2025 | 16·20
GIJÓN | JUNIO

9º CONGRESO **FORESTAL** ESPAÑOL

9CFE-1580

Actas del Noveno Congreso Forestal Español
Edita: **Sociedad Española de Ciencias Forestales. 2025**.
ISBN: **978-84-941695-7-1**

Organiza

Sociedad Española de Ciencias Forestales
SECF

Productividad de nuevas variedades clonales de chopo en función de las condiciones ambientales

ALONSO PONCE, R. (1,2), GARNICA, J. (2), SANTANA, J. (4), MERLO, E. (5), PIÑEIRO, M. (5), ARRAIZA, A. (1), CISNEROS, Ó. (1), CABALLÉ, G. (6)

(1) Fora Forest Technologies SLL

(2) Instituto Universitario de Investigación Gestión Forestal Sostenible iuFOR

(3) Bosques y Ríos SL

(4) Sorbus Bosques Multifuncionales SL

(5) Madera Plus Calidad Forestal SL

(6) Fundación Cesefor

Resumen

La populicultura española carece de diversidad respecto al material vegetal clonal empleado en las plantaciones, comparado con países vecinos como Francia. En este trabajo, desarrollado dentro del Grupo Operativo Biopoptech, se analiza la interacción genotipo-ambiente de 17 clones, expresada en cuanto a su capacidad productiva.

Se tomaron muestras edáficas en 18 ensayos clonales, hasta los 30 cm de profundidad, y se analizaron en laboratorio 15 parámetros relativos a sus propiedades físicas y químicas. Además, se incorporaron 33 variables climáticas. La productividad se evaluó mediante el incremento medio anual en área basimétrica. Las relaciones entre variables abióticas y productividad se analizaron mediante correlación multivariante.

Los resultados revelan diferencias significativas en la productividad, con tres de los cuatro clones “clásicos” por debajo de la media del conjunto, mientras que los clones AF4, 100BYR y Hoogvorst exhiben productividades medias superiores. Además, se ha entrenado un modelo capaz de predecir insesgadamente la productividad a partir de variables ambientales, con un RMSE del 20,2% y explicando casi el 73% de la variabilidad.

Los resultados de este trabajo constituyen un significativo avance para poder respaldar en el futuro la toma de decisiones relativas a la elección del clon óptimo en cada estación.

Palabras clave

Populicultura, crecimiento, propiedades edáficas, condiciones climáticas, modelo predictivo.

1. Introducción

El chopo constituye una especie de gran relevancia para el sector maderero en España , particularmente en el contexto de la producción de tableros contrachapados (Muñoz García 2018). La baja densidad de su madera, su color claro y su alta estabilidad dimensional la hacen particularmente idónea para aplicaciones industriales de muy diverso tipo (AEFCON et al., 2020). A pesar de ello, la disponibilidad de chopo como materia prima para la industria maderera se enfrenta a un escenario de escasez, en un momento en el que el consumo de esta madera por parte de la industria nacional se ha duplicado y donde, ante la falta de materia prima, las empresas españolas se ven obligadas a importarlas de otros países, como Francia (GARCÍA HERNÁNDEZ, 2018).

Las plantaciones de chopo están pues disminuyendo, y los factores que han propiciado esta situación son múltiples. Entre ellos se encuentran las trabas administrativas, la competencia con cultivos agrícolas menos rentables pero que disfrutan de líneas de subvención en el marco de la PAC y la creciente susceptibilidad de los clones más utilizados a plagas, enfermedades y a las cambiantes condiciones climáticas. Así, por ejemplo, el clon I-214, base de la populicultura española e italiana se ha visto afectado de forma repetida en los últimos años por ataques de pulgón lanígero (*Phloeomizus passerinii*). Lo mismo ocurre con el clon Beaupré, en este caso afectado por la roya (*Melampsora* spp.). En general, los cinco clones prevalentes a día de hoy (MC, Raspalje y Unal, además de los dos ya mencionados) muestran en los últimos años signos evidentes de peor desempeño en términos generales. De acuerdo con los datos de plantas comercializadas de cada clon en los últimos cinco años por los viveros registrados en Castilla y León, estos cinco clones de chopos suponen el 94,3% de todas las plantaciones (RUEDA et al., 2019).

Si bien estos clones han demostrado una alta productividad en condiciones óptimas, su amplia utilización ha generado una homogenización genética que aumenta significativamente la vulnerabilidad de las plantaciones frente a factores adversos como enfermedades, plagas y las nuevas condiciones climáticas derivadas del cambio global, tales como el incremento de temperaturas, la alteración en los regímenes hídricos y la mayor frecuencia de fenómenos meteorológicos extremos (IPCC, 2023). En este contexto, los ensayos de evaluación clonal desempeñan un papel crucial (RUEDA & VILLAR, 2024). Estos ensayos, diseñados bajo principios estadísticos robustos, permiten analizar la respuesta productiva de distintos clones frente a una variedad de condiciones ambientales.

La variabilidad clonal en chopo en cuanto a crecimiento y propiedades tecnológicas de la madera es ampliamente reconocida y estudiada en otros países, con programas de selección clonal adaptados a las condiciones específicas del sitio y las condiciones medioambientales (RUEDA et al., 2016). En España se realizó un trabajo muy importante de selección clonal de chopo pero que no se mantuvo en el tiempo. No obstante, es factible aprovechar el trabajo realizado en otros lugares y experimentar en su territorio los clones seleccionados y que ya están disponibles comercialmente o en avanzado estado de selección y éste es el caso de España. Estos nuevos clones suponen una oportunidad para optimizar el material clonal de nuestro país y aportar mayor diversidad incrementando la resiliencia, productividad y calidad tecnológica de nuestras choperas.

En el presente trabajo, dentro del marco del Grupo Operativo BIOPOPTECH, se han

evaluado 17 clones en cuanto a su capacidad productiva en función de variables climáticas y edáficas, en 18 ensayos establecidos en distintas condiciones ambientales de las cuencas del Duero y del Ebro.

2. Objetivos

El objetivo general de este trabajo es analizar la capacidad productiva de 17 clones de chopo en función de las condiciones ecológicas (edáficas y climáticas) de la estación, en las cuencas del Duero y del Ebro. Los objetivos específicos derivados del anterior son: (i) evaluar el efecto de 48 variables ecológicas sobre el crecimiento de 17 clones de chopo en 18 ensayos ya establecidos; (ii) generar un modelo predictivo para el crecimiento de cada clon en localizaciones donde no se ha testado.

3. Metodología

Ensayos clonales

Se ha contado con un total de 18 ensayos clonales repartidos por las cuencas del Duero (León y Zamora) y del Ebro (La Rioja) (Figura 1), en los que se han analizado un total de 17 clones. Estos experimentos forman parte de un conjunto de ensayos clonales establecidos por Bosques y Ríos SL desde hace años. La lista de ensayos y clones se muestran en las Tabla 1 y 2, respectivamente. El marco de plantación es siempre 6x6 salvo en el caso de Turcia_2014, que es 5x5. En todos los casos se ha plantado ha raíz profunda. En la Tabla 1 se detallan tanto los nombres de los ensayos como su edad en 2024, la cuenca hidrográfica en la que se ubican y el número de réplicas (bloques) existentes en el ensayo.

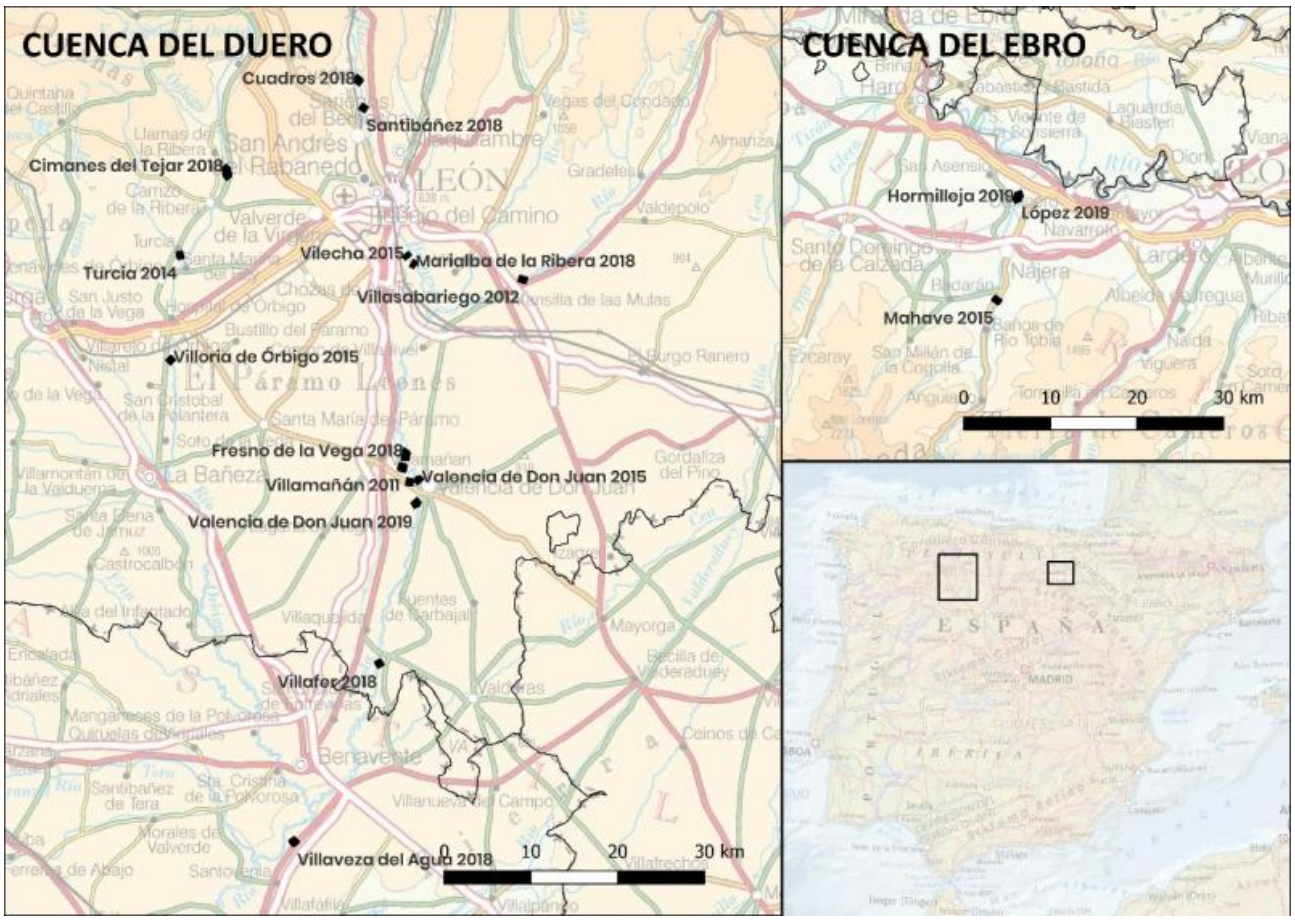

Figura 1. Localización de los 18 ensayos clonales considerados.

Tabla 1. Lista de ensayos evaluados en este trabajo. La edad se corresponde a la que tenían tras el periodo vegetativo de 2024.

Ensayo	Cuenca	Edad	Bloques
CIMANES DEL TEJAR 2018	DUERO	6	3
CUADROS 2018	DUERO	6	3
FRESNO DE LA VEGA 2018	DUERO	6	3
MARIALBA DE LA RIBERA 2018	DUERO	6	3
SANTIBÁÑEZ 2018	DUERO	6	3
TURCIA 2014	DUERO	10	1
VALENCIA DE DON JUAN 2010	DUERO	14	1
VALENCIA DE DON JUAN 2015	DUERO	9	1
VALENCIA DE DON JUAN 2019	DUERO	5	3
VILECHA 2015	DUERO	9	1
VILLAFER 2018	DUERO	6	3
VILLAMAÑÁN 2011	DUERO	13	3
VILLASABARIEGO 2012	DUERO	12	3
VILLAVEZA DEL AGUA 2018	DUERO	6	3
VILLORIA DE ÓRBIGO 2015	DUERO	9	1
HORMILLEJA 2019	EBRO	5	3
LÓPEZ 2019	EBRO	5	3
MAHAVE 2015	EBRO	9	3

Los clones considerados se exponen en la Tabla 2. No todos los clones están presentes en todos los ensayos.

Tabla 2. Lista de clones evaluados en este trabajo.

Clon	Origen	Clon	Origen
39 BYR		HOOGVORST	INBO
100 BYR		I214	CREA
AF4	ALASIA	KOSTER	3C2A
AF8	ALASIA	RASPALJE	INBO
AF13	ALASIA	SENNA	CREA
ALERAMO	CREA	TUCANO	CREA
BEAUPRE	INBO	UNAL	INBO
DIVA	CREA	VESTEN	INBO
GUARDI	CREA		

Clon	Origen	Clon	Origen

Evaluación de la productividad

Los diámetros normales de todos los árboles de estos ensayos son medidos anualmente, a partir del séptimo año, por el personal de Bosques y Ríos, aunque para este trabajo se midieron también los de 6 años. En el caso de los tres ensayos con sólo 5 años, los diámetros fueron medidos *ad hoc* por el personal de Madera Plus SL.

Con estos datos es factible calcular el crecimiento anual medio en área basimétrica a una edad de referencia dada (IGM, m^2/ha/año), y utilizar esa variable como evaluadora de la productividad. No se utilizó el crecimiento diametral ya que no todos los ensayos presentaban el mismo marco de plantación, con lo que la diferente densidad introduciría ruido en el análisis. La edad de referencia escogida fue 7 años. Esta elección ha supuesto un compromiso entre utilizar una edad lo más avanzada posible, y la disponibilidad de datos. Así, de los 18 ensayos, 8 superaban dicha edad (y por lo tanto su área basimétrica puede ser directamente calculada), mientras que otros 7 tenían 6 años y otros tres, 5. A estos ensayos se les simuló el crecimiento hasta los 7 años mediante la aplicación de las ecuaciones de índice de sitio de CASTEDO-DORADO (2023). En esta publicación sólo se establecieron curvas para I-214, Beaupré, MC, Raspalje y Unal, por lo que para los clones distintos de éstos se ha calculado la media de las cinco curvas disponibles a la hora de realizar la simulación.

Variables edáficas

La toma de muestras edáficas se realizó de forma estandarizada, cumpliendo los siguientes pasos:

- Determinación de puntos de muestreo distribuidos por cada bloque. Normalmente se trató de tres puntos por bloque.

- Ejecución de catas de hasta 50 cm de profundidad en los puntos aleatorios de muestreo, utilizando azadón, pico y pala franca. Una vez abierta y profundizada la cata, se extrajo tierra practicando un corte vertical entre los 25 y los 50 cm de profundidad, cuidando de no tomar material de la zona superficial (de influencia antrópica), y se embolsó, etiquetando la submuestra correspondiente.

- Mezcla de las submuestras de tierra de cada cata perteneciente a una misma unidad (bloque de ensayo clonal), conformando así la muestra definitiva. Si hubiera variaciones de calidad de terreno en esas unidades, se conformarán tantas muestras diferentes para esa unidad como rodales discriminados hubiera, anotándose a qué clones dentro del bloque correspondía la nueva muestra.

- La muestra de suelo definitiva se preparó de acuerdo con las exigencias del laboratorio certificado (extracción de elementos contaminantes o que puedan alterar los análisis, como piedras grandes, plásticos, restos vegetales caídos accidentalmente).

Las variables edáficas calculadas a partir de los datos de laboratorio, así como sus métodos de laboratorio, unidades y acrónimos se muestran en la Tabla 3. En el caso de la variable GRUESOS, se trata de una variable ordinal, estimada en campo con criterio experto, con tres niveles (alto, medio, bajo).

Tabla 3. Variables evaluadoras de las propiedades edáficas empleadas en el análisis.

Variable	Acrónimo	Ud	Método
pH	PH	pH	Potenciometría
pH (extracto KCl)	PHKCL	pH	Potenciometría
Materia orgánica oxidable	MO	%	Potenciometría
Carbonato cálcico equivalente	CACO3	%	Potenciometría
Nitrógeno total	N	%	Conductividad térmica
Fósforo	P	mg/kg	Espectrofotometría UV-VIS
Potasio	K	mg/kg	Espectrometría ICP-OES
Calcio	CA	mg/kg	Espectrometría ICP-OES
Magnesio	MG	mg/kg	Espectrometría ICP-OES
Sodio	NA	mg/kg	Espectrometría ICP-OES
Arcilla < 0,002 mm	ARC	%	Sedimentación discontinua
Limo (0.002 - 0.05 mm)	LIM	%	Sedimentación discontinua
Arena fina (0,05 - 2 mm)	ARE	%	Sedimentación discontinua
Capacidad de intercambio catiónico efectiva	CIC	mEq/100 g	Espectrofotometría UV-VIS
Presencia de gruesos (>2 mm)	GRUESOS	ordinal	Pericial (alta, media, baja)

Variables climáticas

Se han calculado, para cada ensayo, 33 variables climáticas a partir de los datos termopluviométricos mensuales proporcionados por los modelos del Atlas Climático Ibérico (AEMET, 2011). La definición, acrónimo y variables se detallan en la Tabla 4.

Tabla 4. Variables climáticas empleadas en el análisis.

Variable	Acrónimo	Ud	Definición
Precipitación total	PT	mm	Suma de las 12 precipitaciones mensuales
Precipitación de primavera	PP	mm	Suma de las precipitaciones de primavera (marzo, abril, mayo)
Precipitación estival	PV	mm	Suma de las precipitaciones de verano (junio, julio, agosto).
Precipitación de otoño	PO	mm	Suma de las precipitaciones de otoño (septiembre, octubre, noviembre).
Precipitación invernal	PI	mm	Suma de las precipitaciones de invierno (diciembre, enero, febrero)
Temperatura media anual	TM	ºC	Media de las 12 temperaturas medias mensuales
Temperatura media de primavera	TMP	ºC	Media de las temperaturas medias de los meses de primavera
Temperatura media estival	TMV	ºC	Media de las temperaturas medias de los meses de verano
Temperatura media de otoño	TMO	ºC	Media de las temperaturas medias de los meses de otoño
Temperatura media invernal	TMI	ºC	Media de las temperaturas medias de los meses de invierno
Temperatura máxima del mes más cálido	TXC	ºC	Temperatura máxima del mes con temperatura media más alta
Temperatura máxima de primavera	TXP	ºC	Media de las temperaturas máximas de primavera
Temperatura máxima estival	TXV	ºC	Media de las temperaturas máximas de verano
Temperatura máxima de otoño	TXO	ºC	Media de las temperaturas máximas de otoño
Temperatura máxima invernal	TXI	ºC	Media de las temperaturas máximas de invierno
Temperatura mínima del mes más frío	TNF	ºC	Temperatura mínima del mes con temperatura media más baja
Temperatura mínima de primavera	TNP	ºC	Media de las temperaturas mínimas de primavera
Temperatura mínima estival	TNV	ºC	Media de las temperaturas mínimas de verano
Temperatura mínima de otoño	TNO	ºC	Media de las temperaturas mínimas de otoño
Temperatura mínima invernal	TNI	ºC	Media de las temperaturas mínimas de invierno
Oscilación media	OSCM	ºC	Diferencia entre la temperatura media del mes más cálido y la del mes más frío
Oscilación máxima	OSCXN	ºC	Diferencia entre la temperatura máxima del mes más cálido y la

Variable	Acrónimo	Ud	Definición
			mínima del mes más frío (TXC-TNF)
Número de meses fríos	NMF	meses	Número de meses en los que T?6
Duración de la sequía	DSQ	meses	Numero de meses en los que la curva de 2 x temperaturas está por encima de la de precipitaciones.
Intensidad de la sequía	ISQ	º/1	Cociente entre el área seca y el área húmeda
Periodo vegetativo	G	meses	Número de meses en los que T?6ºC y P?2T
Evapotranspiración	ETP	mm	Suma de las 12 ETP según Thornthwaite
Suma de déficits	DEF	mm	Suma de las diferencias ETP-P mensuales, siempre que ETP>P
Suma de superávits	SUP	mm	Suma de las diferencias P-ETP mensuales, siempre que P>ETP
Índice hídrico	IH	adim	Indicador síntesis de SUP, DEF y ETP según la fórmula IH=(100SUP-60DEF)/ETP
Índice de Vernet	IVERNET	ºC/mm	(Vernet y Vernet, 1966)
Índice de aridez de Martonne	MARTONNE	mm/ºC	PT/(TM+10)
Índice de aridez de Emberger	EMBERGER	mm/K²	100*PT/(TMC²-TMF²)) siendo TMC la temperatura media del mes más cálido, y TMF la temperatura media del mes más frío, ambas expresadas en grados Kelvin.

Análisis estadístico

Como se ha explicado, el objetivo del análisis es evaluar el efecto de 48 variables ecológicas (15 edáficas y 33 climáticas) sobre la productividad de 17 clones en 18 sitios de ensayos. Así pues, se trata de un ensayo de múltiples ambientes caracterizados por numerosas variables ecológicas que va a servir para predecir el crecimiento de un clon en sitios donde no se ha testado.

Cuando se trata de predecir el efecto del ambiente sobre el crecimiento de diferentes genotipos, la regresión Finlay-Wilkinson es uno de los métodos más directos. La idea de este método es que el efecto de cada ambiente se caracteriza por la media del efecto de los genotipos. Una vez definido el efecto ambiental, se pueden buscar modelos para incluir covariables ecológicas para predecir este efecto. La creciente disponibilidad de información temática ambiental favorece esta línea en el sector forestal y agrícola. PIEPHO & BLANCON (2023) han hecho una revisión de diferentes alternativas para el desarrollo de estos modelos. FERNANDES et al. (2024) proponen el cálculo de los

efectos genéticos de cada clon en cada sitio de ensayo para después integrarlos con los efectos ambientales mediante modelos de aprendizaje automático. A partir de las propuestas de estos autores se plantea:

- Emplear un modelo mixto para extraer el efecto del ambiente.
    - Los ensayos no están equilibrados. Los modelos mixtos con estimación de las varianzas mediante máxima verosimilitud restringida (REML) y mejor predicción insesgada de los efectos (BLUP) permiten abordar análisis con datos desequilibrados, tanto en el diseño estadístico como en el balance de entradas genéticas. Además, permiten la comparación entre ensayos en distintos ambientes o establecer comparaciones a lo largo del tiempo (FERREIRA et al., 2023; RESENDE, 2016).
- Emplear *gradient boosting* como modelo de regresión para predecir el efecto ambiental en función de las variables ecológicas. *Gradient boosting* es un método de aprendizaje automático especialmente adecuado para modelizar el crecimiento cuando están involucradas un número elevado de variables ecológicas. Estas variables están en gran parte correlacionadas entre sí, cada una explica una pequeña parte del efecto de interés y suelen intervenir en forma no lineal. Este tipo de estructura se modeliza adecuadamente mediante esta técnica. En la evaluación de este modelo se han seguido los siguientes pasos:
    - Selección y transformación de variables (*feature engineering*). Para facilitar el proceso y mejorar los resultados, se ha aplicado un proceso habitual en aprendizaje automático, la creación de nuevas variables a partir de las variables originales. En particular, se observa que el resultado del ajuste mejora al emplear la intersección de variables ambientales. Es deseable que el modelo sea parsimonioso y que no incorpore más variables de las necesarias para facilitar el proceso de predicción. En esta línea, se ha empleado el estadístico MIC (*maximal information coefficient*) por su velocidad y versatilidad para encontrar dependencia entre grandes conjuntos de variables y detectar relaciones lineales y no lineales. Una descripción se puede encontrar en RESHEF et al. (2016). Con este algoritmo se ha realizado una selección inicial de las variables ecológicas más estrechamente relacionadas con el efecto ambiental y se repite el proceso con las variables intersección. Se ha empleado la librería Minerva del programa estadístico R (FILOSI et al., 2021).
    - Ajuste de parámetros del modelo de regresión por *gradient boosting*. El esquema general de estos modelos es crear árboles de decisión sencillos que mejoran ligeramente el resultado del árbol anterior. El algoritmo elegido es *Extreme Gradient Boosting*, ampliamente reconocido como un sistema de clasificación y regresión rápido y eficaz. Se ha empleado la librería XGBoost del programa estadístico R (CHEN & GUESTRIN, 2023). Un aspecto importante en el proceso de creación del modelo XGBoost es el ajuste de los hiperparámetros, que son parámetros que no se ajustan en

el proceso normal de entrenamiento del modelo. Los hiperparámetros ajustados son *eta* (tasa de aprendizaje), *max_depth* (profundidad máxima del árbol), *min_child_weight* (peso mínimo de las ramas), *subsample* (proporción de la muestra empleada en entrenamiento) y *colsample_bytree* (porcentaje de las variables empleadas en la construcción del siguiente árbol). El ajuste de hiperparámetros se ha realizado mediante optimización bayesiana minimizando el cuadrado de la suma de residuos, con la librería ParBayesianOptimization del programa R (WILSON, 2022).

    - Para calibrar el efecto del posible sobreajuste en el entrenamiento se aplica validación cruzada con una proporción entre entrenamiento y testado de 0,75/0,25. El proceso de ajuste y validación se repite 100 veces para generar una distribución del sesgo y del error medio cuadrático.

- Evaluar el ajuste del modelo final, que incluye los términos intercepción y efecto genético del modelo mixto más la predicción del efecto ambiental del modelo de regresión por *gradient boosting*.

4. Resultados

Variabilidad de la productividad y de los parámetros ecológicos

En la Tabla 5 se muestran el número de observaciones y los valores medios, mínimo, máximo y desviación típica de IGM para cada clon en estudio. Así mismo, en la Figura 2 se muestra el gráfico de los efectos medios para cada clon para esa misma variable (ver más abajo el apartado referente al modelo mixto).

Tabla 5. Estadísticas descriptivas básicas de la variable evaluadora de la productividad (IGM, m²/ha) para cada clon estudiado.

Clon	n	Media	Mín	Máx	Desvest
39 BYR	24	0.89	0.25	2.22	0.49
100 BYR	14	1.21	0.57	2.19	0.36
AF4	18	1.23	0.52	1.69	0.34
AF8	18	1.07	0.40	2.01	0.42
AF13	36	1.24	0.63	2.10	0.42
ALERAMO	34	1.17	0.57	2.20	0.40
BEAUPRE	29	0.81	0.25	1.79	0.37
DIVA	34	1.12	0.48	2.26	0.41
GUARDI	29	1.24	0.43	2.58	0.55
HOOGVORST	34	1.14	0.51	2.32	0.41
I214	40	1.10	0.46	2.12	0.36
KOSTER	29	1.05	0.48	1.64	0.31
RASPALJE	32	0.94	0.44	1.74	0.40
SENNA	9	1.34	0.67	2.27	0.56
TUCANO	9	1.38	0.63	2.25	0.59
UNAL	22	0.86	0.39	1.41	0.34
VESTEN	9	1.00	0.71	1.63	0.28

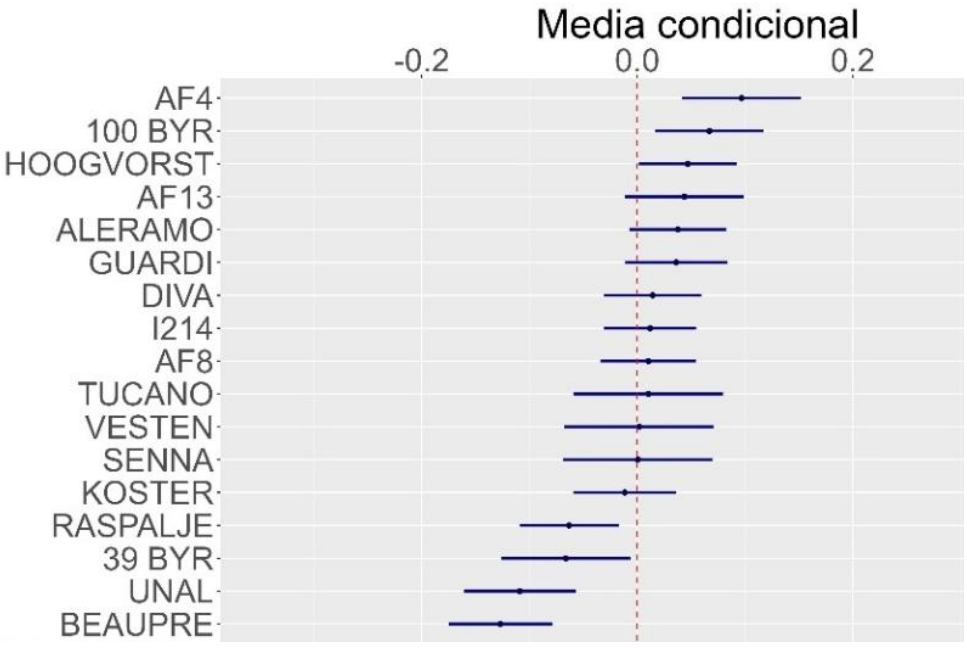

Figura 2. Efectos medios (± 2·sd) de la variable evaluadora de la productividad (IGM) para cada clon en estudio. La línea roja vertical representa el efecto medio de todos los clones.

Si bien ningún clon sobresale sobre el resto (los valores medios oscilan entre 0,81 m²/ha/año para Beaupré y 1,38 para Tucano), es destacable que tres de los cuatro clones "clásicos" testados, junto con el 39BYR, presentan peor productividad que el resto. Únicamente el clon I-214 muestra un comportamiento parecido al del resto de los clones nuevos. Los clones AF4, 100BYR y Hoogvorst manifiestan crecimientos medios significativamente mayores que la media. Los valores extremos máximos de crecimiento encontrados los presenta el clon Guardi (2.58 m²/ha/año en el ensayo Valencia de Don Juan 2019).

En cuanto a las propiedades edáficas, en la Tabla 6 se detallan los valores medios y la desviación típica encontrados, para cada clon en estudio.

Tabla 6. Media (y desviación típica) de las variables evaluadoras de las propiedades edáficas, para cada clon estudiado.

CLON	PH	PHKCL	MO	CACO3	N	P	K	CA	MG	NA	ARC	LIM	ARE	CIC
39 BYR	8.01 (0.55)	7.49 (0.75)	1.3 (0.98)	2.5 (1.52)	0.13 (0.11)	2.28 (1.39)	33.8 (15.8)	2422.7 (1280)	115.6 (60)	12.9 (7.6)	8.5 (3.9)	24.9 (11.5)	66.6 (15.1)	5.81 (2.15)
100 BYR	8.15 (0.46)	7.39 (0.49)	1.81 (1.42)	2.55 (2.49)	0.17 (0.16)	3.99 (6.87)	41.8 (23.1)	2556.7 (1311)	122.3 (48.2)	14.7 (10)	7.3 (3.7)	24.1 (12.9)	68.6 (15.8)	6.71 (3.16)
AF4	8.09 (0.5)	7.41 (0.58)	1.6 (1.47)	2.63 (2.52)	0.13 (0.14)	3.19 (1.85)	36.4 (26.2)	2459.9 (1284.6)	121.7 (50.8)	11 (7.9)	8.2 (3.7)	24.6 (12.9)	67.1 (16.1)	6.29 (2.75)
AF8	8.22 (0.39)	7.59 (0.5)	1.75 (1.29)	2.88 (2.33)	0.16 (0.15)	2.73 (1.6)	36.8 (21.9)	2917.8 (1196.5)	118.4 (57.6)	17.8 (25.7)	7.4 (3.1)	25.6 (12)	67 (14.2)	6.32 (2.6)
AF13	8.09 (0.49)	7.41 (0.57)	1.94 (1.55)	1.66 (1.58)	0.2 (0.18)	2.4 (1.18)	31.2 (16.6)	2531 (1377.8)	137.1 (68.3)	23.2 (34.9)	7.4 (3.8)	24.1 (13.2)	68.5 (16)	6.43 (2.63)
ALERAMO	8.18 (0.41)	7.48 (0.49)	1.77 (1.23)	2.73 (2.36)	0.17 (0.14)	3.62 (5.8)	42 (20.7)	2840.3 (1255.9)	122.9 (59.5)	17.7 (26.4)	7.6 (3.4)	25.5 (12.5)	66.9 (15)	6.71 (2.98)
BEAUPRE	8.09 (0.45)	7.48 (0.61)	1.65 (1.39)	2.13 (1.2)	0.14 (0.13)	3.46 (6.21)	37.1 (23.4)	2362.8 (1044.2)	118.1 (56.2)	11.4 (6.8)	7.7 (3.3)	24.2 (12.6)	68.1 (15)	6.46 (3.46)
DIVA	8.18 (0.41)	7.48 (0.49)	1.77 (1.23)	2.73 (2.36)	0.17 (0.14)	3.62 (5.8)	42 (20.7)	2840.3 (1255.9)	122.9 (59.5)	17.7 (26.4)	7.6 (3.4)	25.5 (12.5)	66.9 (15)	6.71 (2.98)
GUARDI	8.17 (0.45)	7.51 (0.57)	1.7 (1.44)	2.56 (2.4)	0.16 (0.16)	2.76 (1.73)	35.3 (22.3)	2746.2 (1286.5)	128.7 (60)	20.2 (27.9)	7.6 (3.4)	25 (11.9)	67.3 (14.6)	6.34 (2.66)
HOOGVORST	8.09 (0.41)	7.5 (0.56)	1.92 (1.34)	2.97 (2.34)	0.18 (0.14)	3.48 (5.85)	44 (24.4)	2722.1 (1272.4)	124.6 (53.5)	12.6 (8.6)	7.6 (3.4)	26.6 (13.5)	65.8 (15.7)	7.18 (3.48)
I214	8.19 (0.41)	7.55 (0.53)	1.59 (1.19)	2.67 (2.08)	0.13 (0.11)	3.47 (5.37)	39.5 (23.4)	2707.1 (1167.3)	117.5 (56.1)	15.6 (24.3)	7.8 (3.1)	25 (11.5)	67.2 (13.9)	6.26 (2.74)
KOSTER	8.1 (0.46)	7.42 (0.59)	1.81 (1.45)	2.21 (1.63)	0.17 (0.16)	2.18 (1.16)	36.3 (21.5)	2478.2 (1165.8)	124.6 (57.4)	14.8 (9.2)	7.5 (3.5)	25.4 (13.1)	67.1 (15.3)	6.64 (3.35)
RASPALJE	8.17 (0.41)	7.54 (0.5)	1.74 (1.25)	2.83 (2.4)	0.16 (0.15)	2.88 (1.63)	38.1 (21.9)	2969.5 (1240)	126.8 (60.4)	17 (27.1)	7.9 (3.4)	26.1 (12.2)	65.9 (14.6)	6.64 (2.69)
SENNA	8.04 (0.7)	7.16 (0.7)	1.18 (0.61)	1.29 (0.74)	0.08 (0.08)	2.93 (1.03)	29.7 (10)	2211.3 (1458.9)	110.7 (83.1)	31.6 (48.3)	6.3 (2.1)	19 (11.1)	74.7 (13.2)	4.47 (1.77)
TUCANO	8.04 (0.7)	7.16 (0.7)	1.18 (0.61)	1.29 (0.74)	0.08 (0.08)	2.93 (1.03)	29.7 (10)	2211.3 (1458.9)	110.7 (83.1)	31.6 (48.3)	6.3 (2.1)	19 (11.1)	74.7 (13.2)	4.47 (1.77)
UNAL	8.1 (0.47)	7.55 (0.64)	1.79 (1.13)	2.57 (1.73)	0.16 (0.15)	2.47 (1.27)	33.6 (23.5)	2705.4 (1180.5)	113.8 (54.2)	12 (8.8)	6.8 (2.6)	23.7 (12)	69.5 (13.4)	6.39 (3.51)
VESTEN	7.97 (0.67)	7.28 (0.85)	1.45 (1.24)	2.34 (1.52)	0.07 (0.1)	2.61 (1.22)	28.3 (15.1)	1942.8 (1198)	88.2 (32.2)	16.2 (9.9)	5.8 (1.8)	19.5 (11.3)	74.6 (12.9)	4.67 (2.21)

Por último, en lo que respecta a las variables climáticas, dado su elevado número, no se muestran en forma de tabla. No obstante, los valores medios y la desviación típica de las cuatro variables que han tenido más influencia en los modelos (ver apartado *Regresión efecto ambiental ~ variables ambientales*) se muestran en la Tabla 7.

CLON	PV	EMBERGER	TMI	NMF
39 BYR	87.3 (23.1)	5.12 (0.93)	4.6 (0.76)	2.70 (0.68)
100 BYR	83.4 (11.5)	5.47 (1.21)	4.57 (0.71)	2.71 (0.72)
AF4	83.9 (12.2)	5.48 (1.16)	4.56 (0.8)	2.71 (0.81)
AF8	86.8 (17.4)	5.23 (1.09)	4.78 (0.81)	2.51 (0.79)
AF13	81.4 (9.6)	5.73 (1.28)	4.24 (0.27)	3.04 (0.29)
ALERAMO	88.1 (17.2)	5.31 (1.06)	4.8 (0.83)	2.50 (0.82)
BEAUPRE	83.5 (16.7)	5.23 (1.04)	4.42 (0.56)	2.85 (0.52)
DIVA	88.1 (17.2)	5.32 (1.06)	4.8 (0.84)	2.5 (0.82)
GUARDI	81.2 (11.4)	5.25 (1.21)	4.53 (0.65)	2.74 (0.66)
HOOGVORST	88.7 (16.8)	5.37 (1.03)	4.73 (0.86)	2.56 (0.84)
I214	84.9 (16.9)	5.04 (0.95)	4.73 (0.79)	2.55 (0.76)
KOSTER	84.1 (17.2)	5.37 (1.17)	4.47 (0.58)	2.82 (0.54)
RASPALJE	88.9 (17.2)	5.4 (1.03)	4.77 (0.87)	2.52 (0.85)
SENNA	77.9 (8.6)	4.97 (1.1)	4.38 (0.46)	2.87 (0.47)
TUCANO	77.9 (8.6)	4.97 (1.1)	4.38 (0.46)	2.87 (0.47)
UNAL	86.1 (18.1)	5.28 (0.91)	4.45 (0.61)	2.83 (0.55)
VESTEN	77.1 (9.1)	4.89 (1.16)	4.36 (0.46)	2.89 (0.47)

Modelo mixto

En la Tabla 8 se resume la significación de los efectos aleatorios. Tanto el efecto del clon como el de la interacción bloque:ensayo son significativos. Se comprueba que el efecto ambiental es superior al genético.

Tabla 8. Significación de los efectos aleatorios mediante comparación de la razón de verosimilitud (variable dependiente log(IGM+1)

Efecto	npar	logLik	AIC	LRT	Df	Pr. Chisq.
	4	237.515	-467.03			
Clon	3	204.049			1	2.812e-16
Bloque:Ensayo	3	90.005	-174.01		1	< 2.2e-16

Los gráficos de residuos recogidos en la Figura 3 indican que la transformación logarítmica genera una distribución adecuada de los datos para la modelización con un modelo mixto. Por último, los estadísticos de la bondad del ajuste del modelo mixto se muestran en la Tabla 9. El modelo absorbe algo más del 72% de la varianza y el error medio cuadrático de la predicción no llega al 15%.

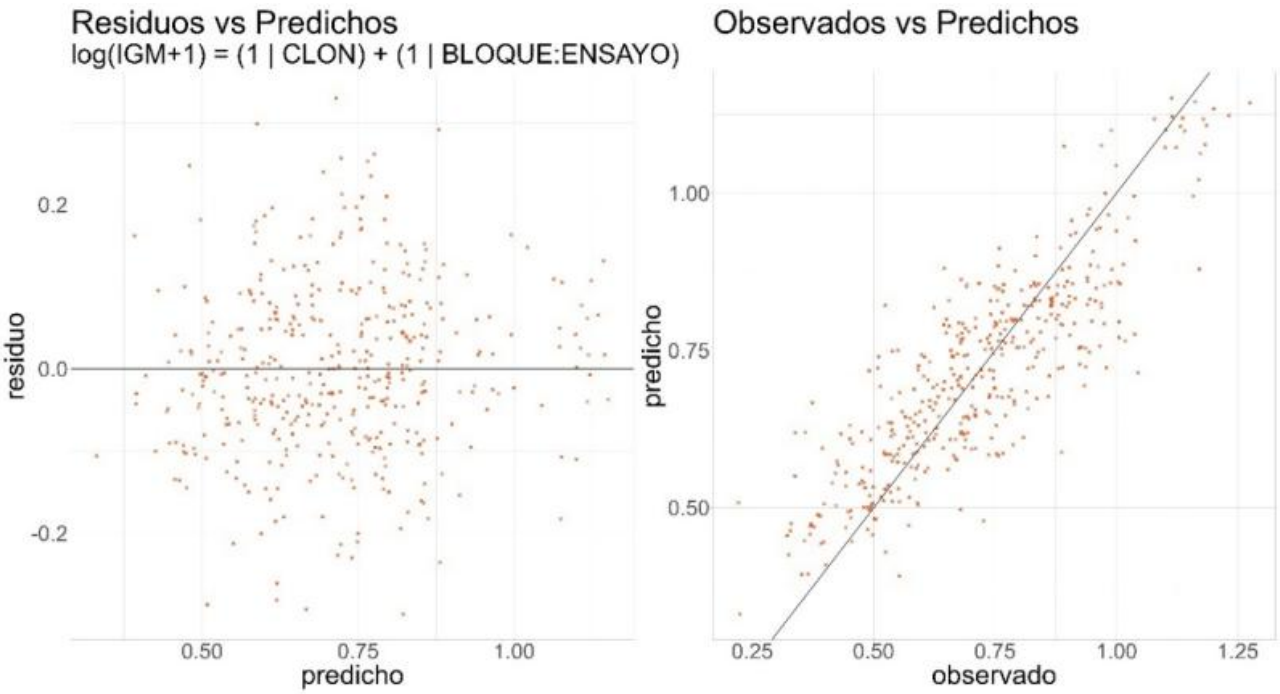

Figura 3. Gráfico de residuos del modelo mixto.

Tabla 9. Estadísticos de ajuste del modelo mixto.

R2	RMSE	RMSE (%)
72.22	0.104	14.50%

Regresión efecto ambiental ~ variables ambientales

La primera selección de variables se ha realizado a partir de MIC >= 0.85 (32 de 48 variables ambientales seleccionadas) y la segunda a partir de MIC >= 0.98 (75 de 496 variables intersección seleccionadas). El valor de MIC en la segunda selección indica la existencia de relaciones muy significativas entre el efecto ambiental y el efecto combinado de pares de variables ambientales, de mayor intensidad que el que se encuentra para variables ambientales individuales.

El proceso de ajuste de hiperparámetros propone la combinación de parámetros expuesta en la Tabla 10.

Tabla 10. Hiperparámetros del modelo XGBoost

Hiperparámetro	Valor
eta (tasa de aprendizaje)	0.05
max_depth (profundidad máxima del árbol)	7
min_child_weight (peso mínimo de las ramas)	7
subsample (proporción de la muestra empleada en entrenamiento)	0.8
colsample_bytree (porcentaje de las variables empleadas en la construcción del siguiente árbol)	0.9249

Las distribuciones del sesgo y del porcentaje de RMSE derivadas de la simulación de 100 validaciones cruzadas se resumen en los percentiles y media de la Tabla 11. Se observa que el modelo es insesgado y con tendencia a subestimar el efecto ambiental predicho, pero los errores se mantienen en la mayoría de las ocasiones por debajo del 8% en valor absoluto, por lo que se considera que el resultada es adecuado. Los abultados valores mínimo y máximo del porcentaje del RMSE indican no obstante que se puede producir sobreajuste en algunas simulaciones, previsiblemente las que impliquen modelizar con la mayoría de los valores en una cuenca hidrográfica y predecir en la otra.

Tabla 11. Distribución del sesgo y el %RMSE de las 100 validaciones cruzadas simuladas.

estadístico	sesgo	RMSE (%)
Min.	-0.00155058	-318.137424
1st Qu.	-3.49e-05	-7.87191212
Median	1.29e-06	-2.76900831
Mean	1.17e-05	-5.09963028
3rd Qu.	6.23e-05	-1.04715129
Max.	0.00272656	201.933327

El modelo final incorpora a los términos de intercepción y del efecto genético (de cada clon) del modelo mixto, el efecto ambiental previsto por el modelo de *gradient boosting*. Los resultados se resumen en la Tabla 12 y en la Figura 4. Se observa que el ajuste mantiene un porcentaje similar al del modelo mixto de varianza explicada. Aumenta cerca de seis puntos el porcentaje del error al emplear las variables ambientales directamente en la predicción, pero la posibilidad de emplear estas variables en la valoración de la productividad esperable para los 17 clones es una ventaja que compensa esta pérdida de precisión.

Tabla 12. Estadísticos de ajuste del modelo final

R2	sesgo	RMSE	RMSE (%)
72.96	0.015	0.220	20.18

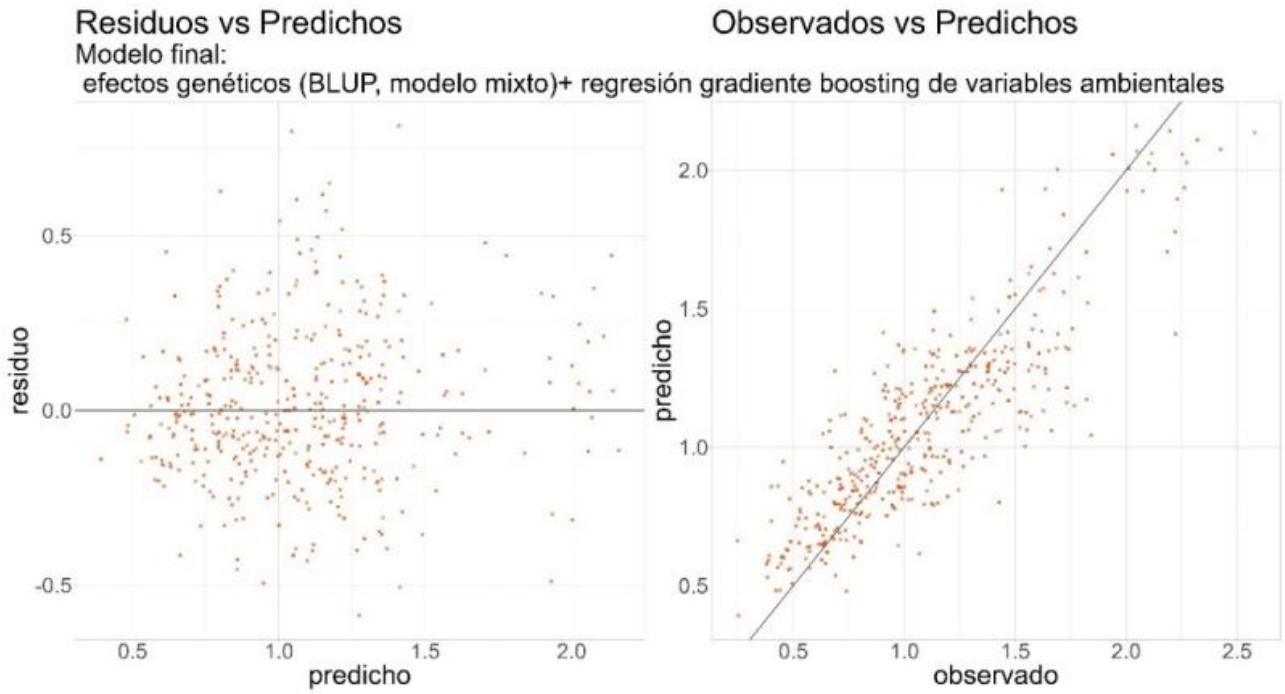

Figura 4. Gráfico de residuos del modelo final.

5. Discusión

En este trabajo se ha examinado la capacidad productiva, evaluada mediante el crecimiento medio en área basimétrica a los 7 años, de 17 clones de chopo, distribuidos en 18 ensayos clonales en las cuencas del Duero y del Ebro, y se ha analizado la relación de dicha variable con atributos edáficos y climáticos, con el objetivo de generar un modelo predictivo de la productividad y así poder tomar decisiones bien informadas sobre la idoneidad de utilizar uno u otro clon en nuevas plantaciones.

Si bien no se ha detectado que ningún clon sobresalga sobre el resto en cuanto a su productividad, es destacable que 3 de los cuatro clones analizados presentaron crecimientos significativamente por debajo de la media del conjunto de clones. Esto no ocurre, al menos parcialmente, en experimentos realizados en Francia, donde el clon Raspalje presentó crecimientos homologables a los del resto de clones (HAVET et al., 2022) o en experiencias en Castilla y León, donde los clones Raspalje y Unal ofrecieron un desempeño incluso mejor que la media de los clones testados (RUEDA & VILLAR, 2024) si bien en estos trabajos no se testaban ninguno de los clones nuevos analizados aquí, salvo Koster en el caso francés.

Los clones AF4, 100BYR y Hoogvorst demostraron crecimientos medios significativamente mayores que la media, según el modelo mixto ajustado. Además, la diferencia máxima de productividad media encontrada entre todos los clones y ensayos alcanza los 0,57 m²/ha/año, lo que implica que a fin de turno podemos encontrar

diferencias de más de 8,5 m²/ha entre dos plantaciones distintas, lo que a su vez puede traducirse, aproximadamente, en más de 100 m³/ha (volumen maderable) y 8.000 € de discrepancia en cuanto a rentabilidad, por término medio, o incluso a arruinar completamente la inversión, el los casos más extremos, al no alcanzar la chopera los diámetros mínimos para ser comercializable. Este hecho apuntala la necesidad ineludible de identificar nuevos clones que puedan sustituir paulatinamente a los actualmente prevalentes.

Además, la trascendencia de conocer a fondo las relaciones entre las condiciones abióticas de una estación y la productividad potencial de los distintos clones de chopo es indudablemente enorme, pero no abundan los trabajos que lo aborden con datos edáficos exhaustivos, análisis estadísticos sólidos y para un número elevado de clones (HAVET et al., 2022), sino que en general el sector son más frecuentes los estudios parciales (RUEDA & BENGOA, 2018) o de carácter cualitativo (RUEDA et al., 2016) que, no por útiles, son insuficientes para realizar predicciones de una precisión satisfactoria. El modelo predictivo entrenado en este trabajo para calcular la productividad a partir de variables ambientales demuestra que estas últimas son más influyentes que el efecto genético, es decir que, en general, la capacidad productiva viene determinada con mayor intensidad por las condiciones edafoclimáticas de la plantación que por el clon de que se trate. Sin embargo, no todos los clones se comportan igual en todos los ambientes, por lo que una predicción precisa *a priori* resulta crítica si se quieren obtener resultados económicamente rentables de la plantación. El modelo entrenado a partir de variables ambientales ofrece predicciones insesgadas de IGM (0.015 m²/ha/año de error) y con un error medio cuadrático del 20%, lo cual brinda la posibilidad de seleccionar el o los clones de mayor productividad prevista apoyándose en datos objetivos en lugar de en intuiciones o “modas”.

Si bien el objetivo de este modelo es únicamente predictivo y no explicativo, sí se pueden detectar algunos patrones de utilidad para futuros análisis con vocación interpretativa. Así, la combinación de variables con mayor MIC son ARC*PV y ARC*TMI (0.99999585 en ambos casos) , aunque el resto de las combinaciones retenidas están muy cerca de estos valores. En general, las variables de textura son las de mayor peso entre las edáficas, aunque es destacable que CACO3 aparece entre las de mayor MIC. En este caso puede tratarse de un efecto confundido porque las marcadas diferencias entre la cuenca del Duero y del Ebro en cuanto a CACO3 (1,82% y 5,60%, respectivamente) están acompañadas de diferencias climáticas. En cuanto a las variables climáticas, PV, TMI, NMF y EMBERGER son las de mayor intensidad en la relación entre las variables combinadas, aunque no se observa un patrón en cuanto a variables de aridez, termicidad o precipitación. En la primera selección, las variables con mayor intensidad son las climáticas térmicas, pero la combinación con variables edáficas o pluviométricas hace que este efecto se diluya en la segunda selección.

Por último, dado que este tipo de modelos no son fácilmente aplicables, se ha desarrollado una aplicación web que proporciona las predicciones para los 17 clones de manera ágil y sencilla, sin más que introducir las coordenadas de la ubicación de la futura plantación y las variables edáficas tomadas en campo. Para para aliviar la carga de análisis edáficos, se ha implementado una versión del modelo que utiliza el mínimo

conjunto de variables que produce resultados, que en el caso de las edáficas incluye a ARE, LIM, ARC, CACO3, N, CA, P, NA y GRUESOS. Las variables climáticas las recupera la propia aplicación a partir de los modelos del Atlas Climático Ibérico (AEMET, 2011). El resultado devuelto al usuario es el valor de IGM previsto para los 17 clones, ordenados de mayor a menor. El análisis se circunscribe a fincas ubicadas en las cuencas del Duero y del Ebro dentro del territorio español.

6. Conclusiones

Las conclusiones más relevantes que se pueden extraer de este trabajo son:

1. Se han detectado diferencias significativas en la productividad de los distintos clones, evaluada a través del crecimiento medio en área basimétrica a los 7 años de la plantación, lo cual puede derivar a fin de turno en una gran disparidad de rentabilidad económica según el clon empleado.

2. Tres de los cuatro clones “clásicos” testados (Beaupré, Raspalje y Unal) presentan productividades significativamente inferiores a la media.

3. Por el contrario, tres clones nuevos (AF4, 100BYR y Hoogvorst) manifiestan crecimientos significativamente superiores a la media.

4. El modelo predictivo entrenado es capaz de estimar la productividad de manera insesgada y con un error medio cuadrático medio en torno al 20%.

5. Las condiciones ambientales tienen más influencia en la productividad que el origen genético.

6. Las variables texturales son las más influyentes de entre las edáficas.

7. De entre las climáticas, PV, TMI, NMF y EMBERGER exhiben la mayor influencia en el modelo entrenado, si bien no se observa ningún patrón en cuanto a aridez, termicidad o precipitación.

7. Agradecimientos

Los autores desean expresar su agradecimiento al conjunto de socios del Grupo Operativo Biopoptech. Este trabajo ha sido financiado con fondos procedentes del Instrumento de Recuperación Europeo (EU Next Generation), a través de la Dirección General de Desarrollo Rural, Innovación y Formación Agroalimentaria (Ministerio de Agricultura, Pesca y Alimentación).

8. Bibliografía

AEFCON; COSE; FAFCYLE; JCYL; SOMACYL; 2020. Beneficios ambientales y socioeconómicos de la populicultura en la cuenca del Duero. Propuestas del sector para el III Ciclo de Planificación Hidrológica. 72 pp. https://www.chduero.es/documents/20126/1213596/N730.pdf

AEMET, 2011. Atlas climático ibérico/Iberian climate atlas, Agencia Estatal de Meteorología, Ministerio de Medio Ambiente y Medio Rural y Marino. Agencia Estatal de Meteorología-Instituto de Meteorología de Portugal. 79 pp.

CASTEDO-DORADO, F.; RODRÍGUEZ-GONZÁLVEZ, P.; GONZÁLEZ-FERREIRO, E.; RAMOS-GONZÁLEZ, G.; ÁLVAREZ-ESTEBAN, R. Y ÁLVAREZ-TABOADA, F.; 2023. A user-driven, reduced whole-stand growth model for major hybrid poplar clones in Spain. *Scand. J. For. Res.* 38, 329-343.

CHEN, T. Y GUESTRIN, C., 2016. XGBoost: A Scalable Tree Boosting System, en: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. Association for Computing Machinery. New York, NY, USA, pp. 785–794.

FERNANDES, I.K., VIEIRA, C.C., DIAS, K.O.G., FERNANDES, S.B., 2024. Using machine learning to combine genetic and environmental data for maize grain yield predictions across multi-environment trials. Theor Appl Genet 137, 189. https://doi.org/10.1007/s00122-024-04687-w

FERREIRA, F.M., CHAVES, S.F. DA S., PEIXOTO, M.A., ALVES, R.S., COELHO, I.F., RESENDE, M.D.V. DE, SANTOS, G.A. DOS, BHERING, L.L., 2023. Multi-trait multi-environment models for selecting high-performance and stable eucalyptus clones. Acta Sci., Agron. 45, e61626. https://doi.org/10.4025/actasciagron.v45i1.61626

FILOSI, M., VISINTAINER, R., ALBANESE, D., RICCADONNA, S., JURMAN, G., FURLANELLO, C., 2021. minerva: Maximal Information-Based Nonparametric Exploration for Variable Analysis.

GARCÍA HERNÁNDEZ, I. (2018). Consumo de chopo por la industria en España: evolución y necesidades futuras. https://www.simposiodelchopo.es/sites/default/files/actas/actas.pdf

HAVET, N.; BONIGEN, J.; HIGUET, V.; DE BACKER, T.; BESNARD, C.; CANO, B.; 2022. Synthèse du Réseau Régional d'Expérimentation et de Démonstration Hauts-de-France sur le peuplier. CNPF. 188 pp.

MUÑOZ GARCÍA, A.; 2018. El Tablero contrachapado de chopo, aplicación noble y mayoritaria para la madera procedente de plantaciones de chopo. https://www.simposiodelchopo.es/sites/default/files/actas/actas.pdf

IPCC; 2023: *Climate Change 2023: Synthesis Report.* Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647.

PIEPHO, H.-P., BLANCON, J., 2023. Extending Finlay–Wilkinson regression with environmental covariates. Plant Breeding 142, 621–631. https://doi.org/10.1111/pbr.13130

RESENDE, M.D.V. DE, 2016. Software Selegen-REML/BLUP: a useful tool for plant breeding. Crop Breed. Appl. Biotechnol. 16, 330–339. https://doi.org/10.1590/1984-70332016v16n4a49

RESHEF, Y.A., RESHEF, D.N., FINUCANE, H.K., SABETI, P.C., MITZENMACHER, M., 2016. Measuring Dependence Powerfully and Equitably. Journal of Machine Learning Research 17, 1–63

RUEDA J. (2016). Clones de chopos del Catálogo Nacional de Materiales de Base. Consejería de Fomento y Medio Ambiente, Junta de Castilla y León. Valladolid. Versión febrero de 2016. 72 pp.

RUEDA, J.; BENGOA, J.; 2018. Nota breve sobre el crecimiento comparado de dos clones de chopo en suelos con diferente pH. https://www.simposiodelchopo.es/sites/default/files/actas/actas.pdf

RUEDA J.; GARCÍA CABALLERO J.L.; CUEVAS Y.; GARCÍA-JIMÉNEZ C.; VILLAR C.; 2019 Cultivo de chopos en Castilla y León. Consejería de Fomento y Medio Ambiente, Junta de Castilla y León. Valladolid. 116 pp.

RUEDA J, VILLAR C (2024): Parcela de experimentación de clones de chopos ZA-7 “Calzada de Tera”. Junta de Castilla y León, Consejería de Medio Ambiente, Vivienda y Ordenación del Territorio. Valladolid, 24 pp

VERNET, J.L.; VERNET, P.; 1966. Sur un indice bioclimatique applicable aux climats de la France. Nat. Monspel. Ser. Bot. 17, 253-261.

WILSON, S., 2022. ParBayesianOptimization: Parallel Bayesian Optimization of Hyperparameters.