichos de cada árbol con los valores observados; ecuaciones de supervivencia (BRAVO-OVIEDO et al., 2006), que actúan sobre el factor de expansión de los árboles, pero se monitorizan sobre la densidad y área basimétrica de la masa; y masa incorporada (BRAVO et al., 2008), que se aplican sobre los datos de parcela y se monitorizan, también, sobre la densidad y área basimétrica de la masa.

De cara a mejorar las predicciones de las ecuaciones de supervivencia, se propone la creación de un valor umbral que determine cuándo es necesaria su aplicación. Este valor umbral funciona como un valor de referencia, de manera que si es superado se aplica la ecuación, mientras que si no se supera se asume que todos los árboles de la parcela sobreviven. Se ha diseñado siguiendo la metodología propuesta por RODRÍGUEZ et al. (2020) en base a Índice de Reineke (REINEKE, 1933) e Índice de Reineke Máximo (RODRÍGUEZ et al., 2020), utilizando los coeficientes específicos de esta ecuación propuestos por DEL RÍO et al., (2006) para los modelos estudiados. El Índice de Reineke funciona en este valor umbral como dato del estado actual de la masa, mientras que el Índice de Reineke Máximo funciona como valor umbral (de referencia, determinando la máxima capacidad de carga de la masa en función de las condiciones climáticas, para lo que usa el Índice de aridez de Martonne (MARTONNE, 1926)), de manera que la ecuación de supervivencia se aplica sólo cuando $I_r > I_m$.

En cuanto a las ecuaciones de masa incorporada, se ha decidido estudiar la influencia de la “masa incorporada técnica” sobre las predicciones, entendiendo por esta la masa incorporada en los círculos concéntricos de las parcelas cuyo diámetro mínimo inventariable es mayor que 7,5 cm; y la alteración del umbral de aplicación establecido para cada modelo. Una explicación más extensa sobre los casos de estudio puede consultarse en VÁZQUEZ-VELOSO (2021), así como el aspecto de las ecuaciones originales y explicaciones más detalladas.

Análisis estadístico

Para el análisis estadístico se ha utilizado el software SAS/STAT (Versión 3.8 (Edición Enterprise); SAS, 2011), por el cual se hace una comparación de los datos observados (IFN3) con los predichos (IFN2 + simulaciones). Mientras que este mismo procedimiento ha sido empleado por estudios enfocados en evaluación de modelos (MOORE, 2010; ROXBURGH et al., 2019), para el presente estudio se ha propuesto el empleo del coeficiente de correlación de concordancia para variables continuas, también conocido como coeficiente de Lin (LAWRENCE & LIN, 1989; LIN, 2000). Se calcula también el factor de precisión (C_b) de cada caso de estudio, siendo este la medida de adecuación de la recta de ajuste a la bisectriz del primer cuadrante, como el cociente entre r (coeficiente de correlación de Pearson) y r_c (coeficiente de correlación de Lin).

El valor del coeficiente de Lin se relaciona con un grado de concordancia entre las predicciones y las observaciones, tal y como se muestra la Tabla 1, desarrollada para evaluar la precisión de la toma de datos de aparatos de laboratorio, y que tomaremos como una aproximación en nuestro estudio. Empleando esta aproximación, podemos agrupar cualitativamente los resultados de nuestras predicciones, lo que resulta más sencillo desde el punto de vista de la evaluación global de la mejora del modelo. Siguiendo este criterio, los ajustes realizados sobre las ecuaciones originales pretenden

mejorar el grado de concordancia de las predicciones, intentando lograr un grado de concordancia moderado o lo más cerca posible a este para cada caso de estudio. Para ello, se corregirán los problemas de sesgo, ajuste, o ambos, detectados en las ecuaciones. Para este análisis se han estudiado el diámetro normal (d) y la altura total (h) para los datos de árbol; y la densidad (N), área basimétrica (G) y altura dominante (H_0) para los datos de parcelas, siendo elegidas como las variables más representativas de las ecuaciones a analizar.

Tabla 1. Clasificación de los grados de concordancia para cada rango de valores del coeficiente de Lin, adaptada de CAMACHO-SANDOVAL (2008).

Grado de concordancia según el valor del coeficiente de Lin	
Grado de concordancia	Valor del coeficiente de Lin
Casi perfecta	> 0,99
Sustancial	0,95 – 0,99
Moderada	0,90 – 0,95
Pobre	< 0,90

4. Resultados

Mediante la metodología desarrollada se ha conseguido mejorar las predicciones de las ecuaciones de crecimiento en diámetro y altura para ambos modelos. Por el contrario, los casos de estudio ensayados para la mejora de las ecuaciones de supervivencia y masa incorporada no han supuesto mejoras respecto al modelo original.

Ecuaciones de crecimiento en diámetro y altura

Tras una primera simulación de referencia en la que se han utilizado los modelos IBERO^{PT} e IBERO^{PS} originales, han sido detectados errores de sesgo para las ecuaciones de crecimiento en diámetro y altura de ambos modelos, mientras que los ajustes de las ecuaciones han sido satisfactorios. Para su corrección, se ha realizado una calibración conjunta de ambas ecuaciones en dos pasos:

- calibración de la ecuación de crecimiento en diámetro, dado que es una variable necesaria para conocer la altura del árbol
- calibración de la ecuación de crecimiento en altura, utilizando las predicciones de crecimiento en diámetro ya calibradas

Las ecuaciones resultantes tras la calibración se muestran a continuación, destacando en negrita la calibración realizada, y siendo 1 y 2 las ecuaciones de crecimiento en diámetro de IBERO^{PT} e IBERO^{PS}, y 3 y 4 las ecuaciones de crecimiento en altura para IBERO^{PT} e IBERO^{PS}, respectivamente.

$$d_{i5} = 1,18 + e^{0,2030 \cdot \ln(d \cdot 10) + 0,4414 \cdot \ln((cr+0,2)/1,2) + 0,8379 \cdot \ln(SI) - 0,1295 \cdot \sqrt{G} - (0,0007 \cdot bal^2) / (\ln(d \cdot 10))} / 10 \quad (\text{Ecuación 1})$$

$$d_{i5} = 1,70 + e^{-0,3711 + 0,2525 \cdot \ln(d \cdot 10) + 0,7090 \cdot \ln((cr+0,2)/1,2) + 0,9087 \cdot \ln(SI) - 0,1545 \cdot \sqrt{G} - (0,0004 \cdot bal^2) / (\ln(d \cdot 10))} / 10 \quad (\text{Ecuación 2})$$

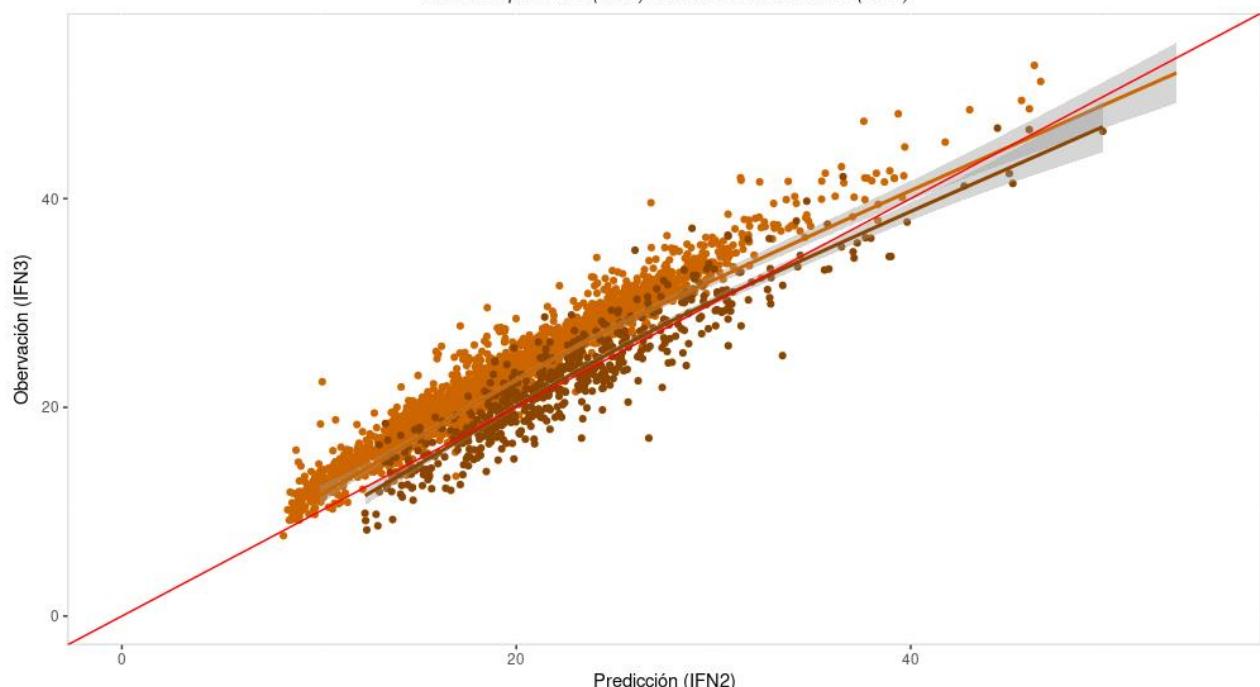
$$h_{i5} = 0,42 + e^{0,2160 + 0,4033 \cdot \ln(d_{i5}/2) - 1,1272 \cdot \ln(d \cdot 10) + 1,1810 \cdot \ln(h \cdot 100) + 3,0162 \cdot cr} / 100 \quad (\text{Ecuación 3})$$

$$h_{i5} = 1,14 + e^{3,1222 - 0,4939 \cdot \ln(d_{i5}/2) + 1,3763 \cdot \ln(SI) - 0,0061 \cdot bal + 0,1876 \cdot \ln(cr)} / 100 \quad (\text{Ecuación 4})$$

siendo:

- d el diámetro normal, medido a la altura del pecho (1,30 m), en cm
- $d_{5\text{a}}$ el incremento del diámetro normal para un periodo de 5 años, en cm
- h la altura total del árbol en m
- $h_{5\text{a}}$ el incremento de la altura total para un periodo de 5 años, en m
- bal el área basimétrica acumulada de los árboles más gruesos que el árbol sujeto, en m^2ha^{-1}
- cr el índice de copa viva, en %
- G el área basimétrica de la masa, en m^2ha^{-1}
- SI el Índice de Sitio, en m, obtenido de BRAVO-OVIEDO et al. (2004) y BRAVO & MONTERO (2001)

En la Figura 2
IBEROPS: modelo original vs calibrado
Diámetros predichos (IFN2) vs diámetros observados (IFN3)



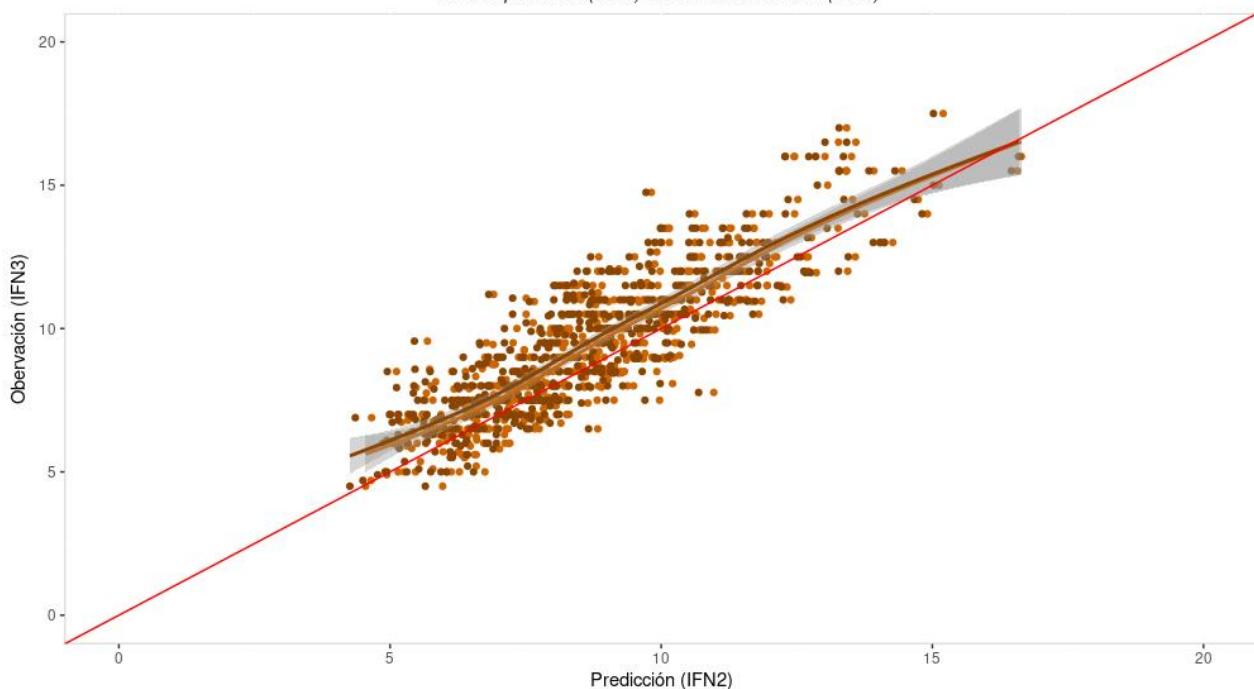
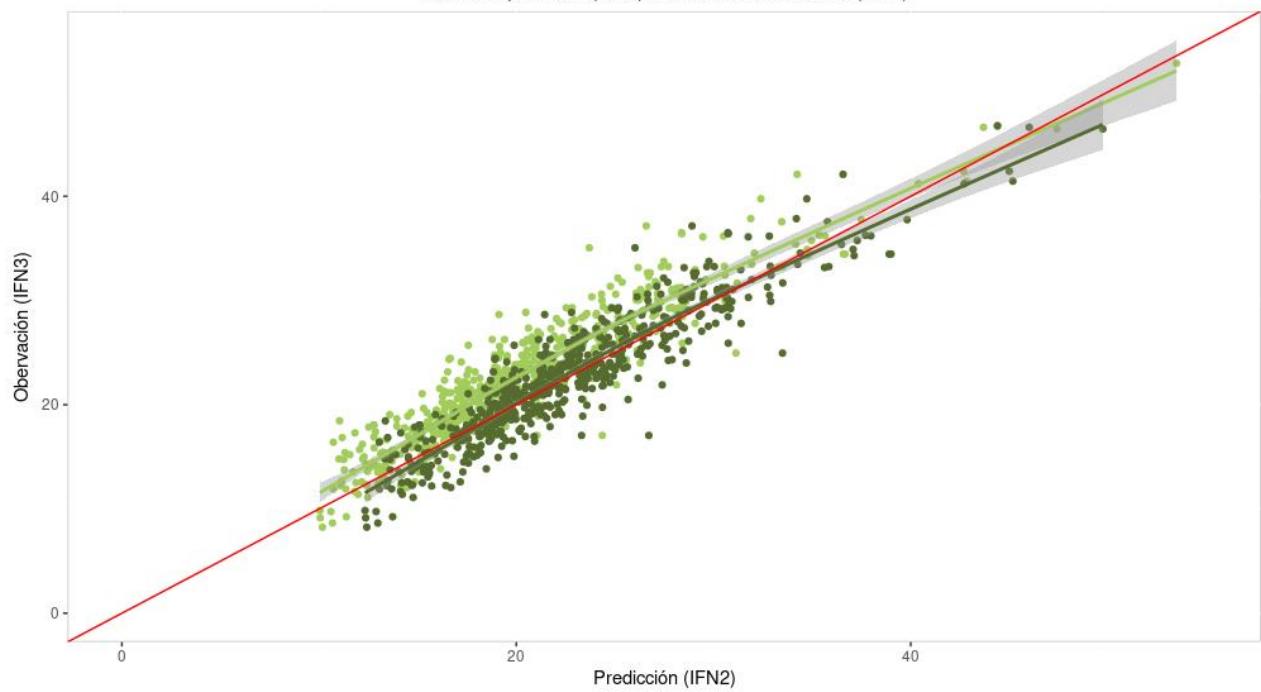
IBEROPS: modelo original vs calibrado
Alturas predichas (IFN2) vs alturas observadas (IFN3)


Figura 2, diseñada a partir de las recomendaciones de PIÑEIRO et al., (2008), se puede comparar la mejora en los datos predichos respecto a los observados para ambas ecuaciones y modelos. Además, para las ecuaciones calibradas se han aumentado tanto la precisión de sus predicciones como las de otras ecuaciones que utilizan estas variables (d y h) como variables independientes para su cálculo, componiendo parte de los mismos modelos. Además, es apreciable una mejora tanto en el ajuste como en la correlación de valores observados y predichos para los análisis realizados tras la calibración del crecimiento, comparándolos con los análisis hechos con las ecuaciones originales, tal y como se aprecia en la Tabla 2.

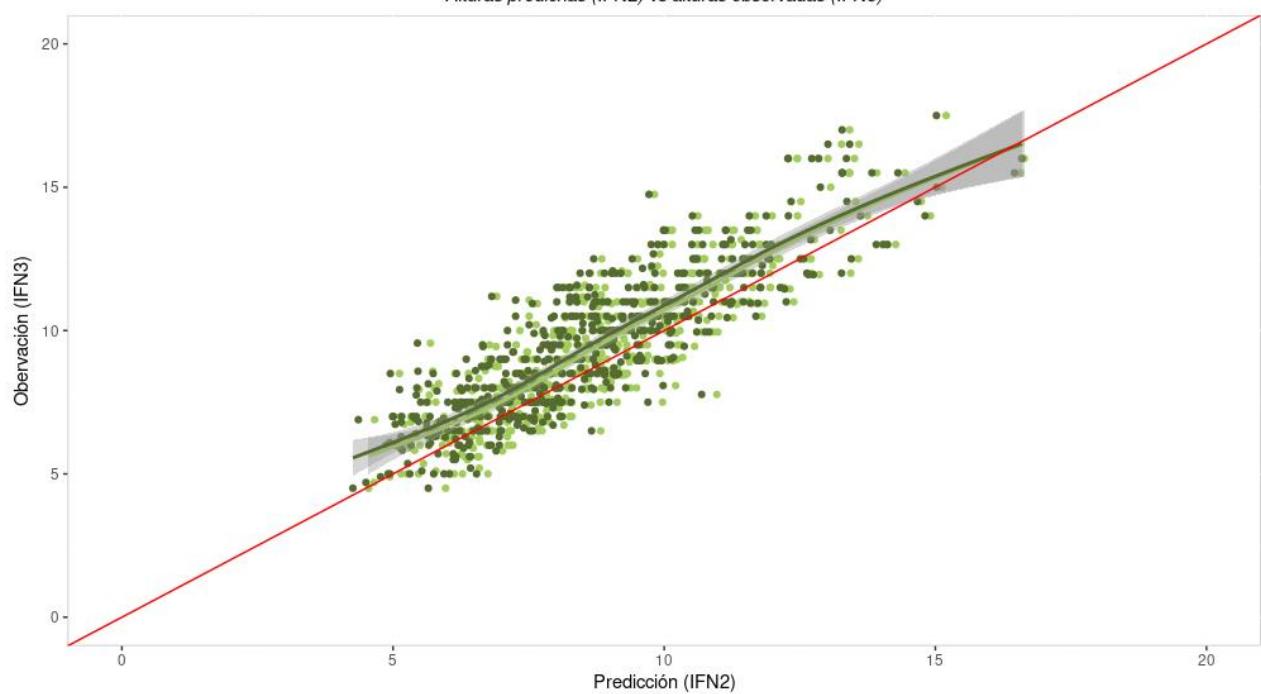
Tabla 2. Resultados obtenidos para el coeficiente de correlación muestral de Pearson (r), el coeficiente de correlación muestral de concordancia de Lin (r_c) y el factor de precisión (C_b) para los distintos casos de estudio de las ecuaciones de crecimiento. Los valores destacados en negrita son los que indican la mejora tras la calibración.

		IBERO ^{PT}		IBERO ^{PS}	
		d	h	d	h
modelo original	r	0,92962	0,86296	0,96995	0,85744
	r_c	0,86221	0,82141	0,83102	0,65214
	C_b	0,92749	0,95185	0,85677	0,76057
d calibrado	r	0,92973	0,86325	0,96990	0,85761
	r_c	0,92727	0,80137	0,96317	0,65380
	C_b	0,99735	0,92832	0,99306	0,76235
d y h calibrado	r	0,92973	0,86336	0,97051	0,86006
	r_c	0,92726	0,85454	0,96389	0,85719
	C_b	0,99734	0,98978	0,99318	0,99666

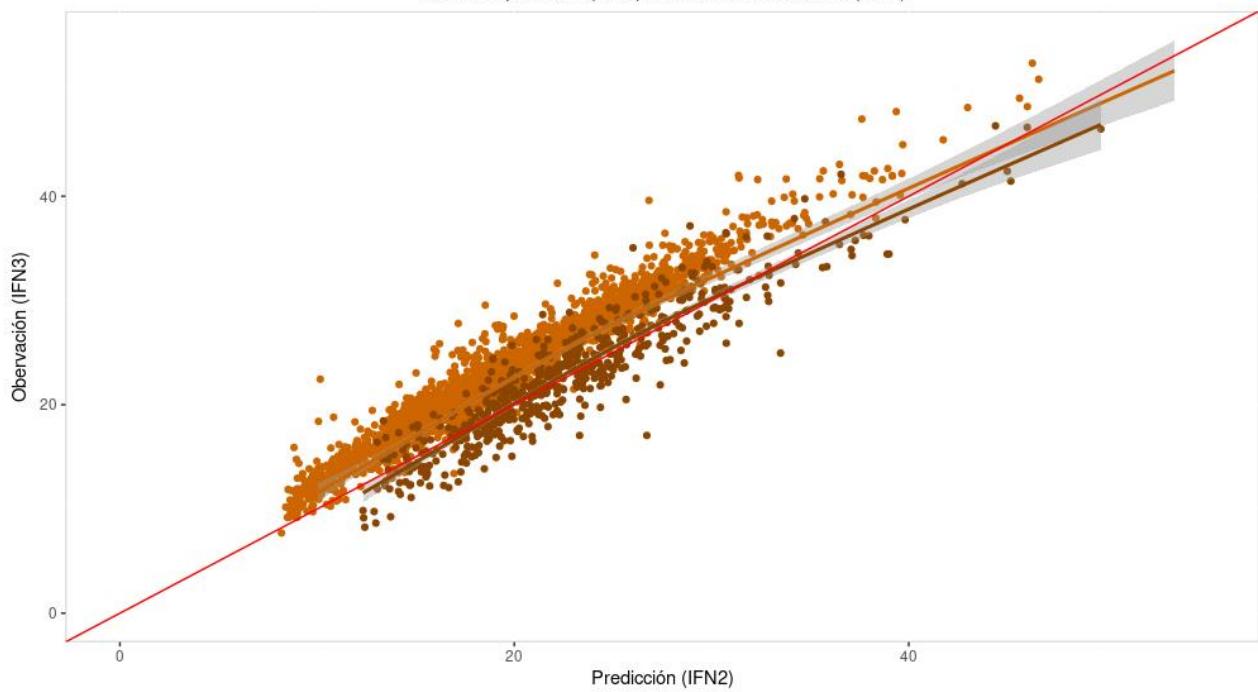
IBEROPT: modelo original vs calibrado
Diámetros predichos (IFN2) vs diámetros observados (IFN3)



IBEROPT: modelo original vs calibrado
Alturas predichas (IFN2) vs alturas observadas (IFN3)



IBEROPS: modelo original vs calibrado
Diámetros predichos (IFN2) vs diámetros observados (IFN3)



IBEROPS: modelo original vs calibrado
Alturas predichas (IFN2) vs alturas observadas (IFN3)

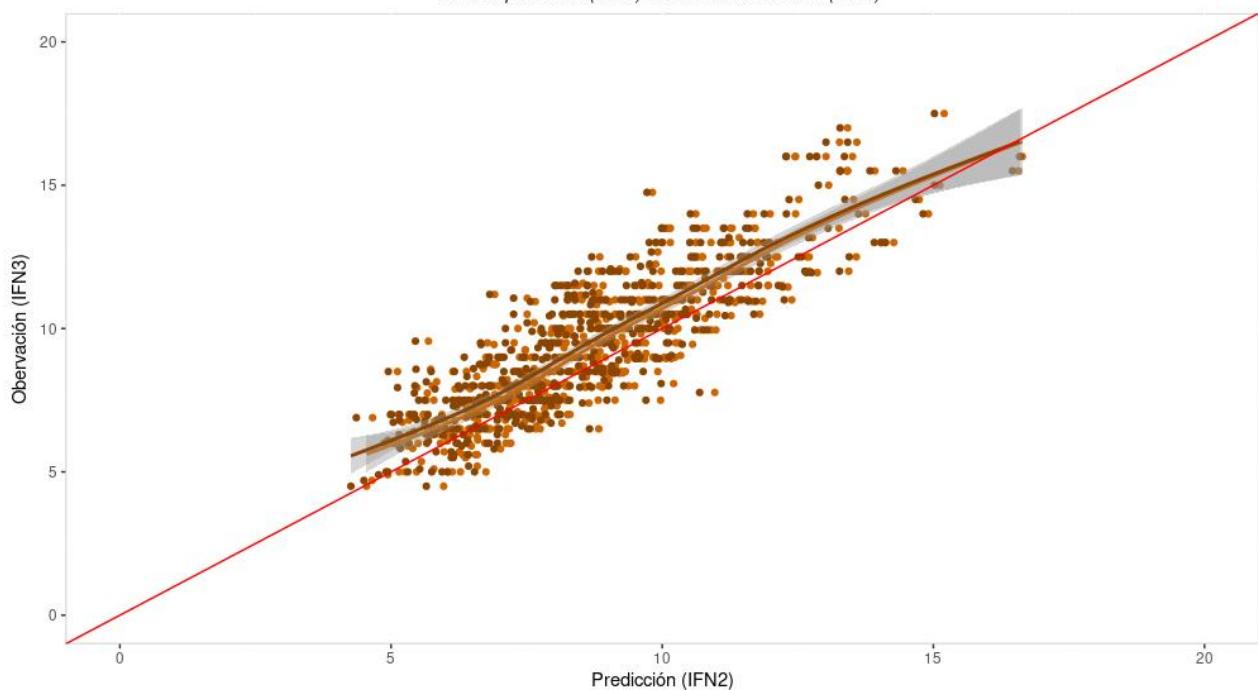


Figura 2. Gráficos de datos predichos (eje horizontal) y observados (eje vertical) generados con R para las variables de diámetro normal y altura del modelo IBEROPT e IBEROPS antes (colores claros) y después (colores oscuros) de su calibración.

Ecuaciones de supervivencia

Una vez calibradas las ecuaciones de crecimiento en diámetro y altura, y conociendo que el ajuste para la variable de densidad (N) no es satisfactorio (Tabla 3), se ha procedido a la evaluación de las ecuaciones de supervivencia estableciendo un valor umbral para su aplicación. Aplicando el criterio anterior sobre los modelos ensayados, los resultados obtenidos tras el análisis no muestran mejoras en la precisión de ajuste de la ecuación ni en el coeficiente de correlación entre valores observados y predichos, siendo ambos valores prácticamente iguales en los distintos casos de estudio planteados (Figura 3). Este resultado no permite sacar conclusiones claras, por lo que se pone de manifiesto la necesidad de continuar desarrollando la metodología para conocer cómo mejorar las predicciones.

Ecuaciones de masa incorporada

Tras realizar el análisis estadístico para cada uno de los valores umbral estudiados, hemos comprobado que el ajuste de la ecuación prácticamente no ha variado, al igual que el factor de precisión (Tabla 3, Figura 3), lo que supone que la modificación del valor umbral no es suficiente para mejorar las predicciones de la masa incorporada. No obstante, la eliminación de la masa incorporada técnica parece aumentar ligeramente la precisión de las predicciones de las variables N y G para ambos modelos (Figura 3), no encontrando mejoras notables para Ho (precisión alta del modelo original). De nuevo, y al igual que para el caso anterior, los resultados no permiten sacar conclusiones claras, por lo que es necesario continuar trabajando en este campo.

Tabla 3. Resultados obtenidos para el coeficiente de correlación muestral de Pearson (r), el coeficiente de correlación muestral de concordancia de Lin (r_c) y el factor de precisión (C_b) para los distintos casos de estudio más representativos de las ecuaciones de supervivencia y masa incorporada. Los valores destacados en negrita son los que mejor representan los cambios entre casos de estudio.

			IBERO ^{PT}			IBERO ^{PS}		
			N	G	Ho	N	G	Ho
modelo	r	0,63443	0,71958	0,79681	0,73423	0,79962	0,89247	
control	r_c	0,59384	0,59026	0,78693	0,56554	0,60565	0,88590	
calibrado	C_b	0,93602	0,82028	0,98760	0,77025	0,75742	0,99264	
supervivencia	r	0,64537	0,73055	0,90901	0,73342	0,80056	0,89499	
	r_c	0,60801	0,60758	0,90626	0,57580	0,61328	0,88730	
	C_b	0,94211	0,83167	0,99697	0,78509	0,76606	0,99141	
masa incorporada	r	0,60844	0,75246	0,88340	0,63553	0,78968	0,89844	
	r_c	0,52626	0,51482	0,86774	0,48410	0,52374	0,89562	
	C_b	0,86493	0,68418	0,98227	0,76173	0,66323	0,99686	
sin masa incorporada técnica	r	0,57210	0,71917	0,80355	0,63340	0,78816	0,89688	
	r_c	0,54384	0,58506	0,79353	0,53583	0,61580	0,89029	
	C_b	0,95060	0,81352	0,98753	0,84596	0,78131	0,99265	

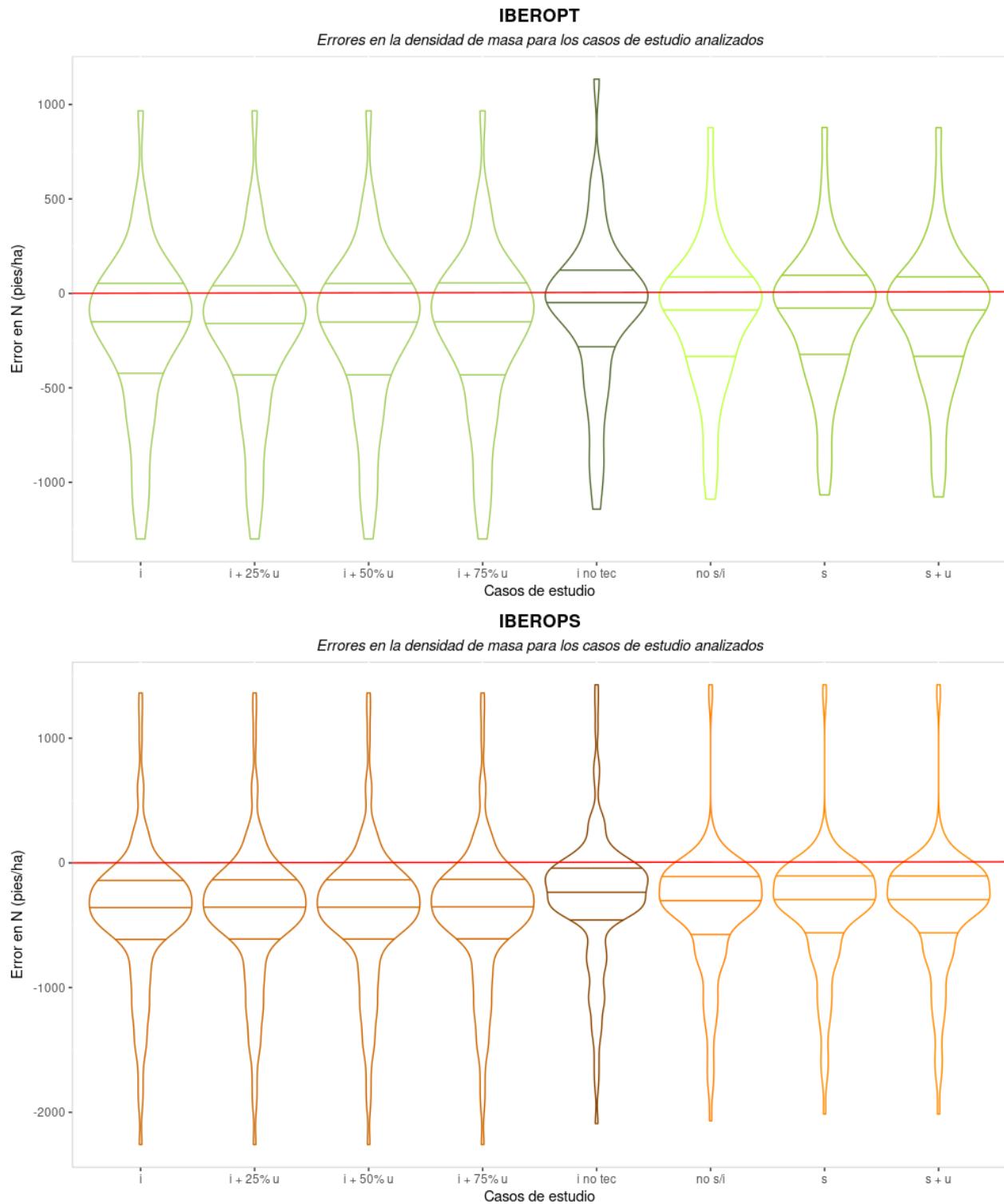


Figura 3. Gráficos de cajas para los errores entre densidad predicha y observada en los modelos IBEROPT (verde) e IBEROPS (naranja) a lo largo de los distintos casos de estudio: i (simulación empleando la ecuación de masa incorporada original); i + 25% u (simulación empleando la ecuación de masa incorporada con valor umbral 0.25); i + 50% u (simulación empleando la ecuación de masa incorporada con valor umbral 0.50); i + 75% u (simulación empleando la ecuación de masa incorporada con valor umbral 0.75); i no tec (simulación empleando la ecuación de masa incorporada sin contemplar la masa incorporada técnica); no s/i (simulación con el modelo control calibrado); s (simulación empleando la ecuación de supervivencia original); s + u (simulación empleando la ecuación de supervivencia tras añadir el valor umbral). Las diferentes tonalidades representan distintos grupos de casos de estudio.

5. Discusión

En base a los resultados obtenidos, podemos decir que la metodología desarrollada en este estudio ha mostrado una buena eficacia para la evaluación y validación conjunta de modelos. No obstante, este trabajo conforma una primera aproximación al desarrollo de la metodología, permitiendo estos resultados depurar la metodología hasta lograr su mejora.

Ecuaciones de crecimiento

Los resultados obtenidos para las ecuaciones de crecimiento han sido muy satisfactorios. La metodología utilizada nos ha permitido detectar errores de sesgo en las ecuaciones de crecimiento en diámetro y altura de ambos modelos, las cuales tendían a subestimar los crecimientos de dichas variables, semejante a lo encontrado por SCHMID et al. (2006) y PRETZSCH et al. (2002) para el simulador SILVA. Sabiendo esto, se ha procedido a la calibración en dos pasos, corrigiendo primero los sesgos de la ecuación de crecimiento en diámetro y posteriormente los sesgos de la ecuación de crecimiento en altura. La calibración de ambas ecuaciones ha supuesto una mejora sustancial en el valor del coeficiente de correlación de Lin (Tabla 2), pasando a encontrarse este valor en un grado de concordancia moderado para el caso del diámetro, y pobre (aunque con una mayor concordancia entre predicciones y observaciones) para el caso de las alturas. Pese a que el grado de concordancia no ha variado en el análisis de crecimiento en altura, es necesario recordar que la clasificación elaborada por CAMACHO-SANDOVAL (2008) ha sido pensada para estudiar la precisión de medida de aparatos de laboratorio, y no para la comparación de valores predichos por una ecuación frente a observados. Esto supone que la mejora sigue siendo interesante, aunque el grado de concordancia no varíe. Además, esta mejora se ve apoyada por incremento en el factor de precisión C_b , lo que indica una mayor adecuación de las predicciones hechas por las ecuaciones calibradas respecto a las originales. Este resultado pone de manifiesto que la técnica de cálculo de las ecuaciones es la adecuada, aunque la precisión de las predicciones puede ser mejorada. De hecho, las ecuaciones originales muestran un alto sesgo en las predicciones para aquellas clases diamétricas extremas en las que se contaba con menos datos en el momento de su creación (<10 y >40 cm), obteniendo una mayor precisión en las clases diamétricas restantes (>10 y <40 cm) (LIZARRALDE, 2008). La importancia de la calibración de estas ecuaciones es clave, dado que las variables de diámetro normal y altura se emplean como variables independientes para el cálculo de variables básicas como el área basimétrica, volumen o biomasa del árbol, además de suponer la base del cálculo de las variables de masa, por lo que su calibración supone una mejora sustancial de los modelos IBEROPT e IBEROPS, hecho que ya se había recalcado en la elaboración inicial de las mismas (LIZARRALDE, 2008). Además de la validación de las ecuaciones con datos de rodales extramaduros, como propone LIZARRALDE (2008), sería interesante hacer validaciones de las ecuaciones por clases diamétricas (VANCLAY & SKOVSGAARD, 1997), dividiendo así la muestra en estratos, pudiendo obtener como resultado variaciones de una misma ecuación principal que permitan mejorar la precisión de las predicciones para cada una de las clases diamétricas estudiadas.

Ecuaciones de supervivencia y masa incorporada

Tras un primer análisis de las predicciones para las variables de masa estudiadas (N , G y H_0), se ha encontrado un ajuste pobre para las variables N y G , con una precisión mayor para la variable H_0 , sin lograr un ajuste moderado para todos los casos de estudio (Tabla 3). Este resultado encontrado en el análisis inicial se debe a la forma en la que se calcula H_0 , dado que únicamente se seleccionan los 100 árboles más grandes para su estimación, sin tener en cuenta datos totales de la masa. Esto supone que la variable H_0 tenga mayor dependencia de las predicciones de altura (calibradas en el apartado anterior) que del efecto de las ecuaciones de supervivencia y masa incorporada (estudiadas en este apartado), y de ahí que sus predicciones sean mejores a lo largo de todos los casos de estudio.

Mientras que para el caso anterior los resultados obtenidos han permitido mejorar la precisión de las predicciones, la eficacia de los casos de estudio planteados no ha resultado satisfactoria para la validación de las ecuaciones de supervivencia y masa incorporada. En ambos casos, se ha estudiado la importancia de un factor umbral que determine cuándo han de aplicarse dichas ecuaciones, creando este valor umbral en base a la densidad de la masa para la ecuación de supervivencia, y modificando el ya existente para el caso de la masa incorporada. Los resultados muestran diferencias estadísticas ínfimas para ambos casos, lo que nos hace entender que la relevancia de este factor es pequeña. Aunque esta alteración de la ecuación de supervivencia, la cual incluye un valor umbral, está diseñada para ser aplicada en modelos dependientes de la distancia (BRAVO-OVIEDO et al., 2006), su aplicación sobre modelos independientes de la distancia podría resultar coherente. Los resultados de nuestro estudio muestran que la importancia de incorporar el valor umbral para determinar cuándo se aplica la ecuación de supervivencia es ínfimo. No obstante, BRAVO-OVIEDO et al. (2006) ha reportado, tras el desarrollo de dichas ecuaciones originales, buenos ajustes para la baja supervivencia de estas masas, con cierta sobreestimación para el modelo IBEROPT e infraestimación para el modelo IBEROPS, cuya precisión no hemos sido capaces de mejorar.

Por otro lado, las ecuaciones de masa incorporada diseñadas en dos pasos sí son de utilidad (BRAVO et al., 2008) aunque, a juzgar por nuestros resultados, la importancia del valor umbral que determina su aplicación parece no ser de gran relevancia. BRAVO-OVIEDO et al. (2006) reporta errores bajos para el modelo IBEROPT, siendo estos mayores para IBEROPS. Mientras que nuestro análisis no ha conseguido aumentar la precisión de las predicciones, sí demuestra que la importancia del valor umbral es menor que las ecuaciones probabilística y determinística, siendo necesario estudiar el comportamiento de ambas ecuaciones para mejorar la precisión de los modelos.

Una de las diferencias de peso en el estudio de estas ecuaciones respecto a las anteriores es el tipo de variables sobre las que se aplican. Mientras que las ecuaciones de crecimiento son aplicadas sobre variables de árbol, y por lo tanto su comparación puede realizarse de manera directa entre el valor predicho y el valor observado para cada árbol individual, las ecuaciones de supervivencia y masa incorporada se aplican sobre variables de masa, y tanto los errores como las carencias en su precisión es el resultado de la acumulación de los errores de cálculo individuales para cada uno de los árboles de la masa. Esto nos ayuda a explicar por qué existe una diferencia tan alta en los valores de los distintos estadísticos estudiados para las variables de masa (N , G y H_0) respecto a las variables de árbol (d y h) (ver Tabla 2 y Tabla 3, respectivamente). Una de las claves que nos ayudan a explicar la dificultad de obtener una alta precisión en las predicciones de estas ecuaciones es el manejo del factor de expansión. Esta variable, empleada para extrapolar los valores de una parcela forestal de superficie conocida a su valor por hectárea, supone una simplificación de la realidad que puede tener grandes variaciones dependiendo del caso de estudio. Mientras que su empleo está justificado desde el punto de vista logístico de la toma de datos, su uso supone una gran dificultad a la hora de realizar los análisis de variables de masa, dado que cada pequeña desviación en las predicciones individuales para cada árbol se multiplica por su propio factor de expansión, amplificándose y acumulándose los errores en los cálculos de las variables de masa.

De cara a nuevos trabajos en esta línea, por medio de este estudio se han detectado una serie de puntos de mejora:

- Variabilidad de datos: la base de datos utilizada está compuesta por parcelas de diversa naturaleza (plantaciones y masas naturales, masas regulares e irregulares...), lo que engloba una serie de casuísticas muy diferentes, introduce una mayor aleatoriedad en los resultados, y dificulta conseguir una mejora en la precisión de las predicciones. Una mayor homogeneidad de los datos permitiría analizar cómo funcionan las ecuaciones analizadas para cada caso en concreto, y hacer pequeñas modificaciones de las mismas dependiendo de la casuística de la parcela a analizar, por ejemplo, introduciendo una variable 'dummy' que determine si la masa es natural o plantación.
- Tamaño muestral: pese a que la gran cantidad de parcelas que componen el IFN, los requerimientos de los modelos reducen mucho la cantidad de parcelas útiles para su estudio, siendo el tamaño muestral de 49 parcelas para el modelo IBEROPT y de 136 parcelas para

IBERO^{PS}. Este es, sin duda, un punto en el que se debería mejorar para obtener resultados estadísticos más robustos. Este aspecto se podrá mejorar una vez se publique la siguiente edición del IFN.

- Edades de masa: mientras que la edad de la masa es una variable de gran importancia, en este estudio no hemos hecho ningún tipo de discriminación o agrupación de parcelas en función de esta variable. Esto podría darnos a conocer si las ecuaciones funcionan mejor para rangos de edades determinados, e incluso nos permitiría ajustar las ecuaciones a distintos rangos de edad de la masa para mejorar las predicciones, tal y como proponen VANCLAY & SKOVSGAARD (1997). No obstante, esto comprometería el tamaño muestral, que se vería aún más reducido para cada grupo de edad, resultando inviable en nuestro caso. El mismo razonamiento es aplicable a las clases diamétricas.
- Periodos de simulación: el empleo de los datos del IFN2 e IFN3 restringe la amplitud del estudio a un periodo de simulación de 10 años (tiempo transcurrido entre ambas ediciones del IFN). La posibilidad de realizar simulaciones con periodos de tiempo mayores nos permitiría conocer si en simulaciones consecutivas los datos predichos tenderían a asemejarse más a los observados o no, eliminando en cierto modo la aleatoriedad que pueda existir en un periodo de tiempo tan corto como el estudio. Esto será posible una vez estén disponibles los datos del IFN4, permitiendo hacer simulaciones de hasta 20 años (con datos del IFN2 e IFN4).
- Reelaboración de las ecuaciones: otra posible vía de mejora es la reelaboración de las ecuaciones de supervivencia y masa incorporada. Buscar un planteamiento común para la aplicación de ambas ecuaciones y rediseñarlas puede ser una opción viable.
- Mejora de la metodología: la metodología diseñada para este estudio ha resultado ser satisfactoria en cuanto a la mejora de las ecuaciones de crecimiento. No obstante, su diseño ha de ser perfeccionado en próximos trabajos para tratar de mejorar las predicciones de las ecuaciones de supervivencia y masa incorporada.

6. Conclusiones

Mediante este trabajo se ha diseñado una metodología para validar modelos forestales, consistente en comparar los valores predichos de una simulación con valores observados en campo. Esta metodología ha permitido evaluar los modelos IBEROPT e IBERO^{PS}, calibrando y validando sus ecuaciones de crecimiento en diámetro y altura, y realizando una primera aproximación a la validación de las ecuaciones de supervivencia y masa incorporada, de carácter más complejo. De este modo, las predicciones de los modelos originales IBEROPT e IBERO^{PS} implementados en SIMANFOR han sido mejoradas, tanto para las variables de diámetro y altura como para las restantes variables de árbol, calculadas a partir de ellas. Este avance supone el primer paso para validar la totalidad del modelo. Futuros trabajos deberán abordar la mejora de las ecuaciones de supervivencia y masa incorporada para poder realizar la validación total de modelos forestales.

7. Agradecimientos

Esta investigación ha hecho uso de los recursos del Centro de Supercomputación de Castilla y León (SCAYLE, www.scayle.es), financiada por el Fondo Europeo de Desarrollo Regional (FEDER). Además, se ha desarrollado parcialmente dentro del marco del proyecto Cross-Forest, cofinanciado por la European Union's Innovation and Networks Executive Agency (INEA), mediante Connecting Europe Facility (CEF) 2014-2020.

8. Bibliografía

AEMET; 2021. AEMET: Agencia Estatal de Meteorología. www.aemet.es/

ALBERDI, I.; SANDOVAL, V.; CONDES, S.; CAÑELLAS, I.; VALLEJO, R.; 2016. El inventario Forestal Nacional español, una herramienta para el conocimiento, la gestión y la conservación de los ecosistemas forestales arbolados. *Revista Ecosistemas*, 25(3), 88-97.

BRAVO, F.; DEL RÍO, M.; DEL PESO, C.; 2002. El inventario forestal nacional: elemento clave para la gestión forestal sostenible. *Fundación General de la Universidad de Valladolid*. Grafistaff.

BRAVO, F.; MONTERO, G.; 2001. Site index estimation in Scots pine (*Pinus sylvestris* L.) stands in the High Ebro Basin (northern Spain) using soil attributes. *Forestry*, 74(4), 395-406. <https://doi.org/https://doi.org/10.1093/forestry/74.4.395>

BRAVO, F.; PANDO, V.; ORDÓÑEZ, C.; LIZARRALDE, I.; 2008. Modelling ingrowth in Mediterranean pine forests: a case study from scots pine (*Pinus sylvestris* L.) and mediterranean maritime pine (*Pinus pinaster* Ait.) stands in Spain. *Investigación agraria. Sistemas y recursos forestales*. 17(3), 250-260.

BRAVO, F.; RODRÍGUEZ, F.; ORDÓÑEZ, C.; BROTO, M.; LIZARRALDE, I.; RUANO, I.; DEL RÍO, M.; CALAMA, R.; 2010. Simanfor: aplicación web para la simulación de alternativas selvícolas. *Divulgación 1er. Trimestre*, (100).

BRAVO, F.; ÁLVAREZ GONZÁLEZ, J. G.; DEL RÍO, M.; BARRIO, M.; BONET LLEDOS, J. A.; BRAVO OVIEDO, A.; CALAMA, R.; CASTEDO DORADO, F.; CRECENTE CAMPO, F.; CONDES, S.; ... 2011. Growth and yield models in Spain: historical overview, contemporary examples and perspectives. *Forest Systems*. 20(2). 315-328.

BRAVO, F.; RODRÍGUEZ, F.; ORDÓÑEZ, C.; 2012. A web-based application to simulate alternatives for sustainable forest management: SIMANFOR. *Forest systems*, 21(1), 4-8.

BRAVO-OVIEDO, A.; DEL RÍO, M.; MONTERO, G.; 2004. Site index curves and growth model for Mediterranean maritime pine (*Pinus pinaster* Ait.) in Spain. *Forest Ecology and Management*, 201(2-3), 187-197. <https://doi.org/https://doi.org/10.1016/j.foreco.2004.06.031>

BRAVO-OVIEDO, A.; STERBA, H.; DEL RÍO, M.; BRAVO, F.; 2006. Competition-induced mortality for Mediterranean *Pinus pinaster* Ait. and *P. sylvestris* L. *Forest Ecology and Management*, 222(1-3), 88-98.

CAMACHO-SANDOVAL, J.; 2008. Coeficiente de concordancia para variables continuas. *Acta Médica Costarricense*, 50(4), 211-212.

DEL RÍO, M.; LÓPEZ, E.; MONTERO, G.; 2006. Manual de gestión para masas procedentes de repoblación de *Pinus pinaster* Ait., *Pinus sylvestris* L. y *Pinus nigra* Arn. en Castilla y León. Junta de Castilla y León, Castilla y León (España).

LAWRENCE, I.; LIN, K.; 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 255-268.

LIN, K.; 2000. A note on the concordance correlation coefficient. *Biometrics*, 56, 324-325.

LIZARRALDE, I.; ORDÓÑEZ, C.; BRAVO-OVIEDO, A.; BRAVO, F. 2010a. IBEROPT: Modelo de dinámica de rodales de *Pinus pinaster* Ait. en el sistema ibérico meridional.

LIZARRALDE, I.; ORDÓÑEZ, C.; BRAVO-OVIEDO, A.; BRAVO, F. 2010b. IBEROPS: Modelo de dinámica de rodales de *Pinus sylvestris* L. en el sistema ibérico meridional.

LIZARRALDE, I.; 2008. Dinámica de rodales y competencia en las masas de pino silvestre (*Pinus sylvestris* L.) y Pino Negral (*Pinus pinaster* Ait.) de los Sistemas Central e Ibérico Meridional (Tesis doctoral). Departamento de Producción Vegetal y Recursos Forestales.

MARTONNE, E.; 1926. L'indice d'aridité. *Bulletin de l'Association de géographes français*, 3(9), 3-5.

MOORE, J. R. (2010). Allometric equations to predict the total above-ground biomass of radiata pine trees. *Annals of Forest Science*, 67(8), 806.

PIÑEIRO, G.; PERELMAN, S.; GUERSCHMAN, J. P.; PARUELO, J. M.; 2008. How to evaluate models: observed vs. predicted or predicted vs. observed? *Ecological modelling*, 216(3-4), 316-322.

PRETZSCH, H.; 2009. Forest dynamics, growth, and yield. En: *Forest dynamics, growth and yield* (pp. 1-39). Springer, Berlin, Heidelberg.

PRETZSCH, H.; BIBER, P.; DURSK, J.; 2002. The single tree-based stand simulator SILVA: construction, application and evaluation. *Forest ecology and management*, 162(1), 3-21.

R CORE TEAM; 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

REINEKE, L. H.; 1933. Perfection a stand-density index for even-aged forest. *Journal of agricultural research*, 46, 627-638.

RODRÍGUEZ, D.; SAN MARTÍN, R.; BRAVO, F.; HERRERO, C.; 2020. Potential climatic influence on maximum stand carrying capacity for 15 Mediterranean coniferous and broadleaf species. *Forest Ecology and Management*, 460, 117824.

ROXBURGH, S. H.; KARUNARATNE, S. B.; PAUL, K. I.; LUCAS, R. M.; ARMSTON, J. D.; SUN, J.; 2019. A revised above-ground maximum biomass layer for the Australian continent. *Forest Ecology and Management*, 432, 264-275.

SAS INSTITUTE; 2011. SAS/STAT 9.3 User's Guide: Survey Data Analysis (book Excerpt). SAS institute.

SCAYLE; 2019. Supercomputación Castilla y León [Accedido 09-05-2021].

SCHMID, S.; ZINGG, A.; BIBER, P.; BUGMANN, H.; 2006. Evaluation of the forest growth model SILVA along an elevational gradient in Switzerland. *European Journal of Forest Research*, 125(1), 43-55.

SECF; 2021. Glosario Técnico Forestal de la Sociedad Español de Ciencias Forestales. <http://secforestales.org/content/glosario-tecnico-forestal>

SIMANFOR; 2021. SIMANFOR - Sistema de Apoyo para la Simulación de Alternativas de Manejo Forestal Sostenible. <https://www.simanfor.es/>

VANCLAY, J. K.; 1994. Modelling forest growth and yield: applications to mixed tropical forests. *CAB international*.

VANCLAY, J. K.; SKOVSGAARD, J.; 1997. Evaluating forest growth models. *Ecological Modelling*, 98(1), 1-12.

VÁZQUEZ-VELOSO, A.; 2021. Evaluación y validación de los modelos de crecimiento forestal IBEROPT e IBEROPS. *Universidad de Valladolid*. Palencia.

WEISKITTEL, A. R.; HANN, D. W.; KERSHAW JR, J. A.; VANCLAY, J. K.; 2011. Forest growth and yield modeling. John Wiley & Sons.