



UNIVERSIDAD DE VALLADOLID

MÁSTER EN
GESTIÓN FORESTAL BASADA EN CIENCIA DE DATOS
(DATAFOREST)

Evaluación y validación de los modelos de crecimiento forestal IBERO^{PT} e IBERO^{PS}

Autor: Aitor Vázquez Veloso

Tutores: Felipe Bravo Oviedo - Valentín Pando Fernández

Junio de 2021



Escuela Técnica Superior
de Ingenierías Agrarias **Palencia**

Índice general

1. Introducción	1
2. Objetivos	2
3. Material y métodos	3
3.1. Datos	3
3.2. Simulaciones de crecimiento	4
3.3. Modelos empleados	5
3.3.1. Ecuaciones de crecimiento	5
3.3.2. Ecuaciones de mortalidad	6
3.3.3. Ecuaciones de masa incorporada	7
3.4. Análisis estadístico	9
4. Resultados	11
4.1. Ecuaciones de crecimiento en diámetro y altura	11
4.2. Ecuaciones de mortalidad	13
4.3. Ecuaciones de masa incorporada	13
5. Discusión	16
5.1. Ecuaciones de crecimiento	16
5.2. Ecuaciones de mortalidad y masa incorporada	17
6. Conclusiones	21
7. Agradecimientos	22

Resumen del TFM

La modelización forestal es una herramienta que nos permite no solo predecir cómo evolucionan las masa forestales bajo diferentes escenarios silvícolas, sino también explicar el estado de la masa forestal bajo determinadas condiciones, ayudando a los gestores en el proceso de toma de decisiones. En este campo se han desarrollado ecuaciones para el cálculo de diferentes variables de interés forestal a nivel de árbol, clases de tamaño y masa, cuya precisión en las predicciones es comúnmente evaluada de manera individual. Para el empleo de las mismas, modelos que agrupan ecuaciones y simuladores que facilitan su utilización han sido desarrollados, 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 conjunto 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 conjunta de modelos. Para ello, se tratan de validar los modelos $IBERO^{PT}$ e $IBERO^{PS}$ 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, mientras que detecta carencias en las ecuaciones de mortalidad y masa incorporada cuyas estimaciones aún no se han podido mejorar.

Palabras clave

desarrollo de metodología, mortalidad, masa incorporada, SIMANFOR, modelización, selvicultura

Abstract

Forest modeling is a tool that allows us not only to predict how forest stands evolve under different silviculture scenarios, also to explain the status of forest stands under different conditions, helping the managers to take decisions. Equations to calculate different forestry variables were developed at the tree, size classes and stand levels, which accuracy on predictions was commonly evaluated individually. To use them, models that group equations and simulators that facilitate their use were also developed, centralizing the information and simplifying the use. However, although individual equations validation is a well-known field, the joint validation of the set of equations and rules that make up the models is an underexplored topic. This study is the first attempt to the development of a methodology that allows us to validate full models. To do it, we treat to validate $IBERO^{PT}$ and $IBERO^{PS}$ models, developed to *Pinus pinaster mesogeensis* and *Pinus sylvestris*, respectively, by using the data of the Spanish National Forest Inventory (SNFI, IFN on Spanish) and SIMANFOR platform to their performance. Our results demonstrate the adequation of the developed methodology in order to improve the accuracy of the growing equations, while it detects deficiencies in the mortality and ingrowth equations whose estimates have not been able to be improved yet.

Keywords

methodology development, mortality, ingrowth, SIMANFOR, modeling, silviculture

Capítulo 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 los resultados de los mismos (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 y col., 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 y col., 2010; Bravo y col., 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 y col., 2016), 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 para su desarrollo 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 y col., 2011), lo cierto es que todavía queda mucho trabajo por hacer para desarrollarlo. 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 conjuntos, que implica la validación conjunta de todas las ecuaciones desarrolladas para una determinada especie y lugar. Dado que esto es un campo novedoso, existe una importante falta de conocimiento acerca de como realizar una validación conjunta de modelos, incluyendo tres partes fundamentales del mismo: crecimiento, mortalidad y masa incorporada.

El propósito de este estudio es sentar las bases para el desarrollo de una primera metodología que permita la validación conjunta de modelos. Para ello, hemos tomado los modelos IBERO^{PT} (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) e IBERO^{PS} (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010), modelos desarrollados para *Pinus pinaster* y *Pinus sylvestris*, como objeto de evaluación. Su empleo se ha realizado a través de la plataforma SIMANFOR (Bravo y col., 2010; Bravo y col., 2012; SIMANFOR, 2021), 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.

Capítulo 2

Objetivos

Los objetivos principales de este estudio son:

1. Desarrollar una metodología que permita la validación conjunta de modelos, incluyendo sus ecuaciones de crecimiento en diámetro y altura, mortalidad y masa incorporada
2. Sistematizar el empleo de la metodología aquí empleada para replicarla con otros modelos empleados en SIMANFOR
3. Analizar la precisión de las predicciones de los modelos $IBERO^{PT}$ e $IBERO^{PS}$, corrigiendo y validando sus ecuaciones o determinando cuales han de ser analizadas bajo otra metodología

Capítulo 3

Material y métodos

3.1. Datos

Con el propósito de validar los modelos de crecimiento, se han utilizado los datos de parcelas forestales permanentes de las ediciones 2 y 3 del Inventario Forestal Nacional Español (IFN), con el objetivo de utilizar 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). Dichas parcelas consisten en cuatro círculos concéntricos de radios 5, 10, 15 y 25 metros, para los cuales fueron medidas múltiples variables de todos los árboles cuyo diámetro a la altura del pecho (1.30 m) supere 7,5, 12,5, 22,5 y 42,5 cm, respectivamente (Alberdi y col., 2016). Los datos utilizados se corresponden con la información de las parcelas y árboles, no siendo necesaria información cartográfica para nuestro propósito. Además, constituyen una base de datos de parcelas permanentes diferente a la utilizada para el desarrollo de los modelos, lo cual es fundamental para un proceso de evaluación (Vanclay, 1994). Tras un preparado 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} (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) e IBERO^{PS} (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010), objeto de validación:

- que la parcela haya sido medida en ambos inventarios, dado que en algunos casos la parcela ha podido desaparecer o sufrir perturbaciones que eliminen la superficie forestal
- 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
- seleccionar las parcelas localizadas en el rango geográfico para el que fueron diseñadas las ecuaciones. Para ambos casos, dicho rango geográfico se reduce al Sistema Ibérico Meridional, comprendiendo las siguientes 4 provincias: Soria, Guadalajara, Cuenca y Teruel

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 2810 árboles *P. sylvestris*, respectivamente, que cumplan dichos criterios.

Conocidas las parcelas que cumplen dichos criterios, se han recalculado aquellas parcelas del IFN3 cuyo tiempo transcurrido entre las mediciones del IFN2 e IFN3 haya sido diferente a 10 años, recalculando los crecimientos en altura y diámetro de los árboles, de manera lineal, a la nueva edad

de la parcela (10 años después de ser medida en el IFN2), permitiendo así hacer coincidir en el tiempo los datos predichos (mediante dos simulaciones de 5 años del IFN2) con los observados (datos del IFN3) para su comparación. Por otro lado, se han utilizado datos de precipitación (mm/año) y temperatura media anual (°C/año) obtenidos de la Estación Española de Meteorología (AEMET, 2021) para evaluar una posible mejora en la aplicación de la ecuación de mortalidad, mediante la creación de un valor de umbral para determinar cuándo es necesaria su aplicación y cuando no, basado en la densidad de la masa.

3.2. Simulaciones de crecimiento

Para las simulaciones del crecimiento se ha utilizado el software SIMANFOR (Bravo y col., 2010; Bravo y col., 2012; SIMANFOR, 2021), utilizando además los modelos $IBERO^{PT}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) e $IBERO^{PS}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) programados para ser utilizados con dicha herramienta. Tras adaptar ligeramente el simulador y los modelos a las necesidades de estas simulaciones, estas 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 ha sido necesario modificar los modelos y repetir las simulaciones, siguiendo el diagrama de flujo general mostrado en la Figura 3.1 y teniendo en cuenta todas las variantes planteadas para el análisis de cada una de las ecuaciones.

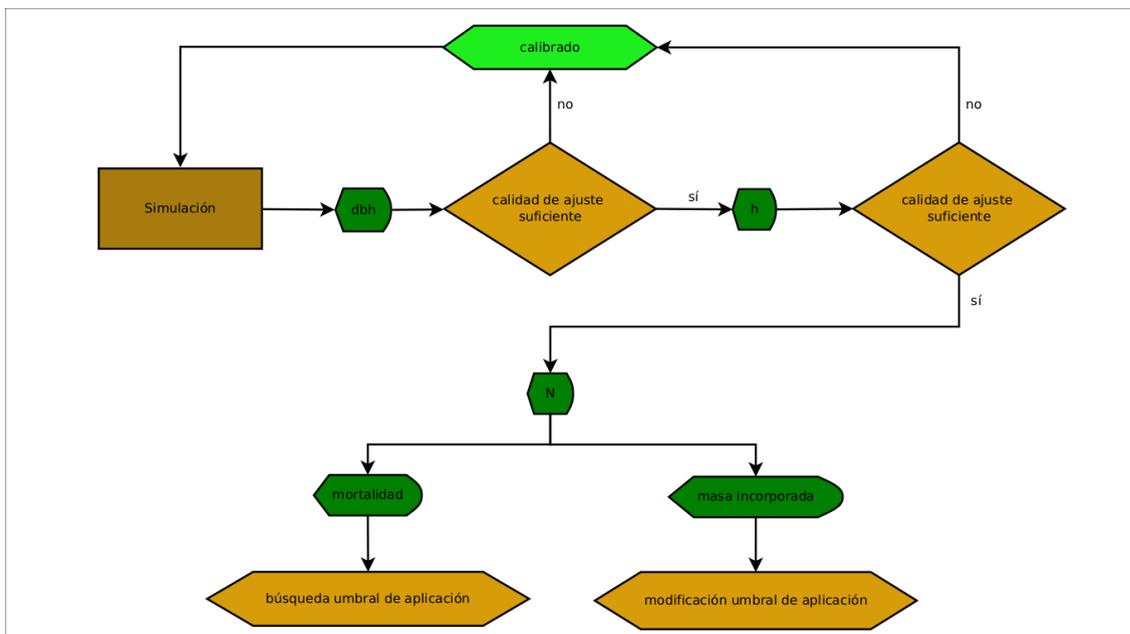


Figura 3.1: Procedimiento empleado para la validación de los modelos $IBERO^{PT}$ e $IBERO^{PS}$

Para cada simulación, los escenarios empleados han consistido en una proyección del crecimiento de 10 años para las parcelas correspondientes al IFN2 (divididas en dos proyecciones de 5 años cada una, siendo este el tiempo de proyección para el que se diseñaron las ecuaciones de los modelos empleados). Para los datos del IFN3 se ha realizado un recálculo de las parcelas,

consistente en realizar un proceso de inicialización en SIMANFOR sin proyecciones en el tiempo, para así poder aplicar el cálculo de todas aquellas variables que permite el simulador (copa, biomasa, volumen...). Este procedimiento nos permite tener las mismas variables disponibles en ambos inventarios, calculadas por el mismo simulador y con las mismas ecuaciones para su posterior comparación, evitando así errores relacionados con diferencias en la transcripción de las ecuaciones o con la precisión de cálculo de distintas plataformas.

Tras realizar las simulaciones, los datos de salida han sido procesados con R (R Core Team, 2019), con el objetivo de unir los resultados de la simulación para cada una de las parcelas en un único archivo, dotándolos de una mejor estructura de cara a su análisis. Tras este proceso, se obtiene para cada caso de estudio 2 archivos diferentes, uno para árboles y otro para parcelas, englobando los datos predichos (IFN2) y observados (IFN3) de cada uno de ellos, con los que se realizan los análisis estadísticos descritos posteriormente.

3.3. Modelos empleados

Los modelos utilizados con objetivo de ser calibrado son $IBERO^{PT}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) e $IBERO^{PS}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010), ambos modelos de crecimiento de árbol individual independientes de la distancia, programados en SIMANFOR (Bravo y col., 2010; Bravo y col., 2012; SIMANFOR, 2021) y adaptados ligeramente a las necesidades de este estudio para cada una de las ecuaciones a analizar, tal y como se describe en el apartado siguiente. Las ecuaciones de ambos modelos fueron desarrolladas para el Sistema Ibérico Meridional en el caso de $IBERO^{PT}$, comprendiendo las provincias de Soria, Guadalajara, Cuenca y Teruel, y para los Sistemas Ibérico y Central en el caso de $IBERO^{PS}$, 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 del árbol 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 por hectárea). De este modo, nuestro estudio se ha centrado en el análisis de las ecuaciones detalladas a continuación.

3.3.1. Ecuaciones de crecimiento

Ambos modelos cuentan con sus propias ecuaciones de crecimiento en diámetro y en altura, las cuales se aplican una a continuación de la otra, siendo el valores resultante del crecimiento en diámetro una de las variables independientes para el cálculo del crecimiento en altura.

Las ecuaciones de crecimiento en diámetro 3.1 y 3.2 (Lizarralde, 2008) correspondientes a los modelos $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente, se muestran a continuación:

$$d_{i5} = e^{0,2030*\ln(d*10)+0,4414*\ln(\frac{cr+0,2}{1,2})+0,8379*\ln(SI)-0,1295*\sqrt{G}-\frac{0,0007*bal^2}{\ln(d*10)}} / 10 \quad (3.1)$$

$$d_{i5} = e^{-0,37110+0,2525*\ln(d*10)+0,7090*\ln(\frac{cr+0,2}{1,2})+0,9087*\ln(SI)-0,1545*\sqrt{G}-0,0004*\frac{bal^2}{\ln(d*10)}} / 10 \quad (3.2)$$

Donde:

- d_{i5} es el incremento del diámetro normal para un periodo de 5 años, en cm
- d es el diámetro normal, medido a la altura del pecho (1,30 m), en cm
- cr es el índice de copa viva, en %

- bal es el área basimétrica acumulada de los árboles más gruesos que el árbol sujeto, en m^2/ha
- G es el área basimétrica de la masa, en m^2/ha
- SI es el Índice de Sitio, en m

Por otro lado, las ecuaciones de crecimiento en altura 3.3 y 3.4 (Lizarralde, 2008) correspondientes a los modelos $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente, son las siguientes:

$$h_{i5} = e^{0,21603+0,40329*\ln(\frac{d_{i5}}{2})-1,12721*\ln(d*10)+1,18099*\ln(h*100)+3,01622*cr}/100 \quad (3.3)$$

$$h_{i5} = e^{3,1222-0,4939*\ln(d_{i5}*10)+1,3763*\ln SI-0,0061*bal+0,1876*\ln(cr)}/100 \quad (3.4)$$

Donde:

- $hi5$ es el incremento de la altura total para un periodo de 5 años, en m
- h es la altura total del árbol en m

Para el estudio de ambas ecuaciones de crecimiento se han utilizado tanto los valores absolutos de diámetro normal y altura total del árbol como el incremento en dichos valores para el periodo de 10 años transcurridos entre las dos mediciones del IFN. Tras realizar algunas pruebas para probar la viabilidad de ambos planteamientos, se ha decidido estudiar ambas variables utilizando tan solo los valores absolutos de las mismas, dado que las diferencias entre observados y predichos son iguales para ambos casos, mientras que la magnitud de la variable a estudiar nos permite entender mejor en qué rangos de diámetro o altura funciona mejor cada ecuación.

3.3.2. Ecuaciones de mortalidad

Las ecuaciones de mortalidad son aquellas que actúan sobre el factor de expansión de los árboles, reduciéndolo en función de las características que presenta cada árbol respecto a las características de la masa en la que se encuentra. A nivel de masa, la aplicación de esta ecuación supone un descenso en la densidad (N , medida en n° pies/ha), resultante de la aplicación de esta ecuación sobre todos los árboles de la misma.

Las ecuaciones de mortalidad 3.5 y 3.7 utilizada con 3.6 (Bravo-Oviedo y col., 2006) correspondientes a los modelos $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente, se muestran a continuación:

$$P = 1 - \frac{1}{1 + e^{(2,0968 + \frac{4,7358*d}{dg} - 0,0012*SI*G)}} \quad (3.5)$$

$$P = \frac{1}{1 + e^{(-6,8548 + \frac{9,792}{d} + 0,121*bal*cvd + 0,037*SI)}} \quad (3.6)$$

$$cvd = \frac{\sqrt{dg^2 - \bar{d}^2}}{\bar{d}} \quad (3.7)$$

Donde:

- P es la probabilidad de mortalidad, pudiendo tomar valores comprendidos entre 0 (muerto) y 1 (vivo)
- cvd es el coeficiente de variación de la distribución diamétrica

- dg es el diámetro medio cuadrático de la masa, en cm
- \bar{d} es el diámetro medio aritmético de la masa, en cm

De cara a mejorar las predicciones de estas ecuaciones, en este estudio se ha evaluado la posibilidad de establecer un valor de umbral para determinar cuando es necesaria su aplicación y cuando no. Para ello, se ha utilizado la metodología propuesta por Rodríguez de Prado y col. (2020), donde la mortalidad depende de la capacidad máxima de carga del medio. Siguiendo esta metodología, en cada iteración se calculan el Índice de Densidad de la Masa o Índice de Reineke, conocido como SDI (Reineke, 1933), que define la densidad actual de la masa, y el Índice de Densidad Máxima de la Masa, SDI_{max}, que define la densidad máxima que la masa puede soportar en nº de árboles por hectárea. En los modelos utilizados, se emplean los coeficientes específicos para la ecuación SDI propuestos por del Río y col. (2006) para *Pinus pinaster* y *Pinus sylvestris* (ecuaciones 3.8 y 3.9, respectivamente). Por otro lado, las ecuaciones utilizadas para el cálculo de SDI_{max} han sido extraídas de Rodríguez de Prado y col. (2020) (ecuaciones 3.10 y 3.11, respectivamente), y utilizan en su cálculo el Índice de aridez de Martonne (Martonne, 1926), cuyo cálculo (ecuación 3.12) utiliza variables climáticas promedio de la zona de estudio.

$$SDI = N * \left(\frac{25}{dg}\right)^{-1,629} \quad (3.8)$$

$$SDI = N * \left(\frac{25}{dg}\right)^{-1,75} \quad (3.9)$$

$$SDI_{max} = e^{11,641+0,489 \ln(M)+(-1,8491-0,0035*M)*\ln(25)} \quad (3.10)$$

$$SDI_{max} = e^{(12,579+(-1,7462+0,0007*M)*\ln(25))} \quad (3.11)$$

$$M = \frac{Pt}{Tm + 10} \quad (3.12)$$

Donde:

- M es el Índice de aridez de Martonne
- Pt es la precipitación total anual, en mm
- Tm es la temperatura media anual, en °C

Calculados ambos valores en cada proyección (se realizan dos proyecciones de 5 años con los datos de IFN2), la ecuación de mortalidad se activa tan solo cuando los valores de SDI >SDI_{max} para una determinada parcela, actuando sobre el factor de expansión de cada árbol y reduciendo la densidad de la masa por debajo del valor SDI_{max} precalculado.

3.3.3. Ecuaciones de masa incorporada

Al contrario de las ecuaciones de mortalidad, las ecuaciones de masa incorporada son aquellas que contemplan la incorporación de árboles a la masa que en el momento anterior no cumplían los requisitos de diámetro mínimo para ser inventariados. Resulta importante destacar que las ecuaciones de masa incorporada incluidas en estos modelos (Bravo y col., 2008) han sido desarrolladas

con una estructura semejante para la recogida de datos que el IFN, utilizando parcelas circulares concéntricas de radios 5, 10 y 15 metros, en las que varían los diámetros mínimos inventariables de los árboles dependiendo de la proximidad al centro de la parcela. Esto restringe su aplicación a aquellos casos en los que se utilicen datos provenientes del IFN, o bien de inventarios que hayan seguido una metodología semejante, mientras que su aplicación para inventarios que no sigan esta metodología todavía no ha sido evaluada.

Por otro lado, dado que tanto la ecuación de mortalidad como para la ecuación de masa incorporada actúan sobre la misma variable de masa (N), para ambos casos de estudio se ha decidido silenciar la otra ecuación, evitando así que los datos de salida se vean alterados por ambas funciones. Este proceso de silenciar nos permite 'desactivar' la ecuación que no estamos evaluando, anulando su efecto y permitiendo así un estudio independiente y más preciso de cada ecuación por separado. A efectos de la simulación, el proceso de silenciado consiste en eliminar la orden que determina que debe de aplicarse la ecuación de mortalidad o masa incorporada, evitando así que entre en funcionamiento.

Estas ecuaciones se componen de tres procesos:

- el primer paso está compuesto por una ecuación de probabilidad con un valor umbral, a partir del cual se decide si existirá masa incorporada en la parcela de estudio (ecuaciones 3.13 y 3.15 para $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente)
- el segundo paso consta de una ecuación para el cálculo de la cantidad de masa incorporada a la parcela en área basimétrica (G, medida en m²/ha), activándose solamente cuando el valor de probabilidad anterior supera el valor umbral establecido para la especie estudiada (ecuaciones 3.14 y 3.16 y valores umbral de 0,38 y 0,43 para $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente). Este dato es incorporado a nivel de árbol aumentando su factor de expansión, repercutiendo directamente sobre la densidad de la masa (N, medida en n° pies/ha).
- un tercer paso que no interviene en los análisis realizados es la repartición de la masa incorporada en distintas clases diamétricas, acorde con los distintos radios de muestreo que se emplean en el IFN para la toma de datos, cuya repartición no es evaluada en este estudio

$$Pmi = \frac{1}{1 + e^{-(12,3424+0,1108*G-0,6154*dg)}} \quad (3.13)$$

$$G_{incorporada} = 6,7389 - 0,2235 * dg \quad (3.14)$$

$$Pmi = \frac{1}{1 + e^{-(8,2739-0,3022*dg)}} \quad (3.15)$$

$$G_{incorporada} = 5,7855 - 0,1703 * dg \quad (3.16)$$

Donde:

- Pmi es la probabilidad, entre 0 y 1, de que exista masa incorporada
- $G_{incorporada}$ es la cantidad de masa incorporada, en área basimétrica, por parcela (m²/ha)

Dada la complejidad que representa el análisis de estas ecuaciones, se ha decidido analizar como influye sobre la precisión de sus cálculos la aplicación de la 'masa incorporada técnica', 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. Para este último caso, se han repetido las simulaciones estudiando los ajustes

obtenidos para los valores umbral 0,25, 0,50 y 0,75, seleccionados como valores ‘guía’ para conocer la fluctuación en el ajuste, y así determinar si dicho cambio es suficiente para tener una mayor precisión en las predicciones respecto a los valores umbral originales (0.38 para IBERO^{PT} y 0.43 para IBERO^{PS}).

Para la comparación de las ecuaciones de mortalidad y masa incorporada se han utilizado simulaciones control en las que se ha silenciado el efecto de ambas ecuaciones, es decir, la variable de densidad (N) no ha sido alterada en ningún momento. De este modo, podremos conocer qué combinación de ecuaciones (ninguna, solo mortalidad, solo masa incorporada, mortalidad + masa incorporada) supone un mejor ajuste para cada uno de los modelos estudiados sobre la densidad de la masa.

3.4. Análisis estadístico

Para el análisis estadístico se ha utilizado el software SAS/STAT (Versión 3.8 (Edición Enterprise); SAS Institute, 2011), por el cual se hace una comparación de los datos observados (IFN3) con los predichos (IFN2). Mientras que este mismo procedimiento ha sido empleado por estudios enfocados en evaluación de modelos (Moore, 2010; Roxburgh y col., 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).

Dada una muestra aleatoria simple bivalente $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ de un vector aleatorio (X, Y) en R^2 se define el coeficiente de correlación muestral de concordancia de Lin r_c como:

$$r_c = \frac{2s_{xy}}{s_x^2 + s_y^2 + (\bar{x} - \bar{y})^2} \quad (3.17)$$

donde:

- s_{xy} es la covarianza muestral
- s_x^2 y s_y^2 son las varianzas muestrales
- \bar{x} e \bar{y} son las dos medias muestrales de las dos variables aleatorias que componen el vector

Se puede demostrar que, si r es el coeficiente de correlación muestral de Pearson, entonces se verifica que

$$r_c = rC_b \quad (3.18)$$

siendo

$$C_b = \left[\left(v + 1/v + u^2 \right) / 2 \right]^{-1} \quad (3.19)$$

con

$$u = s_x / s_y \quad (3.20)$$

y

$$u = (\bar{x} - \bar{y}) / \sqrt{s_x s_y} \quad (3.21)$$

Además, se verifica que $0 < C_b \leq 1$, siendo C_b el factor de precisión, y por tanto $|r_c| \leq |r|$, es decir, el coeficiente de correlación de concordancia de Lin es siempre menor igual, en módulo, que el coeficiente de correlación de Pearson. El coeficiente de correlación de Pearson r es una medida de precisión de la recta de ajuste a los valores de las dos variables y el parámetro C_b es una medida adecuación de la recta de ajuste a la bisectriz del primer cuadrante.

El valor del coeficiente de Lin se relaciona con un grado de concordancia entre las predicciones y las observaciones, tal y como se muestra en el Cuadro 3.1 propuesto por Camacho-Sandoval (2008), el cual ha sido desarrollado 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 para cada ecuación conseguir un grado de concordancia moderado o lo más cerca posible a este. Este proceso se realiza corrigiendo los problemas de sesgo, ajuste, o ambos detectados en las ecuaciones, en cuyo caso será necesario proceder al calibrado o reelaboración de las ecuaciones para obtener una mejor precisión. 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 (Ho) para los datos de parcelas, siendo elegidas como las variables más representativas de las ecuaciones a analizar.

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

Cuadro 3.1: Clasificación de los grados de concordancia para cada rango de valor del coeficiente de Lin (Camacho-Sandoval, 2008)

Capítulo 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 mortalidad y masa incorporadas no han supuesto mejoras respecto al modelo original.

4.1. Ecuaciones de crecimiento en diámetro y altura

Tras una primera simulación de referencia en la que se han utilizado los modelos $IBERO^{PT}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) e $IBERO^{PS}$ (Lizarralde, Ordóñez, Bravo-Oviedo y col., 2010) originales, han sido detectados errores de sesgo para las ecuaciones de crecimiento en diámetro (ecuaciones 3.1 y 3.2) y altura (ecuaciones 3.3 y 3.4) de ambos modelos, mientras que los ajustes de las ecuaciones han satisfactorios. Para su corrección, se ha realizado una calibración conjunta de ambas ecuaciones en dos pasos:

- primero, se ha calibrado la ecuación de crecimiento en diámetro de ambos modelos, dado que el diámetro funciona como variable independiente en la ecuación de crecimiento en altura y, por lo tanto, es un valor que influye en la calibración de la misma. Los resultados de la calibración se muestran en las ecuaciones 4.1 y 4.2 para $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente.

$$d_{i5} = \mathbf{1.18} + e^{0,2030*\ln(d*10)+0,4414*\ln(\frac{cr+0,2}{1,2})+0,8379*\ln(SI)-0,1295*\sqrt{G}-\frac{0,0007*bal^2}{\ln(d*10)}} / 10 \quad (4.1)$$

$$d_{i5} = \mathbf{1.70} + e^{-0,37110+0,2525*\ln(d*10)+0,7090*\ln(\frac{cr+0,2}{1,2})+0,9087*\ln(SI)-0,1545*\sqrt{G}-0,0004*\frac{bal^2}{\ln(d*10)}} / 10 \quad (4.2)$$

- a continuación, se ha repetido el análisis para conocer la magnitud del sesgo en altura, procediendo también a su calibración, y obteniendo como resultado las ecuaciones 4.3 y 4.4 para $IBERO^{PT}$ e $IBERO^{PS}$, respectivamente.

$$h_{i5} = \mathbf{0.42} + e^{0,21603+0,40329*\ln(\frac{d_{i5}}{2})-1,12721*\ln(d*10)+1,18099*\ln(h*100)+3,01622*cr} / 100 \quad (4.3)$$

$$h_{i5} = \mathbf{1.14} + e^{3,1222-0,4939*\ln(d_{i5}*10)+1,3763*\ln(SI)-0,0061*bal+0,1876*\ln(cr)} / 100 \quad (4.4)$$

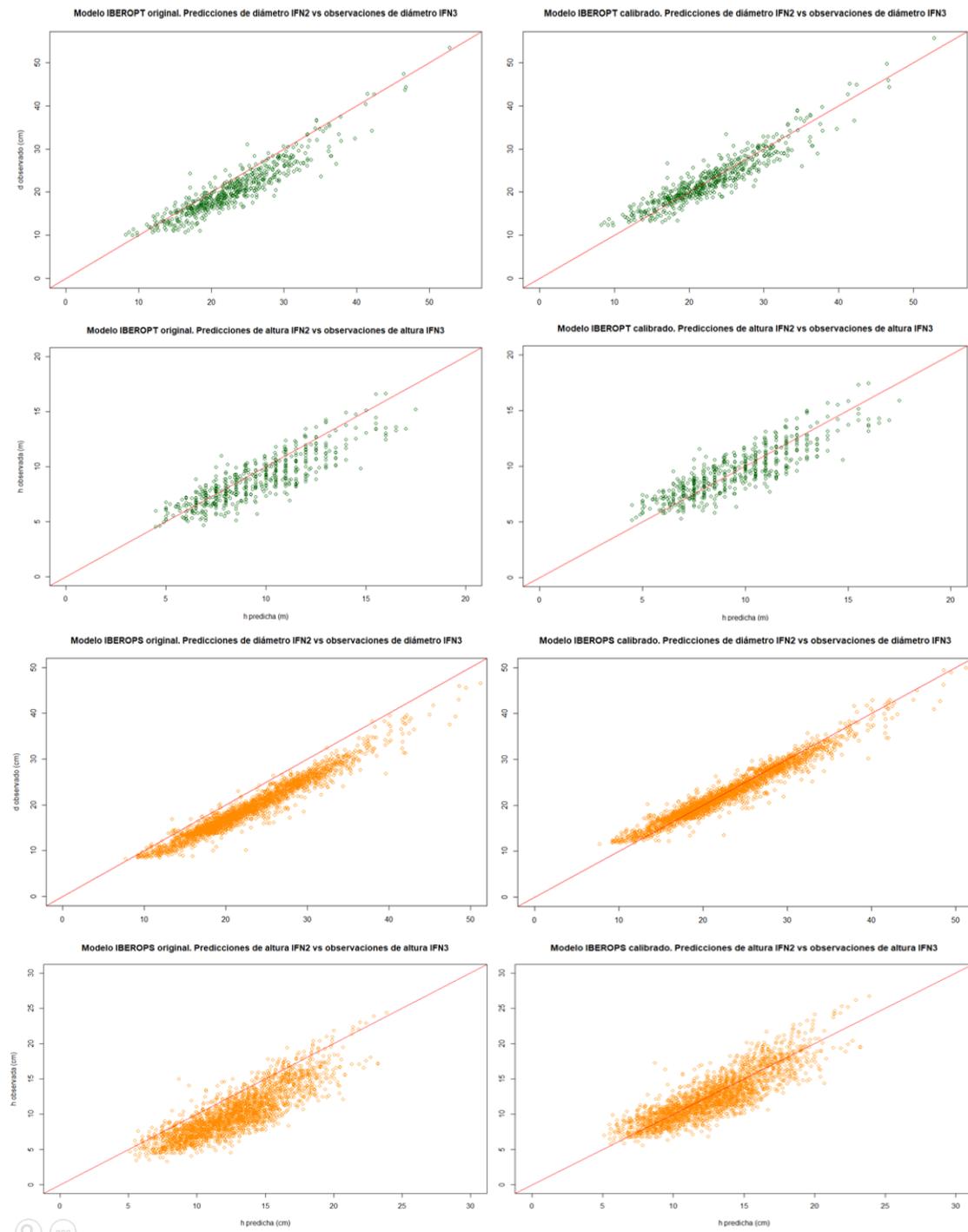


Figura 4.1: Gráficos de datos predichos (eje horizontal) y observados (eje vertical) generados con R para las variables de diámetro normal y altura de los modelos $IBERO^{PT}$ e $IBERO^{PS}$ antes y después de su calibración

De este modo, se han desarrollado las ecuaciones calibradas para crecimiento en diámetro (ecuaciones 4.1 y 4.2) y altura (ecuaciones 4.3 y 4.4) de ambos modelos, cuyas diferencias pueden apreciarse en la Figura 4.1, diseñada a partir de las recomendaciones de Piñeiro y col. (2008). En las ecuaciones calibradas se ha 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, también es apreciable una mejora, tanto en el ajuste como en la correlación de valores observados y predichos, en los análisis realizados tras la calibración del crecimiento frente a los análisis hechos con las ecuaciones originales, tal y como se aprecia en el cuadro 4.1.

Cuadro 4.1: Resultados obtenidos para el coeficiente de correlación muestral de Pearson (r), el coeficiente de correlación muestral de concordancia de Lin (r_c) y factor de precisión (C_b) para los distintos casos de estudio de las ecuaciones de crecimiento.

		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 calibrados	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

4.2. Ecuaciones de mortalidad

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 (ver Cuadro 4.2), se ha procedido a la evaluación de las ecuaciones de mortalidad 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 de esta ecuación (Figura 4.2). Este resultado supone que el criterio aplicado para su evaluación no es suficiente o no es correcto, impidiendo en este caso la mejora de las predicciones.

4.3. 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 (Cuadro 4.2), por lo tanto, 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 4.2), no encontrándose mejoras notables para Ho (precisión en el modelo original alta).

De nuevo y al igual que para el caso anterior, los resultados denotan que el criterio aplicado para

evaluar las ecuaciones no es suficiente, tal y como muestran los resultados gráficos y estadísticos mencionados anteriormente.

Cuadro 4.2: Resultados obtenidos para el coeficiente de correlación muestral de Pearson (r), el coeficiente de correlación muestral de concordancia de Lin (r_c) y factor de precisión (C_b) para los casos de estudio más representativos de las ecuaciones de mortalidad y masa incorporada.

		IBERO ^{PT}			IBERO ^{PS}		
		N	G	Ho	N	G	Ho
modelo control calibrado	r	0,63443	0,71958	0,79681	0,73423	0,79962	0,89247
	r_c	0,59384	0,59026	0,78693	0,56554	0,60565	0,88590
	C_b	0,93602	0,82028	0,98760	0,77025	0,75742	0,99264
mortalidad	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

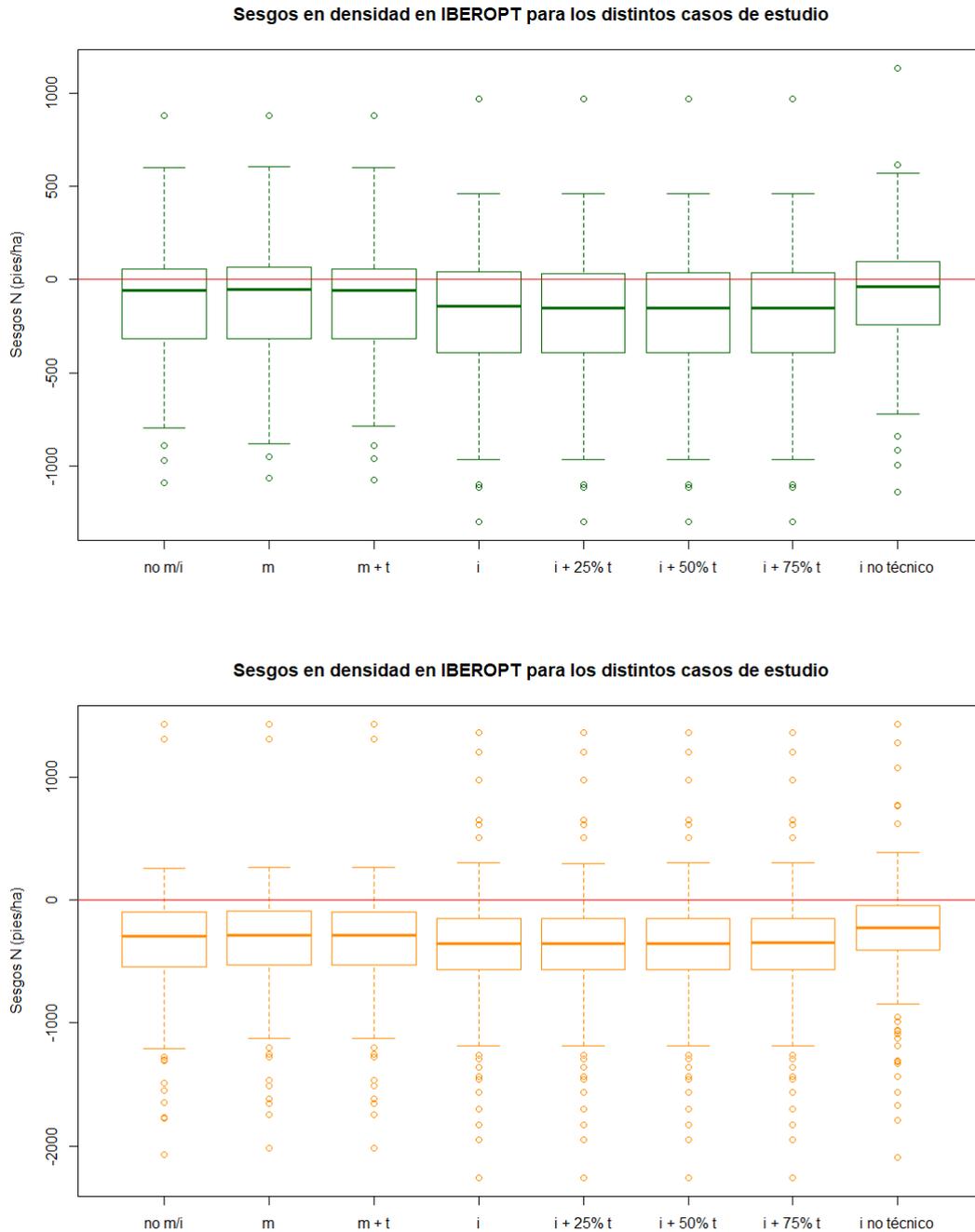


Figura 4.2: Gráficos de cajas para los errores entre densidad predicha y observada en los modelos $IBERO^{PT}$ (verde) e $IBERO^{PS}$ (naranja) a lo largo de los distintos casos de estudio: no m/i (simulación con el modelo control calibrado); m (simulación empleando la ecuación de mortalidad original); m + t (simulación empleando la ecuación de mortalidad tras añadir el valor umbral); i (simulación empleando la ecuación de masa incorporada original); i + 25 % t (simulación empleando la ecuación de masa incorporada con valor umbral 0.25); i + 50 % t (simulación empleando la ecuación de masa incorporada con valor umbral 0.50); i + 75 % t (simulación empleando la ecuación de masa incorporada con valor umbral 0.75); i no técnico (simulación empleando la ecuación de masa incorporada sin contemplar la masa incorporada técnica)

Capítulo 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 avanzando y depurando la metodología hasta lograr mejores técnicas de validación.

5.1. Ecuaciones de crecimiento

Los resultados obtenidos para las ecuaciones de crecimiento han sido muy satisfactorios. La metodología de análisis estadístico 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 y col. (2006) y Pretzsch y col. (2002) para el simulador SILVA. Tras esto, se ha procedido a la calibración en dos pasos, corrigiendo primero los sesgos de la ecuación de crecimiento en diámetro, y posteriormente corrigiendo los sesgos de la ecuación de crecimiento en altura tras el análisis del paso anterior. Esto nos ha permitido tener en cuenta la influencia de haber calibrado el diámetro sobre el cálculo del sesgo en altura, suponiendo una mejora sustancial en las predicciones de los modelos.

La calibración de ambas ecuaciones ha supuesto una mejora sustancial en el valor del coeficiente de correlación de Lin (Cuadro 4.1), 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 para en el análisis de alturas, 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, por lo 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 todavía más. 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 de las ecuaciones 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 en las predicciones de los modelos $IBERO^{PT}$ e $IBERO^{PS}$, hecho que ya se había recalado en la elaboración inicial de las mismas (Lizarralde, 2008). Además de la validación con de las ecuaciones con datos de rodales extramaduros, como se propone en Lizarralde (2008), sería interesante hacer validaciones de las ecuaciones por clases diamétricas, dividiendo así la muestra en estratos tal y como propone Vanclay y Skovsgaard (1997), pudiendo obtener como resultado variaciones de una misma ecuación principal que permitan mejorar precisión de las predicciones para cada una de las clases diamétricas estudiadas.

5.2. Ecuaciones de mortalidad y masa incorporada

Tras un primer análisis de las predicciones para las variables de masa estudiadas (N, G y Ho), se ha encontrado un ajuste pobre para las variables N y G, con una precisión mayor para la variable Ho aunque sin llegar a ser un ajuste moderado para todos los casos de estudio (ver Cuadro 4.2). Este resultado encontrado en el análisis inicial se debe a la forma en la que se calcula Ho, 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 Ho dependa en mayor parte de las predicciones de altura (calibradas en el apartado anterior) que del efecto de las ecuaciones de mortalidad y masa incorporada (estudiadas en este apartado), y de ahí que sus predicciones sean mejores durante 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 han resultado satisfactorios para la validación de las ecuaciones de mortalidad y masa incorporada. En ambos casos, se ha estudiado la importancia de un factor umbral que determine cuando han de aplicarse dichas ecuaciones, creando un valor umbral en base a la densidad de la masa para la ecuación de mortalidad, y modificando el ya existente para el caso de la masa incorporada. Los resultados muestran diferencias ínfimas para ambos casos, lo que nos hace entender que la relevancia de este factor es pequeña.

Aunque esta alteración de la ecuación de mortalidad, la cual incluye un valor umbral, está diseñada para ser aplicada en modelos dependientes de la distancia (Bravo-Oviedo y col., 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 cuando se aplica la ecuación de mortalidad es ínfimo, al menos, con la metodología empleada. No obstante, Bravo-Oviedo y col. (2006) ha reportado tras el desarrollo de dichas ecuaciones originales buenos ajustes para la baja mortalidad de estas masas, con cierta sobreestimación para el modelo $IBERO^{PT}$ e infraestimación para el modelo $IBERO^{PS}$, cuya precisión no hemos sido capaces de mejorar.

Por otro lado, las ecuaciones de masa incorporada diseñadas en dos pasos sí muestran una alta utilidad (Bravo y col., 2008), aunque la importancia del valor umbral que determina su aplicación parece no ser de gran relevancia a juzgar por nuestros resultados. Bravo-Oviedo y col. (2006) reporta errores bajos para el modelo $IBERO^{PT}$, siendo estos mayores para $IBERO^{PS}$. 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. Otro factor importante para próximos estudios es el reparto de la masa incorporada en las

distintas clases diamétricas. Si bien para parcelas de muestreo convencionales la masa incorporada repercute mayormente sobre la menor clase diamétrica (esto es, los árboles que no eran inventariables al comienzo del muestreo y lo son tras su crecimiento) y en menor medida las de clases diamétricas mayores, para el caso de parcelas concéntricas la aproximación es diferente, incluyendo un reparto más equitativo de la masa incorporada en las distintas clases diamétricas (dado que los diámetros mínimos inventariables varían a lo largo de los distintos radios de muestreo). Esto supone que el área basimétrica incorporada represente un determinado número de pies de pequeño tamaño (parcelas convencionales) o un número de pies mucho menor con árboles de tamaños superiores (parcelas concéntricas), alterando considerablemente el funcionamiento de los modelos.

Por otro lado, 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 para 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 mortalidad 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 porqué 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 Cuadros 4.2 y 4.1, respectivamente). Uno 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.

Otra de las peculiaridades que presenta el estudio de estas ecuaciones es el planteamiento de las mismas. En las ecuaciones de mortalidad, su aplicación se realiza teniendo en cuenta los datos de cada uno de los árboles respecto al total de la masa, y decidiendo qué porcentaje de su factor de expansión (es decir, de la cantidad de árboles representados por el árbol sujeto para una masa de 1 ha) sobrevive. Para el caso de las ecuaciones de masa incorporada, se calcula la probabilidad de que exista masa incorporada y la cantidad de esta directamente sobre los datos de masa, siendo necesario repartir posteriormente esta masa incorporada entre las distintas clases diamétricas. En ambos casos el factor de expansión es una variable clave, dado que es la variable que se ve alterada en primera instancia debido al efecto de ambas ecuaciones, y a partir de la cual se recalcula la información de árbol y parcela en el flujo de cálculo de los modelos. Una simplificación que podría resultar interesante sería expresar los datos de masa por unidad de parcela, independientemente de la superficie de la misma. Este planteamiento solucionaría el hecho de trabajar con los factores de expansión de cada árbol, mientras que presentaría problemas tales como son la dificultad de comparar los resultados de distintas masas forestales entre sí (por ser datos referidos a distintos tamaños de parcela) o bien la imposibilidad de trabajar con datos de parcelas concéntricas, como son el caso de las parcelas del IFN, donde los factores de expansión de cada árbol y los diámetros mínimos inventariables varían dependiendo de la distancia de cada árbol al centro de la parcela. Para este último caso, el hecho de encontrar un mejor ajuste al eliminar la masa incorporada técnica en ambos modelos nos hace cuestionarnos si debe de ser aplicada en las clases diamétricas que representen los diámetros mínimos inventariables de casa círculo concéntrico, o si bien deben

seguir una aproximación convencional y aplicarse únicamente sobre la clase diamétrica de menor tamaño. Aunque nuestro resultado apunta a que la segunda aproximación devuelve mejores resultados, un mayor esfuerzo de estudio es necesario para garantizar que nuestros resultados sean concluyentes.

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

- **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 $IBERO^{PT}$ 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, dado que cuanto mayor es el tamaño muestral más fiables son los resultados del análisis. Este aspecto se podrá mejorar una vez se publique el 4IFN.
- **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 propone Vanclay y 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 presentes en la masa.
- **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 mortalidad (Ecuaciones 3.5 y 3.6) y masa incorporada (Ecuaciones 3.14 y 3.16). Buscar un planteamiento común para la aplicación de ambas ecuaciones y rediseñarlas puede ser una opción viable. En este sentido, un planteamiento diferente podría ser reelaborar la ecuación de mortalidad siguiendo la misma estructura que la ecuación de masa incorporada, es decir, empleando los datos de masa para calcular la probabilidad de que exista mortalidad, determinar un factor umbral para su aplicación y diseñar otra ecuación que determine

la mortalidad total de la masa. Por otro lado y dentro del mismo planteamiento, se podría ensayar si el reparto de la mortalidad y masa incorporada de la masa debe realizarse sobre todas las clases diamétricas (parcelas concéntricas) o únicamente sobre la clase diamétrica inventariable menor, dado que nuestros resultados muestran un mejor ajuste para este último caso. De este modo, se podría simplificar el posterior análisis y emplear un procedimiento común para el estudio de ambas ecuaciones.

- **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 mortalidad y masa incorporada.

Capítulo 6

Conclusiones

La modelización es una herramienta de gran importancia para la gestión forestal. Mediante este trabajo se ha diseñado una metodología consistente en comparar los valores predichos para la simulación de diversos casos de estudio con los valores observados para su posterior ajuste en las predicciones. Esta metodología ha permitido evaluar los modelos $IBERO^{PT}$ 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 mortalidad y masa incorporada. De este modo, las predicciones de los modelos originales $IBERO^{PT}$ e $IBERO^{PS}$ implementados en SIMANFOR han sido mejoradas, siendo necesarios futuros trabajos para poder validar la totalidad del modelo.

Capítulo 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.

Particularmente, agradecer la importante contribución de Cristóbal Ordóñez, quien no ha dudado en colaborar en este trabajo y ha estado disponible para todas las consultas necesarias durante el desarrollo del mismo.

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., 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. *Invest Agrar: Sist Recur For*, 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. y col. (2011). Growth and yield models in Spain: historical overview, contemporary examples and perspectives. *Forest Systems*, 2011, vol. 20, núm. 2, p. 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., 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 Senespleda, E. & Montero González, 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). Consejería de Medio ...
- Institute, S. (2011). *SAS/STAT 9.3 User's Guide: Survey Data Analysis (book Excerpt)*. SAS institute.
- 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. (2010). IBEROPS: Modelo de dinámica de rodales de *Pinus sylvestris* L. en el sistema ibérico meridional.
- Lizarralde, I., Ordóñez, C., Bravo-Oviedo, A. & Bravo, F. (2010). IBEROPT: Modelo de dinámica de rodales de *Pinus pinaster* Ait. 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. d. (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. *Forest dynamics, growth and yield* (pp. 1-39). Springer.
- Pretzsch, H., Biber, P. & Ďursk, 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 de Prado, D., San Martín, R., Bravo, F. & de Aza, C. H. (2020). Potential climatic influence on maximum stand carrying capacity for 15 Mediterranean coniferous and broad-leaf 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.
- 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.
- SIMANFOR. (2021). SIMANFOR - Sistema de Apoyo para la Simulación de Alternativas de Manejo Forestal Sostenible.
- 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.
- Weiskittel, A. R., Hann, D. W., Kershaw Jr, J. A. & Vanclay, J. K. (2011). *Forest growth and yield modeling*. John Wiley & Sons.