2025 | 16·20
GIJÓN | JUNIO
9º CONGRESO FORESTAL ESPAÑOL

9CFE-1487

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

Organiza

Recomendaciones para la generación de modelos de distribución de especies en taxones con marcada variabilidad genética: el caso de *Pinus pinaster* Ait.

LOPEZ SENESPLEDA, E. (1), ALONSO PONCE, R. (2,3) y RUIZ-PEINADO, R. (1)
(1) Instituto de Ciencias Forestales ICIFOR, INIA-CSIC.
(2) Fora forest technologies SLL.
(3) Instituto Universitario de Investigación Gestión Forestal Sostenible iuFOR.

Resumen

Los modelos de distribución de especies constituyen una potentísima herramienta de predicción sobre el área de distribución de cualquier taxón, con multitud de aplicaciones no sólo en el ámbito teórico de la ecología y la geobotánica sino también en otras disciplinas aplicadas.

Es sabido que a la hora de obtener modelos fiables es necesario contar con información de calidad que represente completamente el nicho ecológico de la especie, de lo contrario se pueden producir sesgos en la representación de la distribución potencial de las especies.

En el presente trabajo se ha utilizado la aplicación ModERFoRest para generar más de 50 modelos de distribución para analizar la influencia que puede tener el hecho de que el taxón de interés presente diversos subtaxones o rasgos adaptativos dentro del área de interés. Para ello se ha empleado como ejemplo una especie, *Pinus pinaster*, con una marcada variabilidad genética y dos subespecies reconocidas en la península ibérica

Los resultados indican que la calidad de la predicción del modelo de distribución de especies se ve notablemente mermada cuando se utiliza el conjunto de la subespecie, siendo particularmente notorio en el caso de la subespecie mediterránea (más de 25 puntos porcentuales de descenso en la sensibilidad), con mucha mayor variabilidad genética. Por último, se propone una serie de recomendaciones para el buen uso de la aplicación a la hora de predecir la distribución potencial de las especies forestales. .

Palabras clave

Autoecología paramétrica, software, modelización, nicho ecológico.

1. Introducción

Los modelos de distribución de especies (conocidos internacionalmente como SDM, según sus siglas en inglés) constituyen una potentísima herramienta de predicción sobre el área de distribución de cualquier taxón, con multitud de aplicaciones no sólo en el ámbito teórico de la ecología y la geobotánica sino también en disciplinas aplicadas como la restauración de ecosistemas, monitorización y detección temprana de especies invasoras, programas de migración asistida, también para tratar de entender los posibles efectos del cambio climático en la distribución potencial de las especies (FELICÍSIMO et al., 2011) y

como ayuda para elaborar estrategias de conservación y adaptación al cambio climático (GIL, 2017; GUISAN et al., 2013; GUISAN et al., 2017), entre otras aplicaciones.

En general, los SDM tienen una serie de ventajas que los convierten en una herramienta útil para la conservación de la biodiversidad:

- Basan sus predicciones en métodos estadísticos objetivos, lo que reduce los posibles sesgos en los datos de partida (CHEVALIER et al., 2022).
- Se pueden proyectar a diferentes escenarios climáticos futuros, lo que proporciona información de valor acerca de los posibles cambios en la distribución potencial de las especies.
- Son una herramienta muy útil para obtener información sobre la ecología y biogeografía de las especies, y poder evaluar el impacto de los factores ambientales y antropogénicos (GUISAN et al., 2017).

Sin embargo, a su vez, los SDM poseen una serie de limitaciones:

- Pueden verse afectados por la falta de información sobre las variables ambientales que determinan la distribución de las especies.
- Pueden ser difíciles de interpretar y aplicar en la práctica.
- Son muy dependientes de la calidad de la información, y de las posibles relaciones que existan en ella.

Por otro lado, cada vez hay más fuentes de datos accesibles sobre la distribución de especies a disposición de los investigadores y técnicos, lo que dota de medios suficientes para realizar con facilidad multitud de análisis ecológicos de gran transcendencia para la gestión forestal. Entre esas fuentes se pueden citar el Inventario Forestal Nacional (IFN), a punto ya de concluir su cuarta edición, así como el Sistema Global de Información sobre la Biodiversidad (GBIF, por sus siglas en inglés), que es una red internacional e infraestructura de datos financiada por los distintos gobiernos participantes. Si bien las bases de datos del IFN cuentan con una localización de las parcelas de campo con cada vez mayor precisión geográfica, lo que aumenta así su ya enorme potencia para generar valor, la información procedente de la red GBIF suele tener sesgos y errores, lo que puede llevar a enmascararlas predicciones de los modelos y los patrones que tienen que reflejar (VARELA et al., 2014).

Otras posibles fuentes de error o sesgo serían la utilización de datos con sesgo espacial, por ejemplo las localizaciones muestreadas cerca de vías de acceso, o las procedentes de colecciones de museos, cuyos especímenes no se muestrearon para representar la distribución de la especies, la modelización de datos con efecto frontera o límite administrativo, etc. En resumen, la importancia se puede resumir en que el resultado depende más de la calidad de los datos biológicos de entrada que del algoritmo utilizado (LOBO, 2008). En definitiva, un aspecto importante a la hora de obtener modelos fiables es la necesidad de contar con información de calidad tanto sobre la distribución de las especies (en concreto la localización), como para las variables ambientales. Esta información debe representar completamente el nicho ecológico de la especie, de lo contrario se pueden producir sesgos en la representación de la distribución potencial de las especies.

Sin embargo, un aspecto menos estudiado en el marco de los modelos de distribución de especies es la influencia que puede tener en la predicción el hecho de que el taxón analizado presente diversos subtaxones o rasgos adaptativos dentro del área de interés. Dado que estas metodologías se apoyan en la estabilidad (al menos a escala de cientos de años) de las relaciones entre el taxón y las componentes físicas del ámbito geográfico en el que se encuentra (PETERSON, 1999; GUISAN et al., 2006), la presencia de, por ejemplo, plasticidad fenotípica en algunas de sus poblaciones o de gran diversidad genética entre las distintas poblaciones de un mismo taxón puede conducir a que el proceso de modelización genere resultados insatisfactorios, ya que, de alguna manera, estaríamos intentando predecir el área de distribución de más de un "taxón" a la vez.

En el caso de *Pinus pinaster* Ait., una especie con dos subespecies reconocidas en la península Ibérica (la *atlantica* y la *mesogeensis*), las diferencias adaptativas a condiciones locales (su plasticidad ecológica) complican la aplicación de modelos de distribución. Para empezar, las subespecies tienen requisitos ecológicos distintos, lo que puede generar resultados imprecisos si no se tienen en cuenta adecuadamente las diferencias genéticas y ecológicas en los modelos de distribución (GONZÁLEZ et al., 2011; CHARDON et al., 2019, BIAOU et al., 2023). En segundo lugar, los estudios genéticos y ecológicos sobre la especie han establecido 28 regiones de procedencia, lo cual es un reflejo de la gran variabilidad adaptativa que presenta. Por último, este taxón ha sido utilizado profusamente en repoblaciones en España, tanto con fines productivos como protectores, lo que añade complejidad al uso de datos de presencia actual para modelizar su área potencial de distribución.

2. Objetivos

El objetivo general de este trabajo es evaluar la influencia que tienen en el desempeño de los modelos de distribución de especies las agrupaciones de la muestra de presencias actuales de *Pinus pinaster*, basándose en criterios genéticos y ecológicos. Como objetivos específicos se plantean los dos siguientes: (i) generar modelos predictivos específicos para las dos subespecies de *Pinus pinaster* en España y (ii) ofrecer recomendaciones sobre el uso adecuado de los algoritmos de modelización para predecir la distribución de especies forestales, considerando las particularidades genéticas y ecológicas de cada subespecie.

3. Metodología

3.a. Datos de presencia y agrupaciones territoriales

Los datos de presencia de la especie se han extraído de las bases de datos del Inventario Forestal Nacional (IFN), considerando aquellas parcelas con >90% de número de pies y de área basimétrica. En total se ha contado con 16054 localizaciones (*Figura 1*).

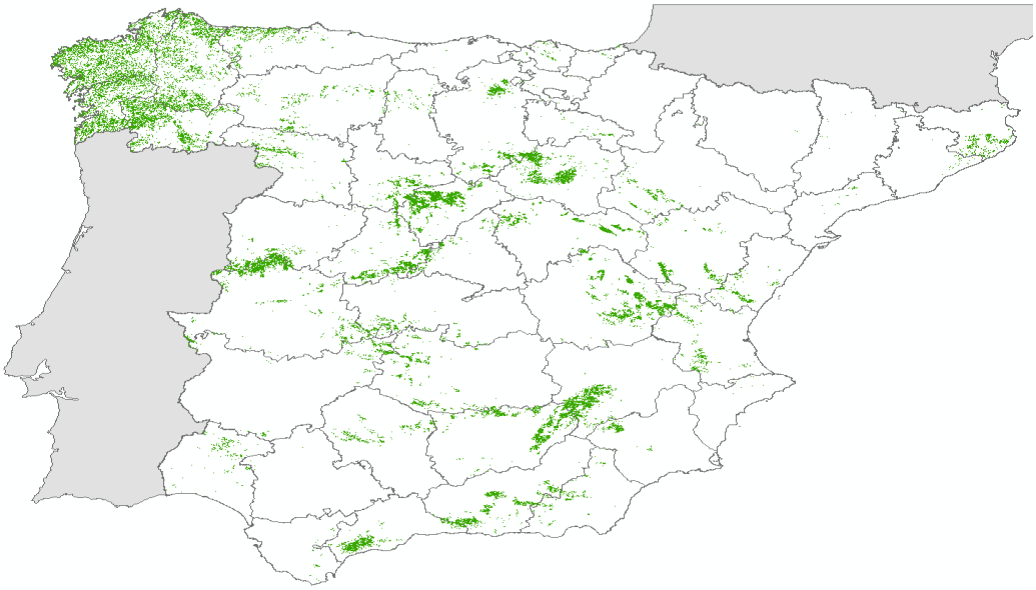

Figura 1. Distribución de la especie Pinus pinaster en la España peninsular.

Para la segmentación de dicha muestra pueden utilizarse diversos criterios, como las tipologías selvícolas (SERRADA et al., 2008), factores geoedáficos (COSTA et al., 1997) o regiones de procedencia (ALIA, 2010; GUADAÑO, 2019). En este trabajo se testaron dos segmentaciones a partir de las regiones de procedencia. Por un lado, se emplearon criterios genéticos, de manera independiente para las dos subespecies y para masas repobladas y naturales (Tabla 1 y Figura 2). Por otro, se siguieron criterios ecológicos basados en parámetros climáticos, igualmente de manera independiente para los dos subtaxones y para los dos orígenes de las masas (Tabla 2 y Figura 3).

Tabla 1. Agrupaciones genéticas de las presencias de Pinus pinaster en el territorio peninsular.

Subespecie	Natural/Repoblación	Grupo Genético	Nº Parcelas
Atlántica	Natural		5455
	Repoblación		148
Mediterránea	Natural	Centro	4673
		Este	1941
		Sur	2310
	Repoblación	Centro	400
	Este	37	
	Sur	1090	
		Total general	16054

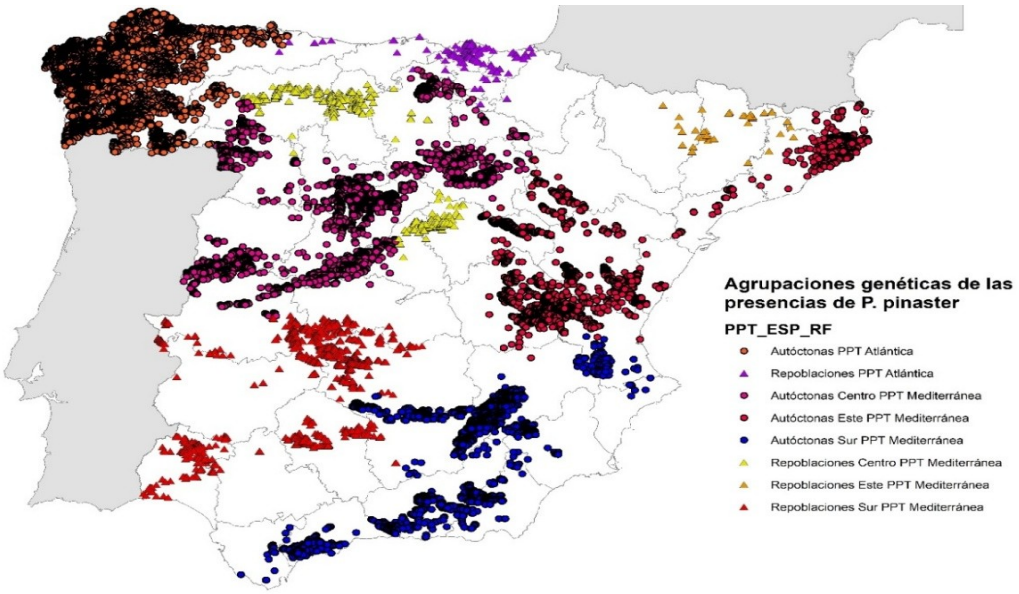

Figura 2. Distribución de las agrupaciones genéticas para la especie Pinus pinaster en la España peninsular.

Tabla 2. Agrupaciones ecológicas de las presencias de Pinus pinaster en el territorio peninsular ([1] Se corresponde con la región 20 de las regiones definidas por el método divisivo)

Subespecie	Natural/Repoblación	Grupo Genético	Nº Parcelas
Atlántica	Natural	Galicia Costera	4160
		Galicia Interior	1295
	Repoblación		148
Mediterránea	Natural	Cádiz	326
		Cataluña-Valencia-Murcia	500
		León	338
		Litoral Cataluña	527
		Málaga-Granada	551
		Meseta Norte y Cuenca	3034
		Meseta Sur	2421
		Sistema Ibérico	1109
		Valle del Tiétar	118
	Repoblación	Repo_DIV20 [1]	134
	Meseta Sur	758	
	Noreste	37	
	Páramos	266	
	Sierra Morena	332	
		Total general	16054

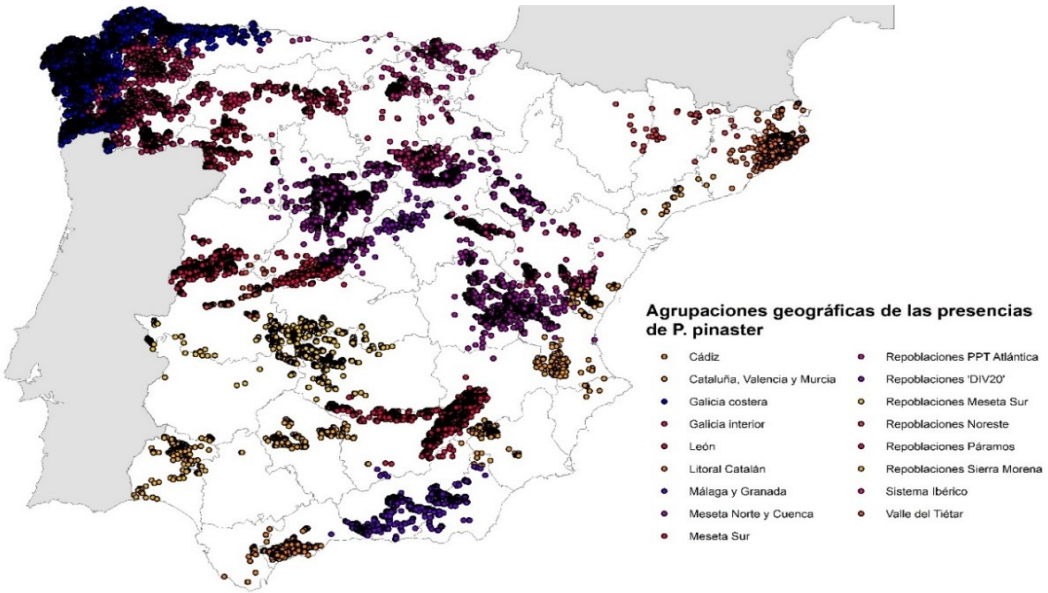

Figura 3. Distribución de las agrupaciones ecológicas para la especie Pinus pinaster en la España peninsular.

3.b. Variables climáticas

Se utilizaron dos conjuntos de variables climáticas para los modelos de distribución de especies. El primer conjunto (C1) incluye variables seleccionadas para un proyecto de mitigación de los efectos del cambio climático (Proyecto MEMECLIM, MITERD - FIC), que son la evapotranspiración potencial (ETP), la oscilación máxima (OSCXN) y la precipitación anual (PT). El segundo conjunto (C2) se compone de variables obtenidas en el proyecto RTA2014 de autoecología de especies forestales, que son el Índice Hídrico (IH), el Índice de Vernet (IVRNT), la precipitación anual (PT) y la temperatura media anual (TM). La definición y unidades de estas variables se detallan en la Tabla 3. Ambas series de datos se han calculado a partir del Atlas Climático Ibérico (AEMET, 2011).

Tabla 3. Variables y unidades de las variables climáticas empleadas en el análisis.

Acrónimo	Variable	Unidades	Definición
ETP	Evapotranspiración	mm	Suma de las 12 ETP mensuales según Thornthwaite
OSCXN	Oscilación máxima	ºC	Diferencia entre la temperatura máxima del mes más cálido y la mínima del mes más frío
PT	Precipitación total	mm	Suma de las 12 precipitaciones mensuales
IH	Índice hídrico	adim	Indicador síntesis de SUP, DEF y ETP según la fórmula IH=(100SUP-60DEF)/ETP*
IVRNT	Índice de Vernet	ºC/mm	(Vernet y Vernet, 1966)
TM	Temperatura media anual	ºC	Media de las 12 temperaturas medias mensuales
* Siendo SUP la suma de las diferencias P (precipitación)-ETP mensuales, siempre que P>ETP, y DEF la suma de las diferencias ETP-P mensuales, siempre que ETP>P			

3.c. Calibración y validación de modelos

Para cada una de las agrupaciones (genéticas y ecológicas) y los dos conjuntos de variables climáticas definidos se generó un modelo de distribución de la "especie" en todo el territorio peninsular español y Baleares. La herramienta utilizada para el proceso de modelización ha sido ModERFoRest (LÓPEZ-SENESPLEDA et al, 2023), una aplicación informática diseñada para apoyar la restauración forestal, especialmente en la elección de especies y su material forestal de reproducción adecuado. Además de modelizar la potencialidad del territorio para las especies, calcula variables climáticas y sintetiza datos edáficos. La aplicación, de uso libre, cuenta con una extensa base de datos sobre 19 taxones arbóreos y una comunidad vegetal, lo que le confiere fiabilidad y potencia. Desarrollada por el INIA-CSIC y colaboradores, ModERFoRest incluye tres módulos principales: MDES (Modelos de Distribución de Especies), AERF (Adecuación para la Restauración Forestal) y SEF (Similitud entre Estaciones Forestales), que emplean dos algoritmos: el IPF (Índice de Potencialidad Factorial) (RUBIO Y SÁNCHEZ PALOMARES, 2006) y el IPA (Índice de Potencialidad Aditivo) (ALONSO PONCE et al., 2010b, 2010a).

Se utilizó el algoritmo IPA para modelizar la distribución de la especie, validando los resultados mediante una submuestra de presencias y realizando 30 réplicas por modelo. En el proceso se estableció el porcentaje de submuestras de validación (Tabla 4) en función del número de parcelas de presencia disponible para cada agrupación (ver Tablas 1 y 2), tratando de asegurar que se alcanza un mínimo de presencias para que los modelos sean lo más fieles posible (PROOSDIJ et al., 2016).

Tabla 4. Porcentajes de la submuestra de validación

Número de parcelas	% validación
>4000	95
3000-4000	85
1000-3000	80
500-1000	60
<500	30

La construcción de la curva de Boyce-Hirzel (LÓPEZ-SENESPLEDA et al., 2023) es crítica en el proceso de modelización. Esta curva es clave para determinar si el modelo es mejor que una predicción al azar, separando en clases (extramarginal, marginal, adecuada, óptima) las áreas predichas por el modelo para la especie. Además, se usaron indicadores como la sensibilidad y el test de Kolmogorov-Smirnov con el fin de evaluar la semejanza entre las muestras de calibración y de validación para distribuciones empíricas. La aplicación emplea la generalización robusta de dicho test (GARCÍA PÉREZ, 2005; MEE, 1990). Al finalizar, el programa genera cinco capas ráster de idoneidad, índices de marginalidad (GÓMEZ-SANZ,

2019) para cada variable ambiental, y un informe resumen en PDF, que incluye los resultados del modelo y su validación. Estos datos permiten realizar un análisis comparativo de los modelos generados y determinar su eficacia.

4. Resultados

Después de un exhaustivo proceso de modelización se han seleccionado un total de 50 modelos, correspondientes a cada grupo definido según los dos criterios (genético y ecológico). Estos modelos se han desarrollado para los dos conjuntos de variables climáticas establecidos, además de otros dos modelos para cada subespecie en su conjunto. En los siguientes apartados se presentan los valores de sensibilidad obtenidos, tanto para la submuestra de calibración como para la de validación, así como el número de réplicas que pasaron la prueba de Kolmogorov-Smirnov. También se muestra una valoración de la calidad de la curva de Boyce-Hirzel para los dos conjuntos de variables climáticas considerados.

4.a. Resultados por grupos genéticos:

La Tabla 5 presenta los resultados de la validación de los modelos obtenidos para los grupos definidos según el criterio genético. Los valores de sensibilidad para la muestra de validación varían, alcanzando entre el 76.7% y el 94.5% en el caso del conjunto de variables C1, con un promedio de 86.5%, y entre el 77.4% y el 93.8%, con un promedio de 86.5%, para las variables C2. Como es de esperar, la sensibilidad de validación siempre es inferior a la de calibración, aunque, salvo en el grupo PPTM_REPO_ESTE, la diferencia es menor a dos puntos porcentuales. Las curvas de Boyce-Hirzel (BH) fueron excelentes en el 62.5% de los casos y normales en el 37.5%, lo que indica el buen desempeño de todos los modelos y su diferencia significativa respecto a un modelo trivial al azar.

Tabla 5. Resumen de la evaluación de los modelos obtenidos para los grupos territoriales establecidos según el criterio genético. C1: modelización utilizando las variables climáticas C1; C2: modelización utilizando las variables climáticas C2. Sens: sensibilidad. K-S: número de réplicas (de las 30 ejecutadas) que han pasado la prueba de Kolomogorov-Smirnof (igualdad de distribuciones). Curva BH: calidad de la curva de validación obtenida (E: excelente; N: normal).

Grupo	Parc	% validac.	C1				C2			
			Sens. Cal.	Sens. Val.	K-S	Curva BH	Sens. Cal.	Sens. Val.	K-S	Curva BH
PPTA_AUTO	5455	95	95.9	94.5	30	E	92.9	91.8	30	E
PPTA_REPO	148	30	93.9	91.0	30	E	93.4	92.4	30	N
PPTM_AUTO_CENTRO	4673	95	90.1	87.8	28	N	88.9	87.3	30	N
PPTM_AUTO_ESTE	1941	80	87.1	86.1	30	N	86.6	86.6	30	E
PPTM_AUTO_SUR	2310	80	83.3	82.2	30	E	78.1	77.4	29	E
PPTM_REPO_CENTRO	400	30	94.2	93.3	30	E	94.5	93.8	29	E
PPTM_REPO_ESTE	37	30	91.2	76.7	29	N	89.7	81.8	30	N
PPTM_REPO_SUR	1090	80	91.7	89.4	29	E	88.4	86.5	30	E

4.b. Resultados por grupos ecológicos:

La Tabla 6 muestra los resultados de la validación de los modelos obtenidos para los grupos definidos según el criterio ecológico. Los valores de sensibilidad para la muestra de validación varían entre el 80.0% y el 97.2% para el conjunto de variables C1, con un valor medio del 90.1%, y entre el 82.4% y el 94.7%, con un valor medio del 90.5%, para las variables C2. Los resultados son muy similares para ambos conjuntos de variables. Al igual que en la agrupación genética, la sensibilidad de validación es solo ligeramente inferior a la de calibración. Sin embargo, el grupo PPTM_REPO_NE muestra una mayor diferencia en sensibilidad (13.5 puntos para C1 y 7.6 para C2), lo que probablemente se debe a la muestra extremadamente menor (37) utilizada para generar los modelos.

Tabla 6. Resumen de la evaluación de los modelos obtenidos para los grupos territoriales establecidos según el criterio ecológico. C1: modelización utilizando las variables climáticas C1; C2: modelización utilizando las variables climáticas C2. Sens: sensibilidad. K-S: número de réplicas (de las 30 ejecutadas) que han pasado la prueba de Kolomogorov-Smirnof (igualdad de distribuciones). Curva BH: calidad de la curva de validación obtenida (E: excelente; N: normal).

Grupo	Parc	% validac.	C1				C2			
			Sens. Cal.	Sens. Val.	K-S	Curva BH	Sens. Cal.	Sens. Val.	K-S	Curva BH
CADIZ	326	30	88.1	86.2	28	E	88.3	86.2	29	E
CAT_VAL_MUR	500	30	92.0	90.6	30	E	88.9	88.0	29	N
GAL_COST	4160	95	97.0	95.6	28	E	94.1	92.7	28	E
GAL_INT	1295	80	92.3	91.0	26	E	93.0	91.7	29	N
LEON	338	30	98.2	97.2	24	E	96.4	94.5	27	E
LIT_CAT	527	60	94.5	93.5	25	N	94.0	92.5	29	E
MA_GR	551	60	94.0	91.2	27	N	93.5	92.1	27	N
MESET_NORTE_CU	3034	85	87.1	86.6	30	E	91.2	90.4	30	E
MESET_SUR	2421	80	82.3	80.9	27	N	89.5	89.0	28	E
PPTA_REPO	148	30	93.9	91.0	30	E	93.4	92.4	28	N
REPO_DIV20	134	30	95.1	91.8	27	N	94.7	91.4	25	E
REPO_MESET_SUR	758	60	90.7	89.7	29	E	90.1	91.6	28	E
REPO_NE	37	30	93.5	80.0	29	N	90.0	82.4	28	N
REPO_PARAM	266	30	95.9	94.3	25	E	95.2	94.7	25	N
REPO_SMORENA	332	30	91.0	90.8	29	E	93.2	92.5	26	E
SIST_IBER	1109	80	94.3	92.8	29	E	89.8	90.0	30	E
TIETAR	118	30	91.2	88.6	27	E	88.2	86.5	27	N

4.c. Resultados por subespecies:

Con el fin de evaluar el desempeño de los diferentes modelos, se realizaron modelizaciones considerando el conjunto total de observaciones, separadas únicamente por subespecies. Los resultados obtenidos, que se presentan en la Tabla 7 y la Figura 3, muestran un desempeño significativamente peor del modelo para la subespecie mediterránea en comparación con la subespecie atlántica.

Tabla 7. Resumen de la evaluación de los modelos obtenidos para las dos subespecies (PPTA: Pinus pinaster ssp. atlantica; PPTM: Pinus pinaster ssp. mesogeensis). C1: modelización utilizando las variables climáticas C1; C2: modelización utilizando las variables climáticas C2. Sens: sensibilidad. K-S: número de réplicas (de las 30 ejecutadas) que han pasado la prueba de Kolomogorov-Smirnof (igualdad de distribuciones). Curva BH: calidad de la curva de validación obtenida (E: excelente; N: normal).

Grupo	Parc	% validac.	C1				C2			
			Sens. Cal.	Sens. Val.	K-S	Curva BH	Sens. Cal.	Sens. Val.	K-S	Curva BH
PPTA	5603	95	95.7	94.2	29	E	91.6	90.2	30	N
PPTM	10451	95	80.0	78.9	29	E	66.6	66.4	29	N

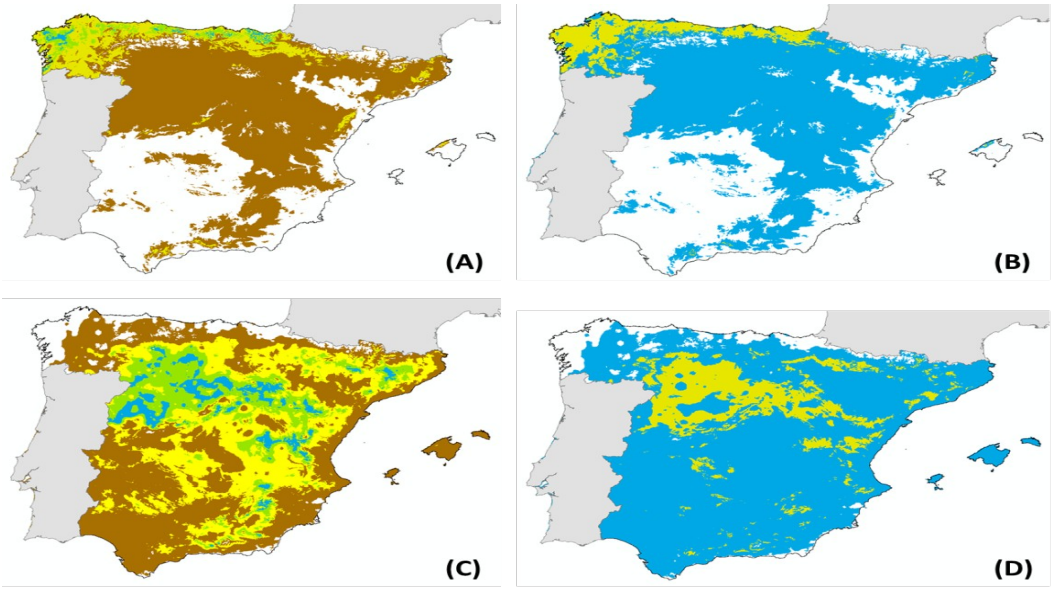

Figura 4. Resultado de los modelos para cada subespecie y sus desviaciones típicas: (A) modelo de aptitud para la subespecie atlántica, (B) desviación típica del modelo para la subespecie atlántica, (C) modelo de aptitud para la subespecie mesogeensis, (D) desviación típica del modelo para la subespecie mesogeensis. Obsérvese en las figuras B y D que los valores de desviación típica son mayores donde el valor medio del modelo es mayor, pero sin embargo, siguen siendo inferiores al 10% (color azul: desviación inferior al 1%, color verde: desviación entre el 1-10%).

4.d. Comparando resultados:

En las siguientes figuras se presentan las comparaciones de la sensibilidad de la muestra de validación de las distintas agrupaciones con la obtenida para la subespecie en su conjunto.

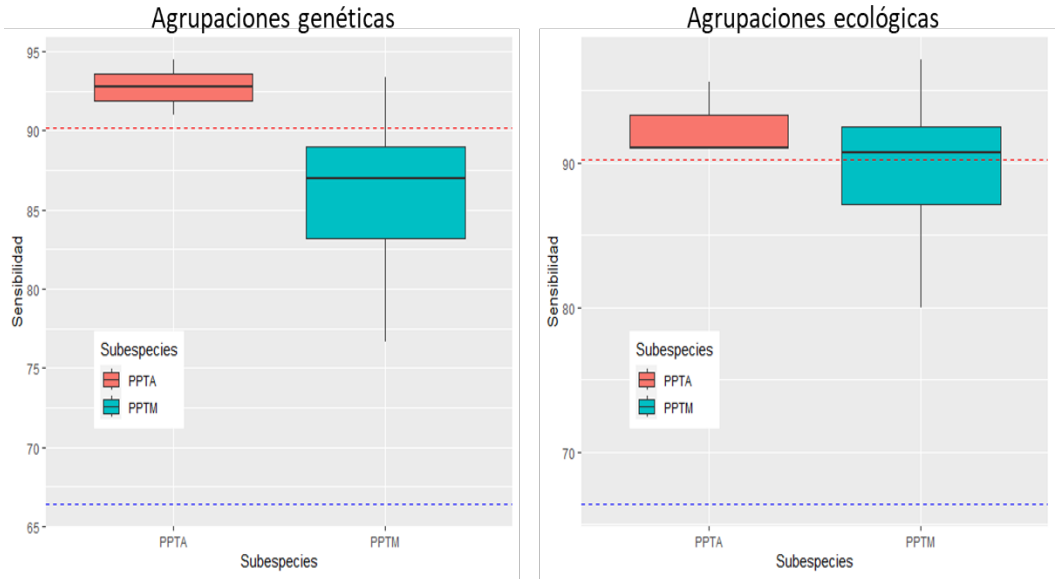

Figura 5. Sensibilidad de validación por agrupaciones genéticas y ecológicas ysubespecies para el conjunto de variables climáticas C2. Las líneas discontinuas representan la sensibilidad de validación del modelo obtenido con todas las observaciones disponibles de cada subespecie, sin considerar grupos (rojo, Pinus pinaster ssp. atlantica; azul, Pinus pinaster ssp. mesogeensis).

De ambas figuras se deduce que el desempeño de los modelos generados a partir de las agrupaciones definidas es significativamente mejor que el de la subespecie respectiva en conjunto. Particularmente llamativo resulta el caso de la subespecie mediterránea, con más de 20 puntos de diferencia y casi 25 puntos para las agrupaciones genéticas y ecológicas, respectivamente. En la subespecie atlántica las diferencias no son tan notables, pero son igualmente significativas (93.2 y 93.5%, para las agrupaciones genética y ecológica, respectivamente, frente a 90.2% de la subespecie en su conjunto).

Además, se analizó el diferente desempeño de la capacidad predictiva de los modelos según fuera el origen de la masa (repoblada o natural, Figura 5A, según el conjunto de variables empleadas (C1 o C2, Figura 5B) o según el tamaño de la muestra (menos o más de 1000 observaciones, Figura 5C). De todas estas comparaciones se deduce que no existen diferencias significativas en ninguno de los casos

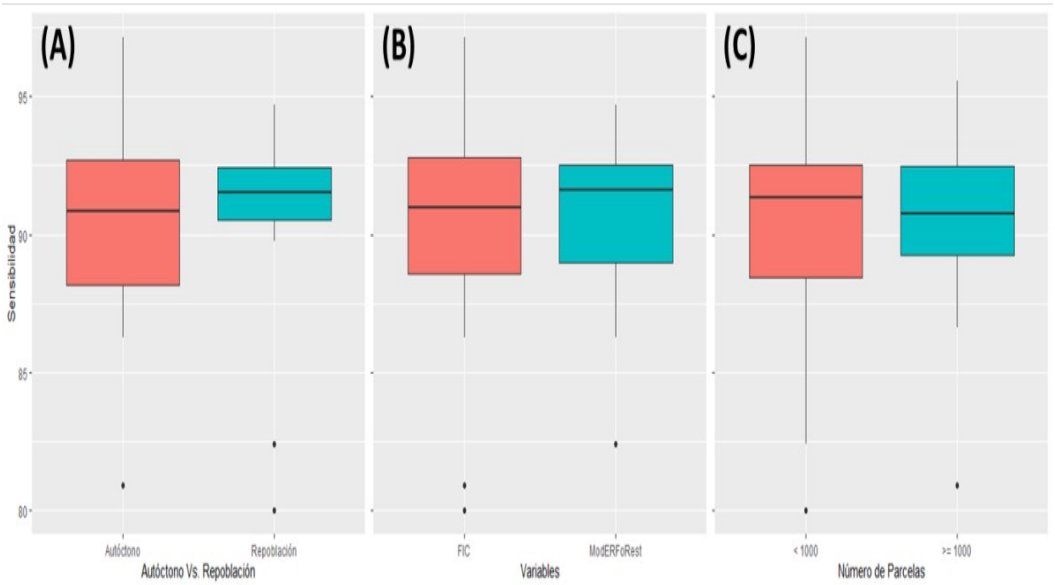

Figura 6. Sensibilidad de validación: (A) por origen de la masa, usando variables C2; (B) por conjunto de variables C1: FIC y C2: ModERFoRest, y (C) por tamaños muestrales usando variables C2.

5. Discusión

En este trabajo se compara el desempeño que muestran los SDM para predecir el rango de distribución de P. pinaster cuando se utiliza todo el rango de distribución peninsular de la especie con cuando se utilizan subpoblaciones basadas en criterios genéticos o ecológicos. La asunción de que todas las poblaciones de una especie han de responder de manera equivalente al clima ha sido analizada en pocos estudios pero demuestran, como en nuestro trabajo, que incorporar la variación intraespecífica mejora la capacidad predictiva de los modelos, tanto en taxones vegetales (BIAOU et al., 2023; CHARDON et al., 2019) como animales (GONZALEZ et al., 2011). Las dos agrupaciones utilizadas en este trabajo, como se ha explicado, se basan en criterios genéticos y ecológicos y han demostrado una mejor capacidad predictiva que los modelos construidos con todo el rango de cada una de las subespecies peninsulares de *P. pinaster*, por lo que se alinean con los estudios mencionados. Sin embargo, la manera en que se incorpore dicha variación es muy relevante, ya que divisiones de la población basadas en criterios meramente geográficos no parecen aportar grandes mejoras a la predicción que se obtiene utilizando el conjunto de la población.

Dado que las interrelaciones tanto entre el número de predictores del modelo como la complejidad y la naturaleza de las interacciones entre especie y ambiente influyen en el tamaño de muestra necesario para obtener una buena predicción mediante SDM, es muy complicado proponer un umbral mínimo para el número de observaciones necesarias (MOUDRÝ et al., 2024). En nuestro caso, la muestra disponible para cada agrupación definida es muy amplia (las únicas excepciones son PPTM_REPO_ESTE para el criterio genético, y REPO_NE para el ecológico), y aunque se ha sometido a los modelos a un exigente proceso de validación, con porcentajes muy altos de observaciones reservadas para tal propósito, no se han encontrado diferencias en el desempeño del modelo según el tamaño de la muestra, probablemente por estar muy por encima del umbral mínimo necesario,

que en algunos casos puede llegar a ser tan bajo como 10 para obtener predicciones útiles (SOULTAN & SAFI, 2017).

Así, a la vista del comportamiento de los distintos modelos, y en función de las agrupaciones realizadas, se pueden hacer una serie de recomendaciones de uso de la aplicación y sus algoritmos, cara a la obtención de mejores resultados y su interpretación. Estas son:

- Dividir los datos de presencias en el caso de especies de extensa distribución, o de nicho amplio, mediante algún tipo de criterio ecológico o genético. Esta división de los datos suele, en general, mejorar la precisión de las predicciones y el desempeño de los modelos.
- Modificar los porcentajes de las submuestras de validación, en función del número de datos de presencia disponibles, con el fin de mejorar el rendimiento del programa. Se recomiendan los porcentajes sugeridos en la tabla 4 al haber sido probados con éxito por los autores.
- Por último, contrastar la información generada con las segmentaciones, frente a la que se puede obtener del conjunto total de datos sin dividir, y comprobar las mejoras en la elaboración de los modelos.

6. Conclusiones

En este estudio se ha analizado la influencia de la variabilidad genética y los rasgos adaptativos dentro de una especie en la generación de modelos de distribución, utilizando como ejemplo *Pinus pinaster*, que presenta dos subespecies reconocidas. Se empleó el algoritmo de índice de potencialidad aditivo (IPA) en la herramienta ModERFoRest y dos conjuntos de variables climáticas. Las principales conclusiones son las siguientes:

1. Se ajustaron 50 modelos de distribución, siendo todos significativamente diferentes a un modelo aleatorio mediante la validación con las curvas de Boyce-Hirzel.
2. ModERFoRest facilitó un proceso de modelización y validación rápido y eficiente.
3. La calidad de la predicción disminuye al usar la subespecie en su conjunto, especialmente en la subespecie mediterránea, con una disminución de más del 25% en la sensibilidad. La subespecie atlántica mostró una menor caída, de aproximadamente 3 puntos.
4. La capacidad predictiva fue mayor cuando se usaron agrupaciones genéticas o ecológicas, siendo más evidente en el caso de las genéticas.
5. No se encontraron diferencias significativas según el origen de la masa, las variables climáticas usadas o el tamaño de la muestra.

7. Agradecimientos

Este trabajo ha sido posible realizarlo gracias al contrato entre el Ministerio para la Transición Ecológica y el Reto Demográfico y Fora Forest Technologies SLL, y la colaboración del ICIFOR.

8. Bibliografía

AEMET; IMP; 2011. Atlas climático ibérico/Iberian climate atlas. 79 pp.

ALIA, R.; 2010. Regiones de procedencia de especies forestales en España. Organismo Autónomo de Parques Nacionales. 364pp.

ALONSO PONCE, R.; ÁGUEDA, B.; ÁGREDA, T.; MODREGO, M.P.; ALDEA, J.; MARTÍNEZ-PEÑA, F.; 2010a. Un modelo de potencialidad climática para la trufa negra (*Tuber melanosporum*) en Teruel (España). For. Syst. 19, 208–220.

ALONSO PONCE, R.; ÁGUEDA, B.; ÁGREDA, T.; MODREGO, M.P.; ALDEA, J.; FERNÁNDEZ-TOIRÁN-Toirán, L.M.; MARTÍNEZ-PEÑA, F.; 2011. Rockroses and Boletus edulis ectomycorrhizal association: Realized niche and climatic suitability in Spain. Fungal Ecol. 4, 224–232. doi:10.1016/j.funeco.2010.10.002.

ALONSO PONCE, R.; LÓPEZ-SENESPLEDA, E.; SÁNCHEZ PALOMARES, O.; 2010b. A novel application of the ecological field theory to the definition of physiographic and climatic potential areas of forest species. Eur. J. For. Res. 129, 119–131. doi: 10.1007/s10342-008-0254-2.

BIAOU, S.; GOUWAKINNOU, G.N.; NOULÈKOUN, F.; SALAKO, K.V.; KPOVIWANOU, J.M.R.H.; HOUEHANOU, T.D.; BIAOU, H.S.S.; 2023. Incorporating intraspecific variation into species distribution models improves climate change analyses of a widespread West African tree species (Pterocarpus erinaceus Poir, Fabaceae). Glob. Ecol. Conserv., 45, 1-16.

BOYCE, M.S.; VERNIER, P.R.; NIELSEN, S.E.; SCHMIEGELOW, F.K.A.; 2002. Evaluating resource selection functions. Ecol. Modell. 157, 281–300.

CHARDON, N.I.; PIRONON, S.; PETERSON, M.L.; DOAK, D.F.; 2019. Incorporating intraspecific variation into species distribution models improves distribution predictions, but cannot predict species traits for a wide-spread plant species. Ecography, 42, 1-15.

CHEVALIER, M., ZARZO-ARIAS, A.; GUÉLAT, J.; MATEO, R.G.; GUISAN, A.; 2022. Accounting for niche truncation to improve spatial and temporal predictions of species distributions. Front. Ecol. Evol., 10.

COSTA, M.; MORLA, C.; SAINZ, H.; (Eds.); 1997. Los bosques ibéricos. Una interpretación geobotánica. Ed. Planeta. 598 pp.

FELICISIMO, A.M.; MUÑOZ, J.; VILLALBA, C.J.; MATEO, R.G.; 2011. Impactos, vulnerabilidad y adaptación al cambio climático de la biodiversidad española. 2. Flora y vegetación. Madrid: Oficina Española de Cambio Climático. Ministerio de Medio Ambiente y Medio Rural y Marino. 550 pp.

FIELDING, A.H.; BELL, J.; 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv., 24 (1), 38-49.

GANDULLO, J.; SÁNCHEZ PALOMARES, O.; 2000. Programa Pinares. Montes 60, 5–9.

GANDULLO, J.M.; SÁNCHEZ PALOMARES, O.; 1994. Estaciones ecológicas de los pinares españoles. ICONA, Madrid.

GARCÍA PÉREZ, A.; 2005. Métodos avanzados de estadística aplicada. Métodos robustos y de remuestreo. UNED, Madrid.

GIL, G.E.; 2017. Modelos de distribución y su aplicación en la conservación de mamíferos terrestres sudamericanos. Tesis Doctoral. Facultad de Ciencias. UAM. 287 pp.

GÓMEZ-SANZ, V.; 2019. Site-scale ecological marginality: Evaluation model and application to a case study. Ecol. Modell. 408. doi:10.1016/j.ecolmodel.2019.108739.

GÓMEZ-SANZ, V.; SERRADA, R.; ELENA ROSELLÓ, R.; FERNÁNDEZ YUSTE, J.A.; GARCÍA VIÑAS, J.I.; LÓPEZ LEIVA, C.; 2014. Medio físico de rodales selectos de especies de forestales: Pinus sylvestris L. ECOGESFOR.

GONZÁLEZ, S.C.; SOTO-CENTENO, J.A.; REED, D.L.; 2011. Population distribution models: species distributions are better modeled using biologically relevant data partitions. Ecology, 11, 20.

GUADAÑO, C.; 2019. Actualización de las regiones de procedencia y desarrollo de Indicadores del estado de conservación y uso de los recursos genéticos para tres especies del género Pinus. TFM. ETSIA. UVa. 82 pp.

GUISAN, A.; BROENNIMANN, O.; ENGLER, R.; VUST, M.; YOCCOZ, N.G.; LEHMANN, A.; ZIMMERMANN, N.E.; 2006. Using niche-based models to improve the sampling of rare species. Conserv. Biol., 20, 501-511.

GUISAN, A.; TINGLEY, R.; BAUMARTNER, J.B.; NAUJOKAITIS-LEWIS, I.;, SUTCLIFE, P.R.; TULLOCH, A.I.T.; REGAN, T.J.; BROTONS, L.; MCDONALD-MADDEN, E.; MANTYKA-PRINGLE, C.; MARTIN, T.G.; RHODES, J.R.; MAGGINI, R.; SETTERFIELD, S.A.; ELITH, J.; SCHWARTZ, M.W.; WINTLE, B.A.; BROENNIMANN, O.; AUSTIN, M.; FERRIER, S.; KEARNEY, M.R.; POSSINGHAM, H.P.; BUCKLEY, Y.M; 2013. Predicting species distributions for conservation decisions. Ecol. Lett., 16 (12), 1424-1435.

GUISAN, A.; THUILLER, W.; ZIMMERMANN, N.E.; 2017. Habitat Suitability and

Distribution Models. With Applications in R. Cambridge University Press. 462 pp.

HIRZEL, A.; LE LAY, G.; HELFER, V.; RANDIN, C.F.; GUISAN, A.; 2006. Evaluating the ability of hábitat suitability models to predict species presences. Ecol. Modell. 199, 142–152.

LOBO, J.M.; 2008. More complex distribution models or more representative data? Biodivers. Inform., 5, pp 14-19.

LÓPEZ-SENESPLEDA, E.; ALONSO PONCE, R.; RUIZ-PEINADO, R.; GÓMEZ, V.; SERRADA, R.; MONTERO, G.; 2023. ModERFoRest: "Modelling Environmental Requirements for Forest Restoration". A new software for assessing the environmental performance of forest species. For. Syst., 32 (1), eRC01.

LÓPEZ-SENESPLEDA, E.; MONTERO, G.; RUIZ-PEINADO, R.; ALONSO PONCE, R.; SERRADA, R.; SÁNCHEZ PALOMARES, O.; 2018. Cincuenta años de autoecología forestal paramétrica en España. Foresta 70, 40–47.

MEE, R.W.; 1990. Confidence-Intervals for Probabilities and Tolerance Regions Based on a Generalization of the Mann-Whitney Statistic. J. Am. Stat. Assoc. 85, 793–800.

MOUDRÝ, V.; BAZZICHETTO, M.; REMELGADO, R.; DEVILLERS, R.; LENOIR, J.; MATEO, R. G.; LEMBRECHTS, J. J.; SILLERO, N.; LECOURS, V.; CORD, A. F.; BARTÁK, V.; BALEJ, P.; ROCCHINI, D.; TORRESANI, M.; ARENAS-CASTRO, S.; MAN, M.; PRAJZLEROVÁ, D.; GDULOVÁ, K.; PROŠEK, J.; … ŠÍMOVÁ, P.; 2024. Optimising occurrence data in species distribution models: sample size, positional uncertainty, and sampling bias matter. *Ecography*, *2024*(12), e07294. https://doi.org/https://doi.org/10.1111/ecog.07294

PETERSON, A.T.; 1999. Conservatism of ecological niches in evolutionary time. Sci., 285, 1265-1267.

PROOSDIJ, A.S.J.; SOSEF, M.S.M.; WIERINGA, J.J.; RAES, N.; 2016. Minimum required number of specimen records to develop accurate species distribution models. Ecography, 39, 542-552.

PHILLIPS, SJ.; DUDIK, M.; SCHAPIRE, R.E.; 2004. A máximum entropy approach to species distribution modeling. In: Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004.

SERRADA, R.; MONTERO, G.; REQUE, J.A.; 2008. Compendio de Selvicultura Aplicada en España. INIA-MICIN. 1178 pp.

SOULTAN, A.; SAFI, K.; 2017. The interplay of various sources of noise on reliability of species distribution models hinges on ecological specialisation. *PLoS ONE*, *12*(11), e0187906. https://doi.org/10.1371/journal.pone.0187906

RUBIO, A.; SÁNCHEZ PALOMARES, O.; 2006. Physiographic and climatic potential areas for Fagus sylvatica L. based on habitat suitability indicator models. Forestry, 79, 439–451.

VARELA, S.; MATEO, R.G.; GARCÍA-VALDÉS, R.; FERNÁNDEZ-GONZÁLEZ, F.; 2014. Macroecología y ecoinformática: sesgos, errores y predicciones en el modelado de distribuciones. Ecosistemas 23(1): 46-53.

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

WU, H.; SHARPE, P.J.H.; WALKER, J.; PENRIDGE, L.K.; 1985. Ecological field theory: a spatial analysis of resource interference among plants. Ecol. Modell. 29, 215–243.