

ECOINFORMÁTICA APLICADA A LA  
CONSERVACIÓN: SIMULACIÓN DE EFECTOS DEL  
CAMBIO GLOBAL EN LA DISTRIBUCIÓN DE LA  
FLORA DE ANDALUCÍA

BLAS M. BENITO DE PANDO

Licenciado en Biología



Departamento de Botánica

Facultad de Ciencias

Universidad de Granada

14 de septiembre de 2009

Editor: Editorial de la Universidad de Granada  
Autor: Blas M. Benito de Pando  
D.L.: GR. 3482-2009  
ISBN: 978-84-692-6380-8

*Ecoinformática Aplicada a la Conservación: Simulación de Efectos del  
Cambio Global en la Distribución de la Flora de Andalucía*

© Blas BP 2009

Autor

BLAS M. BENITO DE PANDO  
Investigador Predoctoral  
Departamento de Botánica  
Universidad de Granada

Directores

DR. JULIO PEÑAS DE GILES  
Profesor Titular de Universidad  
Departamento de Botánica  
Universidad de Granada

DR. JUAN LORITE MORENO  
Profesor Contratado Doctor  
Departamento de Botánica  
Universidad de Granada

en Granada

a 14 de septiembre de 2009

TESIS DOCTORAL  
DEPARTAMENTO DE BOTÁNICA  
UNIVERSIDAD DE GRANADA

ECOINFORMÁTICA APLICADA A LA  
CONSERVACIÓN: SIMULACIÓN DE EFECTOS DEL  
CAMBIO GLOBAL EN LA DISTRIBUCIÓN DE LA  
FLORA DE ANDALUCÍA

Memoria presentada por  
el Licenciado en Biología  
BLAS M. BENITO DE PANDO  
para aspirar al Grado de Doctor  
por la Universidad de Granada

**Fdo. Blas Manuel Benito de Pando**

DIRECTORES

**Fdo. Dr. Julio Peñas de Giles**

**Fdo. Dr. Juan Lorite Moreno**

14 de septiembre de 2009



Dedicado a vosotros, mi familia,  
porque os quiero.



*La belleza de la Tierra  
ha sido esculpida por el cambio  
el cambio suave, casi imperceptible,  
y el cambio repentino y violento.  
En el Cosmos no hay lugar  
que esté a salvo del cambio.*

– CARL SAGAN (1934-1996)

## AGRADECIMIENTOS

---

A mi pequeña Ana, me ha llevado de la mano estos últimos años, y con ella quisiera ir de la mano hasta el final. Ha vivido a mi lado el esfuerzo que ha supuesto esta Tesis Doctoral y me ha animado con su sonrisa cuando faltaban las fuerzas. Con nuestros pequeños Chispa y Bola hemos formado nuestra diminuta *familia del metro cuadrado*. Ana, te quiero.

Mi madre, Genoveva, se dejó la piel para traerme a este lugar y ayudarme a buscar un futuro. Su fuerza y sus principios están conmigo en todo momento, y su apoyo, al otro lado del teléfono, allí en San Fernando, siempre que lo necesito. Mamá, me has dado alas para que pueda volar, y siempre te lo agradeceré.

A mi hermano Jaime. el tiempo y la distancia han afianzado el vínculo que me une con él.

Debo el mayor agradecimiento a Julio Peñas. Un día hace años, un profesor de botánica al que no conocía me dio la oportunidad de trabajar con él, sin sospechar que estaba cambiando mi vida. Julio, mi jefe, me abrió de ese modo las puertas al mundo de la investigación, una ilusión de la infancia convertida en realidad. Julio, mi amigo, comprendió mi vocación, y me dio todas las facilidades para avanzar por el camino que me ha traído hasta aquí. Julio, siempre te estaré agradecido.

A Juan Lorite, un gran amigo que siempre está ahí, dispuesto a dejar lo que está haciendo para escuchar y ayudar. Su trato a los demás es un ejemplo para mí, y la pasión



que siente por su trabajo, realmente contagiosa. Me siento afortunado de tenerlo cerca.

A José Ángel, que desde la lejana Argentina ha estado apoyándome en los últimos meses de este trabajo.

A Jorge, un hermano para mí, porque una gran parte del camino la hicimos juntos. Añoro las largas charlas de madrugada de dos apasionados de la biología, y los fantásticos días de escalada que hemos disfrutado.

A Curro, Ramón y Raúl, mis amigos en el CEAMA y un equipo de trabajo fantástico, porque me han ayudado a crecer.

A los recién doctorados del departamento, Eva y Diego. Han sido una referencia importante para mí, por su profesionalidad y calidad humana. Además del debido agradecimiento, les deseo la misma suerte que quiero para mí en el camino que tenemos por delante.

A mis compañeros del departamento y el herbario, un excelente grupo humano lleno de ejemplos a seguir, porque veo la pasión en sus ojos cuando hablan de nuestras queridas plantas, y aunque no lo crean, algo me contagian.

A la Comunidad de Software Libre, todos aquellos que han contribuido con su esfuerzo desinteresado, inteligencia y líneas de código, a programar las herramientas gracias a las que este trabajo ha salido a flote. Tomando su ejemplo, podemos hacer un mundo mejor.

A la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía, que ha hecho posible esta tesis financiando el Proyecto de Excelencia *Conservación de flora endémica y amenazada de hábitats frágiles: zonas áridas y altas montañas de Andalucía* (RNM 1067).

A todos, gracias.

## PUBLICACIONES

---

Algunos de los capítulos e ideas expuestos en esta tesis han aparecido en las siguientes publicaciones:

Benito B, Peñas J (2007). Aplicación de modelos de distribución de especies a la conservación de la biodiversidad en el sureste de la Península Ibérica. *Geofocus* 7. 100-109

Benito B, Peñas J (2007). Los modelos de distribución en el trabajo de campo: localización de nuevas poblaciones de especies amenazadas. *IG+* 8, 8-10

Gutiérrez L, del Río J, Navarro FB, Lorite J, Benito B, Peñas J (2007). Novedades sobre la flora amenazada de las zonas áridas interiores de Granada (Hoyas de Guadix y Baza). *La-gascalía* 27, 407-415

Benito B, Peñas J (2008). Greenhouses, land use change, and predictive models: MaxEnt and Geomod working together. En *Modelling Environmental Dynamics: Advances in Geomatic Solutions*. 297-317. Editores: Martin Paegelow y María Teresa Camacho Olmedo. Editorial Springer.

Peñas J, Benito B, Lorite J, Morales MC, Martínez-Ortega MM (2008). *Astragalus peregrinus* subsp. *warionis* and *A. guttatus* (Fabaceae) New for the Iberian Peninsula. *Phyton* 48, 155-167

Benito B, Martínez-Ortega MM, Muñoz LM, Lorite J, Peñas J (2009). Assessing extinction-risk of endangered plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses. *Biodiversity and Conservation* . 18(9), 2509-2520



## RESUMEN

---

La crisis de la biodiversidad que afecta a la Tierra, conducida por el Cambio Global, requiere respuestas coordinadas a escala planetaria, pero las acciones inmediatas deben ser planificadas y ejecutadas a escala local. La información sobre distribución geográfica de la biodiversidad, procesada mediante las tecnologías de la información, puede mejorar nuestra perspectiva sobre el problema, facilitando el desarrollo de modelos que nos ayuden a entender las relaciones entre los factores implicados y abriendo paso al diseño e implementación de soluciones.

En esta Tesis exploro una metodología de análisis y pronóstico, los Modelos de Distribución de Especies (MDE), que son herramientas diseñadas para calcular la distribución geográfica potencial de las especies a partir de registros de presencia y variables ambientales (topografía, clima, usos del suelo, etc.). Los MDE tienen como soporte conceptual las teorías sobre distribución de organismos y como soporte funcional las tecnologías de la información. En el trabajo que presento trato de utilizar y extender las posibilidades de los MDE para ofrecer respuestas a problemas de conservación en la flora de Andalucía.

Esta Tesis está dividida en tres Partes.

En la Parte I, que puede considerarse la Introducción General, trato de describir la *esencia* de los MDE. El Capítulo 1 contiene una introducción a las teorías sobre distribución de organismos en las que se apoyan los MDE, descritas desde la perspectiva de la ecología vegetal. El Capítulo 2) trata sobre el marco disciplinar al que pertenecen los MDE (la Ecoinformática), la naturaleza de los modelos, los MDE, el proceso de modelado, las incertidumbres, limitaciones y aplicaciones de la herramienta.

La Parte II y la Parte III forman el núcleo práctico, que describe los distintos trabajos de simulación realizados. Estas dos partes siguen un orden cronológico en cuanto al orden de realización de los distintos trabajos. Como consecuen-

*Los MDE tienen aplicaciones en estudios de biodiversidad, conservación, biogeografía, evolución, y expansión de enfermedades infecciosas u organismos invasores, entre otros.*

*Los invernaderos ya  
han hecho  
desaparecer  
formaciones  
vegetales singulares  
como los artales en  
El Ejido, y  
poblaciones de  
Linaria nigricans en  
Campohermoso,  
ambas localidades en  
Almería.*

cia existe en esta Tesis una línea ascendente en cuanto a complejidad en la aplicación de los MDE, que coincide con una adquisición progresiva de conocimientos sobre el funcionamiento y aplicaciones de la herramienta. En los trabajos de simulación descritos en esta partes trato de profundizar en dos problemas que ponen en riesgo la diversidad vegetal del sur de la península Ibérica: el Cambio de Uso del Suelo, y el Cambio del Clima.

En la Parte II, uso los invernaderos en el sureste árido ibérico como ejemplo paradigmático del conflicto entre la conservación de la diversidad biológica y el desarrollo económico, para: **a)** desarrollar un método para localizar potenciales conflictos espaciales entre la construcción de nuevos invernaderos y especies vegetales amenazadas; **b)** utilizar una combinación de técnicas para predecir la expansión de invernaderos en los próximos años.

En cuanto al efecto potencial del Cambio del Clima sobre la flora, tratado en la Parte III, presento los resultados de cuatro trabajos: **a)** la primera simulación de alta resolución del desplazamiento potencial del hábitat idóneo de formaciones vegetales en el Parque Nacional de Sierra Nevada, desarrollando una técnica para la detección temprana de frentes de avance y retroceso de las poblaciones, y un método para pronosticar las dinámicas potenciales de sucesión de las formaciones vegetales; **b)** desarrollo de una herramienta informática (un flujo de trabajo automatizado) para facilitar y mejorar las proyecciones de cambio potencial en la distribución de la flora ante distintos escenarios de cambio climático; **c)** aplicación del flujo de trabajo para simular cambios en la idoneidad del hábitat en poblaciones de flora amenazada de zonas áridas (Hoyas de Guadix y Baza) y alta montaña (Sierra Nevada), para comparar el efecto potencial del cambio del clima en la flora rara y amenazada de ambos hábitats; **d)** simulación de cambios en la distribución de 198 especies de la flora andaluza, teniendo en cuenta factores como la migración o la tolerancia a nuevas condiciones, para estudiar el comportamiento potencial de la diversidad vegetal de Andalucía lo largo del siglo XXI.

En los distintos trabajos trato de abordar dos perspectivas: una enfocada en el proceso de modelado, sus limitaciones

y sus posibilidades de mejora, y otra dedicada a la interpretación biológica de los resultados.

Esta no es una tesis experimental, sino una tesis de modelado y simulación, basada en los fundamentos biológicos de la ecología de plantas y en los métodos propios de la Ecoinformática. A consecuencia de los métodos utilizados, los resultados de los trabajos presentados están sujetos a la indeterminación inherente a eventos que aún no han tenido lugar.

*En el Capítulo 2 describo qué es la Ecoinformática, sus objetivos, y las herramientas que utiliza para alcanzarlos.*



## OBJETIVOS

---

Como objetivo general, este trabajo pretende ahondar en las perspectivas de conservación de la flora amenazada de Andalucía mediante el análisis, aplicación, y mejora de los Modelos de Distribución de Especies. Ambas vertientes, conservación y metodología de modelado, tienen sus propios objetivos específicos:

### **Objetivos metodológicos:**

1. Diseñar herramientas para pronosticar riesgos en poblaciones de plantas amenazadas relacionados con cambios en el uso del suelo.
2. Desarrollar un flujo de trabajo para simular de modo automatizado cambios potenciales en la distribución geográfica de la flora según distintos escenarios de cambio climático
3. Mejorar las proyecciones de cambio de distribución de especies ante el cambio del clima incorporando técnicas robustas como el ensamblado de los resultados de distintos algoritmos, y factores que generalmente no son tenidos en cuenta como la distancia de dispersión natural de las plantas, o su capacidad de tolerar condiciones climáticas relativamente adversas.

### **Objetivos de conservación**

1. Realizar un análisis de riesgo de desaparición de poblaciones de plantas amenazadas en entornos de agricultura intensiva bajo plástico.
2. Obtener las primeras simulaciones de efectos del Cambio Climático en la distribución de la flora de Andalucía. Estas simulaciones proporcionan datos valiosos sobre las posibilidades de persistencia de las poblaciones, dinámica del área de ocupación potencial de las especies y variaciones locales en riqueza específica.



Como objetivo transversal, pero no secundario, me propongo demostrar el valor de la Ecoinformática como disciplina emergente que puede aportar nuevos puntos de vista a la investigación biológica con entidad propia, y reivindicar su lugar en el entorno universitario, donde ha sido ignorada, o simplemente desconocida.

*La Ecoinformática ya  
está consolidada en  
USA, y en los  
últimos años se está  
extendiendo en  
Europa. En España,  
por desgracia, vamos  
a la cola...*

## TABLA DE CONTENIDOS

---

### I TEORÍA Y PRÁCTICA DE LOS MODELOS DE DISTRIBUCIÓN DE ESPECIES 1

1	DISTRIBUCIÓN DE PLANTAS	7
1.1	Introducción	7
1.2	Hipótesis del Continuo	7
1.3	Hipótesis de Brown	9
1.4	El Nicho Ecológico	9
1.4.1	Distintos conceptos de Nicho Ecológico	10
1.4.2	Dinámica del Nicho Ecológico	13
1.5	El equilibrio de las especies con el clima	15
1.6	Factores que afectan a la distribución geográfica de las plantas	16
1.6.1	Los gradientes ecológicos	17
1.6.2	El ambiente biótico	19
1.6.3	La capacidad de dispersión	19
1.6.4	Las perturbaciones	19
1.6.5	La capacidad evolutiva	20
1.6.6	La historia biogeográfica	20
2	MODELOS DE DISTRIBUCIÓN	23
2.1	Marco disciplinar	23
2.1.1	Ecoinformática	24
2.1.2	Informática de la Biodiversidad	26
2.2	El concepto de modelo	27
2.3	Definición de Modelo de Distribución	30
2.4	El proceso de modelado	31
2.4.1	Formulación teórica del modelo	32
2.4.2	Captura y preparación de la muestra de presencia	32

### II APLICACIONES DE LOS MDE AL ESTUDIO DE EFECTOS DEL CAMBIO DE USOS DEL SUELO EN LA FLORA AMENAZADA 35

3	LINARIA NIGRICANS VS. GREENHOUSES	39
3.1	Abstract	39

3.2	Introduction	40
3.3	Materials and methods	41
3.3.1	Study area, model species and green-houses	41
3.3.2	Sampling and environmental data	43
3.3.3	Species Distribution Models	44
3.3.4	Evaluation and selection of models	45
3.3.5	Postprocessing of distribution models	46
3.3.6	Extinction-risk model	46
3.4	Results	48
3.5	Discussion	50
3.5.1	Modelling and model selection	50
3.5.2	Extinction-risk model	52
4	MAXENT AND GEOMOD WORKING TOGETHER	55
4.1	Introduction	55
4.2	Test area, data sets and tools	57
4.3	Modelling and evaluation tools	60
4.4	Methodology and practical application	63
4.5	Results	68
4.6	Validation of the results and discussion	73
4.7	Conclusions and perspectives	78
III PRONÓSTICO DE EFECTOS POTENCIALES DEL CAMBIO DEL CLIMA EN LA FLORA 81		
5	CLIMATE CHANGE IN SIERRA NEVADA	85
5.1	Introduction	86
5.2	Materials and Methods	88
5.3	Results	93
5.4	Discussion	95
6	FLUJO DE TRABAJO AUTOMATIZADO	103
6.1	Introducción	103
6.2	Objetivo	104
6.3	Antecedentes	105
6.4	Material y Métodos	106
6.4.1	Herramientas informáticas	106
6.4.2	Proceso de modelado	114
6.5	El Flujo de Trabajo	122
6.5.1	PRP	123
6.5.2	MODPLAN	127

6.6	Resultados del flujo de trabajo	130	
6.6.1	Distribución actual y requerimientos ecológicos de la especie	130	
6.6.2	Modelo de idoneidad del hábitat en el periodo 1971-2000	131	
7	ALTA MONTAÑA Y ZONAS ÁRIDAS	141	
7.1	Introducción	141	
7.1.1	La cuestión biológica	141	
7.1.2	La cuestión metodológica	142	
7.2	Objetivos	143	
7.3	Material y métodos	143	
7.3.1	Áreas de trabajo y especies	143	
7.3.2	Análisis de cambio potencial del clima en SN y GB	145	
7.3.3	Calibrado de simulaciones	147	
7.3.4	Selección de modelos	148	
7.3.5	Análisis de efectos del Cambio del Clima en las especies y hábitats	149	
7.3.6	Análisis de MODPLAN	150	
7.4	Resultados	150	
7.4.1	Tendencias del clima hasta el 2050	150	
7.4.2	Selección de modelos	151	
7.4.3	Efectos del Cambio del Clima en las especies y hábitats	152	
7.4.4	Funcionamiento del ensamblado	153	
7.5	Discusión	155	
7.5.1	Efectos del Cambio del Clima en las especies y hábitats	155	
7.5.2	Funcionamiento del ensamblado	158	
7.5.3	Limitaciones técnicas del flujo de trabajo y posibles mejoras	160	
7.6	Conclusiones	162	
8	CAMBIO CLIMÁTICO Y FLORA DE ANDALUCÍA	163	
8.1	Introducción	163	
8.1.1	En el frente de avance	163	
8.1.2	En el frente de retroceso	165	
8.1.3	Efectos del cambio del clima en la biodiversidad	166	
8.2	Objetivos	167	
8.3	Material y Métodos	167	

8.3.1	Área de trabajo	167
8.3.2	Especies de trabajo	168
8.3.3	Modificaciones a MODPLAN	169
8.3.4	Calibrado y análisis de las simulaciones	176
8.3.5	Variaciones geográficas y temporales en la riqueza de especies	178
8.4	Resultados	179
8.4.1	Variaciones geográficas y temporales en la riqueza de especies	184
8.5	Discusión	188
8.5.1	Limitaciones del algoritmo de migración-tolerancia	188
8.5.2	Interpretación biológica de los resultados	191
8.5.3	El patrón regional de extinción	193
8.6	Conclusiones	194
IV REFLEXIONES Y CONCLUSIONES 195		
9	REFLEXIONES FINALES	197
9.1	La Ecoinformática	197
9.2	Los MDE	198
9.3	El Cambio del Usos del Suelo	199
9.4	El Cambio del Clima	200
9.5	Nuevas perspectivas en los MDE	202
9.6	Pero...	203
10	CONCLUSIONES	205
10.1	Metodología	205
10.2	Cambio de Usos del Suelo	206
10.3	Cambio del Clima	206
V APÉNDICE 209		
A	EVALUACIÓN DE MODELOS	211

Part I

TEORÍA Y PRÁCTICA DE LOS  
MODELOS DE DISTRIBUCIÓN DE  
ESPECIES



## INTRODUCCIÓN A LA PARTE I

---

En general, los MDE se perciben como bonitos mapas de colores. Y en parte esa apreciación tiene algo de cierto, porque aplicándoles un gradiente de colores el resultado puede llegar a ser ciertamente vistoso. Pero la impresión visual del modelo puede ocultar, o al menos, desviar la atención de lo importante: el valor de la herramienta, su potencia y sus aplicaciones.

Son varios los factores que pueden ser causantes de esta interpretación superficial de los resultados de los MDE: la propia historia del desarrollo de la herramienta, que es muy reciente; un cierto acantonamiento en el *lado teórico* de la ciencia, propio de cualquier disciplina en fase de desarrollo, que ha ralentizado su incorporación al mundo real; y su total inexistencia, ni siquiera por mención, en los planes de formación en biología.

Pero esta tendencia está cambiando, y rápidamente. Además de los grandes especialistas internacionales en MDE, son muchos los investigadores de nuestro país que han abierto brecha en este campo, contribuyendo al desarrollo teórico y práctico de la disciplina, y *contagando* a otros sus técnicas y conocimientos, facilitando su cada vez mayor expansión. A la expansión de los modelos de distribución como herramienta entre los científicos españoles han contribuido también la reunión *Predictive Modelling of Species Distribution: New Tools for the XXI Century* (Baeza, noviembre del 2005) a la que acudieron los máximos exponentes internacionales de la disciplina para compartir sus conocimientos con los que por aquellas fechas comenzábamos, y los **Talleres de Modelización** del nodo español de GBIF <sup>1</sup>, que son réplicas de los *Modelling Workshops* internacionales promovidos por GBIF <sup>2</sup>.

La mayor presencia de los MDE en la ciencia de nuestro país puede comprobarse leyendo las tesis doctorales que

*Sin pretender hacer una lista exhaustiva (y que me perdonen los que se sientan omitidos), puede nombrarse a: Araújo, Benito-Garzón, Brotons, Felicísimo, Hortal, Jiménez-Valverde, Lobo, Nogués-Bravo, Real, Seoane, ...*

---

<sup>1</sup> [www.gbif.es](http://www.gbif.es)

<sup>2</sup> [www.gbif.org](http://www.gbif.org)



se han desarrollado sobre el tema: La tesis Marta Benito Garzón (una pionera en este tipo de trabajos), titulada *El efecto del cambio climático sobre las distribuciones de los bosques ibéricos: pasado, presente y futuro*<sup>3</sup> (2006), en la que se exponen metodologías de proyección de modelos de especies arbóreas ante distintos escenarios del pasado y el futuro, con la intención de dilucidar la historia biogeográfica y las posibilidades de conservación de las especies de trabajo.

*La tesis de Marta Benito-Garzón ha sido una fuente de inspiración, y quiero agradecer desde aquí su contribución a mi trabajo.*

Manuel Rubén García, con su trabajo *Modelos Predictivos de Riqueza de Diversidad Vegetal. Comparación y Optimización de Métodos de Modelado Ecológico*<sup>4</sup> (2008), analiza puntos clave de la metodología como el tamaño de muestra, o las diferencias entre resultados de distintos algoritmos. También Tania Delgado, con su trabajo *Evolución de la Diversidad Vegetal en Ecuador ante un Escenario de Cambio Global* (2008) analiza las posibles consecuencias de un cambio drástico en el clima sobre la diversidad vegetal de Ecuador.

*Hay tesis de modelado anteriores, como la de Joaquín Hortal, del 2002, que aplica modelos predictivos a escarabajos coprófagos, pero no ha sido posible examinar su contenido.*

No sólo la ciencia se está beneficiando de la aplicación de los MDE, porque en el campo de la gestión ambiental también se están comenzando a utilizar en la conservación del medio natural y la resolución de conflictos entre conservación y desarrollo económico. Como resultado, son varias las empresas de gestión ambiental que envían técnicos a los distintos cursos de modelado de distribución con el objetivo de incorporar la metodología a sus trabajos.

Y un ejemplo más de crecimiento de la disciplina debería ser este trabajo, en el que se aplican los MDE con distintos objetivos de conservación, intentando vincular en lo posible el plano teórico de construcción del modelo con la realidad que se puede percibir en el campo.

*Modelar, su participio modelado y modelizar son los vocablos apropiados (según el RAE) para referirse al proceso de construcción de un modelo. Modelizado y modelización, muy utilizados, no existen en español.*

Teniendo en cuenta que los MDE no son una herramienta muy conocida, estimo oportuno incluir en esta tesis esta parte introductoria, que considero importante para facilitar la comprensión de los capítulos posteriores. Los capítulos descritos a continuación emanan de un trabajo de revisión exhaustiva de la literatura relativa a los MDE de los últimos 20 años.

<sup>3</sup> [http://www.uam.es/proyectosinv/Mclim/pdf/TesisMarta\\_2006.pdf](http://www.uam.es/proyectosinv/Mclim/pdf/TesisMarta_2006.pdf)

<sup>4</sup> [http://www.unex.es/unex/grupos/grupos/kraken/archivos/ficheros/Tesis\\_RGM.pdf](http://www.unex.es/unex/grupos/grupos/kraken/archivos/ficheros/Tesis_RGM.pdf)

En el Capítulo 1 expongo los conceptos teóricos sobre la distribución geográfica de las plantas que mayor relación directa tienen con los MDE, tratando resumidamente la Hipótesis del Continuo, el concepto de Nicho Ecológico o los factores que afectan a la distribución de las especies, entre otros.

En el Capítulo 2 trato el marco disciplinar de los MDE, introduciendo un concepto poco conocido en nuestro país: Ecoinformática. Otras cuestiones fundamentales, como el concepto de modelo o la definición de MDE acompañan a una sección extensa en la que describo el proceso de modelado. Finalmente menciono otras cuestiones no menos importantes, como las limitaciones, incertidumbres y aplicaciones de los MDE.

*Quien no conozca en profundidad los MDE, encontrará una guía de fundamentos con gran número de referencias útiles en este campo.*



# LA DISTRIBUCIÓN GEOGRÁFICA DE LAS PLANTAS

---

## 1.1 INTRODUCCIÓN

Conocer cómo las condiciones ambientales y los procesos poblacionales determinan la abundancia y distribución de las especies es una cuestión prioritaria de la biogeografía y la ecología (Brown, 1984). La **Ciencia de la Vegetación** es la disciplina que estudia los procesos que determinan los patrones de composición de la vegetación (Austin y Smith, 1989). No hay unanimidad en la comunidad científica respecto a estos procesos, como muestra la naturaleza de las distintas hipótesis que tratan de explicarlos, algunas de ellas enfrentadas, como la Hipótesis de Comunidades de Clements (1936), o la Hipótesis Individualista de Gleason (1926). A continuación resumiré brevemente las teorías sobre la distribución de plantas que mayor influencia tienen actualmente en los Modelos de Distribución de Especies.

## 1.2 HIPÓTESIS DEL CONTINUO

Según Austin y Smith (1989), el concepto de comunidad de especies coocurrentes solo es relevante para un paisaje particular y su patrón de combinaciones de variables ambientales. Ambos autores definen la comunidad como una propiedad del paisaje, ligada a las condiciones climáticas de un momento determinado. Antes de su trabajo de síntesis sobre la Hipótesis del Continuo, *A new model for the continuum concept*, al menos coexistían tres versiones de la hipótesis del continuo:

*La Hipótesis del Continuo es la base conceptual en una gran mayoría de trabajos de modelado de distribución.*

- **Continuo individualista de Gleason (1926):** la hipótesis de Gleason propone que la distribución de las especies es individual e independiente de la distribución de otras especies. La estructura de la vegetación

se debe a un fenómeno especie-escala que involucra la tolerancia de los individuos de diferentes especies a las condiciones ambientales locales, incluyendo la competencia interespecífica y otras interacciones.

- **Continuo de Gauch y Whittaker (1972)**: estos autores proponen que las especies más importantes están regularmente distribuidas a lo largo del gradiente con óptimos uniformemente espaciados, mientras que las especies menores están distribuidas independientemente.
- **Continuo estratificado (Austin, 1985)**: según Austin, cada estrato (por ejemplo, árboles, arbustos y hierbas) tiene un continuo de composición independiente, que puede tener una distribución regular o irregular de especies y con posibles relaciones entre estratos, que pueden ser independientes entre sí o estar correlacionados.

En 1993, **Collins *et al.* (1993)** revisaron el concepto de continuo para proponer una nueva visión jerárquica del mismo. Este nuevo concepto deriva de la hipótesis individualista de **Gleason (1926)**, la estructura jerárquica de las comunidades de **Kolasa (1989)**, y la hipótesis de las especies núcleo-satélite de **Hanski (1991)**.

Según esta hipótesis, algunas especies pueden aparecer a través de todo el gradiente ambiental, otras pueden ocupar grandes porciones del gradiente siendo poco frecuentes, y otras pueden estar muy localmente distribuidas. La amplia distribución de las primeras reduce los efectos de las fluctuaciones ambientales, haciéndolas menos dinámicas que aquellas con rangos ecológicos más restringidos, porque la amplitud de su nicho incorpora esas variaciones dentro de su rango. Las especies con rango ecológico restringido en cambio sí pueden cambiar su distribución y/o patrones de abundancia a lo largo de un gradiente como respuesta a las mismas variaciones del ambiente.

### 1.3 HIPÓTESIS DE BROWN

Según **Brown y Gibson (1983)**, existe un patrón de decrecimiento simétrico y monotónico en abundancia desde el centro del área de distribución hacia los bordes, y una correlación positiva entre la densidad de población local y la extensión de la distribución espacial entre especies similares. La teoría que ofrece para explicar este patrón se basa en las siguientes asunciones:

- La abundancia y distribución de cada especie está determinada por combinaciones de varias variables físicas y bióticas fundamentales para la supervivencia y reproducción de los individuos. Como consecuencia, las variaciones en la densidad de las poblaciones de las especies en el espacio reflejan la probabilidad de aparición de las combinaciones requeridas de variables ambientales.
- Algunos conjuntos de variables están distribuidos independientemente de otras, y hay un grado aleatorio significativo de variación local. La variación ambiental está autocorrelacionada, por lo que la probabilidad de que dos sitios tengan las mismas condiciones es una función inversa de la distancia entre ellos.
- La forma exacta de la variación espacial en la abundancia depende del número y tipo de variables ambientales que afectan a la especie y el patrón espacial de variación de esas variables.

*Estas asunciones prácticamente resumen base conceptual en la que se sustentan los MDE.*

### 1.4 EL NICHO ECOLÓGICO

Según **Soberón (2007)**, una de las principales debilidades teóricas identificadas en la disciplina de los MDE ha sido una difusa utilización del término **nicho ecológico** (**Kearney, 2006**).

#### 1.4.1 Distintos conceptos de Nicho Ecológico

Aunque investigadores anteriores como Darwin o Humboldt ya trabajaban con un concepto similar a la idea de nicho, la primera definición formal se debe a Grinnell (1917), cuyo interés se centraba en los factores que determinan dónde puede encontrarse una especie y qué nichos, generados por el ambiente, están ocupados. En su definición de nicho ecológico, lo interpreta como una descripción de los requerimientos del hábitat para una especie.

A estas variables Grinnell las denomina **escenopoéticas**, con un significado próximo a “describir la escena”.

Más concretamente lo define como *la mínima unidad de distribución espacial a la cual la especie está ligada por su estructura y limitaciones instintivas*. El concepto se basa en la importancia de las variables ambientales (clima y topografía) definidas a una escala geográfica relativamente grosera.

Teniendo en cuenta la definición de Grinnell, el conocimiento del nicho de la especie determinado por los requerimientos del hábitat es esencial para entender y predecir su distribución geográfica, por lo que esta definición ha tenido tradicionalmente una mayor relevancia en biogeografía.

Una década más tarde Elton (1927) revisa el concepto de nicho, pero enfatizando la importancia de los roles funcionales de las especies, afirmando que cada especie tiene un rol exclusivo en el ecosistema, y que este rol no puede ser compartido con otras especies dentro del mismo ecosistema, aunque el mismo nicho sí puede ser ocupado por otras especies en otros ecosistemas.

Elton llama **biónómicas** a las variables que describen gradientes bióticos, como la disponibilidad de recursos o la competencia interespecífica

Según Elton, las dimensiones del nicho estarían formadas por gradientes bióticos relacionados con la biología de la especie, la de sus competidores o depredadores, o gradientes de recursos directamente consumidos. Esta definición del nicho como función ha sido muy utilizada en áreas como la ecología de poblaciones.

En 1957, Hutchinson utiliza características de los nichos de Grinnell y Elton (ver Figura 1) para redefinir el concepto, afirmando que *el nicho ecológico es un hipervolumen en un espacio ecológico multidimensional, determinado por los requerimientos de la especie para reproducirse y sobrevivir*. Según esta definición, cada dimensión del nicho representa una variable ambiental importante para la supervivencia de la especie.

Las variables que dan forma al nicho ecológico de la especie pueden ser bióticas o abióticas, y pueden representarse mediante valores numéricos (como la temperatura), u otros tipos de valores (como la textura del suelo).

El propio **Hutchinson (1978)** apunta la existencia de dos niveles del nicho: el **nicho fundamental** y el **nicho realizado** (o **nicho efectivo**). Hutchinson llama nicho fundamental al espacio multidimensional determinado por todas las variables ambientales (bióticas y abióticas) donde la especie puede reproducirse y sobrevivir. El nicho realizado es un subconjunto del nicho fundamental, donde las especies pueden vivir dada la presencia de otras especies que compiten por los mismos recursos. Ante esta definición, el nicho realizado debería tener una extensión menor que el nicho fundamental y, además, sería exclusivo, de forma que dos especies no pueden solapar sus respectivos nichos realizados en un ambiente estable.

*En la disciplina MDE existe una discusión muy activa sobre la aplicación de los términos nicho fundamental o realizado a los resultados de los modelos.*



**Figura 1:** Conceptos de nicho ecológico de Grinnell, Elton y Hutchinson. Nf se refiere a nicho fundamental, y Nr a nicho realizado.

En el artículo *Grinnellian and Eltonian niches and geographic distribution of species* (**Soberón, 2007**) el autor ofrece una disección teórica de los conceptos de nicho de Grinnell y Elton, con el objetivo de clarificar la utilización de los términos en la disciplina MDE. Según Soberón, para alcanzar una correcta definición del nicho aplicable a los modelos de distribución, es necesario tener en cuenta tres principios:

1. El concepto de nicho tratado como hábitat debe separarse del concepto de nicho tratado como función.



2. Hay diferencias importantes entre los nichos definidos a escalas locales, que tienen en cuenta la ecología, comportamiento y fisiología de pequeñas poblaciones o individuos, y los nichos definidos a grandes escalas geográficas, que solo tienen en cuenta el concepto de distribución.
3. Hay que separar los distintos tipos de variables que componen el nicho en al menos dos grupos: **condiciones** y **recursos**.

*Las condiciones son equivalentes a las variables escenopoéticas de Grinnell, y los recursos son el subconjunto de las variables bionómicas de Elton que la especie consume directamente. La postura de estos autores trata de poner fin a un debate estéril, pero que es el favorito de muchos referees.*

Según Araújo *et al.* (2006) la diferenciación entre nicho fundamental y realizado es artificial, y por tanto, de poca utilidad en el contexto de los modelos de distribución. La alternativa que ofrecen estos autores consiste en desechar ambos conceptos, y aceptar que *cualquier caracterización del nicho es una descripción incompleta de los factores bióticos y abióticos que permiten a la especie satisfacer sus requerimientos ecológicos mínimos*

Esta propuesta coincide con la definición de nicho propuesta por Chase y Leibold (2003) según la que *el nicho ecológico está compuesto por las condiciones ambientales que permiten a las especies satisfacer sus mínimos requerimientos, para que la tasa de nacimientos de una población local sea igual o mayor que la tasa de mortalidad.*

Respecto a las dimensiones que puede llegar a tener el nicho ecológico de una especie hay cierta controversia. Por ejemplo, Brown (1984) indica que el número de dimensiones independientes del nicho ecológico de cualquier especie probablemente oscila entre 5 y 10, ya que cuanto mayor es el número de dimensiones independientes, menor es la posibilidad de que la especie esté presente en algún lugar geográfico, porque las localidades con combinaciones favorables de variables tienen escasa probabilidad de ocurrir por motivos puramente estadísticos. En cambio, Polechová y Storch (2009), en la misma línea que Hutchinson, afirman que es muy difícil describir con propiedad el nicho ecológico de una especie real, porque el número de dimensiones del nicho es potencialmente infinita, y además, los ejes más significativos del nicho pueden ser difíciles de identificar.

*Desde el punto de vista de los MDE, es posible conseguir una buena aproximación al nicho ecológico de una especie con un número limitado de variables.*

#### 1.4.2 *Dinámica del Nicho Ecológico*

La asunción de que el nicho de las especies permanece sin cambios (estasis del nicho ecológico), o cambia muy lentamente a lo largo de cientos de millones de años, actualmente influencia el estudio de la distribución de las especies y ha generado un debate considerable (Wiens y Graham, 2005). La estasis completa o un lento cambio del nicho ecológico se usa como justificación para simular cambios en la distribución de las especies a través del tiempo (Pearman *et al.*, 2008), pronosticar la extinción de especies y evaluar la pérdida de biodiversidad como respuesta al cambio climático (Thomas *et al.*, 2004; Martínez-Meyer *et al.*, 2004; Peterson *et al.*, 2002).

Aunque comprender la dinámica del nicho ecológico es un prerrequisito importante para interpretar estas simulaciones, apenas existen estudios que aclaren si la estasis del nicho es una norma o una excepción. A priori no es posible predecir qué especies pueden cambiar su nicho ecológico en cortos periodos de tiempo, ni están claras las condiciones ecológicas sobre los que esos cambios son posibles. Mientras no se desarrollen más trabajos en esta dirección, el grado en el que la dinámica del nicho puede afectar a los resultados de las simulaciones es desconocido, por lo que es importante analizar las conclusiones de los estudios basados en esta técnica aplicando un principio de precaución. Esta carencia origina una necesidad imperiosa de emprender estudios comparativos que identifiquen los grupos de especies que están caracterizados por la estasis del nicho o cambios rápidos en el mismo. En este campo, los propios modelos de distribución pueden contribuir a la detección de desplazamientos en el nicho que ocurren en las especies a través del tiempo o entre poblaciones separadas en el espacio geográfico (Pearman *et al.*, 2008).

El registro fósil nos muestra que el ambiente y las especies cambian con el tiempo, por lo que los nichos ecológicos no son estables, teniendo en cuenta periodos de tiempo geológico. Sin embargo, algunos estudios revelan una rápida evolución del nicho (periodos menores a 100 años) que en determinados grupos taxonómicos podría estar vinculada a un proceso activo de especiación (Peterson *et al.*, 1999).

*En los distintos capítulos de esta tesis asumo que el nicho no varía en el tiempo, porque los periodos son de unas pocas decenas de años.*

*Por el momento no existe ninguna medida estándar de desplazamiento de nicho.*

*No existen razones para pensar que el nicho de una especie puede desplazarse sobre todas sus dimensiones simultáneamente; probablemente este tipo de cambios solo afecten a una (o unas pocas).*

Un ejemplo de marco conceptual que asume cambios en el nicho ecológico es la Teoría de la Combinación de Habilidades Competitivas (Aarssen, 1983), según la cual las especies con nichos similares incrementan en el tiempo su habilidad competitiva, variando dimensiones de sus nichos para compensar el solapamiento y disminuir la tensión competitiva. Además, las especies pueden responder a los cambios ambientales adaptándose mediante cambios rápidos en sus requerimientos bióticos y abióticos, modificando sus propios nichos y afectando a su vez a los nichos de otros organismos. Cambios en los requerimientos de una sola variable, o un número pequeño de ellas, pueden permitirle colonizar nuevas áreas en la periferia de su distribución donde esos factores eran previamente limitantes (Brown, 1984). Por ejemplo, un reducido conjunto de adaptaciones a nuevos rangos de algunas variables del nicho pueden facilitar la explotación de ambientes muy específicos, como saladares, y las especies capaces de ese cambio pueden obtener una ventaja competitiva, o eludir la competencia con otras especies colonizando el nuevo hábitat (Brown y Gibson, 1983).

Un nicho dinámico es un requisito indispensable para posibilitar la especiación, por lo que debería ser normal encontrar especies con cierta capacidad de expandir o contraer determinadas dimensiones del nicho. Pero es cierto que distintas pruebas (Peterson *et al.*, 1999; Prinzing *et al.*, 2001; Davis *et al.*, 2005; Pearman *et al.*, 2008; Crisp *et al.*, 2009) apoyan que el dinamismo del nicho puede ser variable entre grupos taxonómicos, llegando casi a la estasis en clados evolutivamente longevos. La conservación de nicho es un proceso que puede estar causado por más de un factor a nivel de población. Por ejemplo, un intenso flujo genético puede ser una fuerza importante para prevenir la expansión de nicho. Suponiendo una especie dominante con un amplio rango geográfico, las pequeñas poblaciones en el borde, que serían candidatas a un desplazamiento del nicho para adaptarse a nuevas condiciones, en la práctica tienen muy pocas posibilidades de diferenciación, porque reciben una gran carga genética del núcleo del área de distribución (Kirkpatrick y Barton, 1997).

Por este motivo, las grandes poblaciones de especies muy extendidas parecen ser relativamente resistentes a cambios evolutivos rápidos y direccionales (Brown, 1984).

Una cuestión importante en biogeografía consiste en tratar de explicar por qué algunos clados han dispersado hacia algunos lugares pero no a otros, y la conservación de las dimensiones del nicho relacionadas con el clima parece ser un factor limitante importante en muchos grupos. Este factor podría explicar la Hipótesis del Conservacionismo Tropical (Wiens, 2004), según la cual, muchos grupos presentan una mayor diversidad en los trópicos porque se originaron allí y han tenido más tiempo para especiar en ese lugar, ya que las regiones tropicales fueron más extensas en el pasado (30-40 millones de años). Según esta hipótesis, la dispersión de las especies tropicales hacia regiones templadas es rara, porque estos organismos carecen de adaptaciones para superar inviernos fríos. Otro ejemplo de posible estasis del nicho lo suponen algunos grupos de plantas que están confinados en las regiones templadas, y aparecen solo en uno o dos continentes (Donoghue y Smith, 2004).

La conservación de nicho puede interpretarse como una limitación en la capacidad adaptativa de las especies porque impide la expansión sobre nuevos regímenes climáticos. Desde esta perspectiva la estasis del nicho ecológico convierte al cambio climático en un peligro para la biota de la Tierra, porque obligaría a las especies a trasladar sus rangos geográficos para seguir sus regímenes climáticos ancestrales a una gran velocidad (Wiens y Graham, 2005). En este escenario, las especies con escasa capacidad dispersiva estarán ante un riesgo de extinción más evidente de extinción, que no exime del mismo a otras con mecánicas de dispersión más efectivas.

## 1.5 EL EQUILIBRIO DE LAS ESPECIES CON EL CLIMA

Según las visiones de nicho de Grinnell y Hutchinson, las variables climáticas definen dimensiones importantes del nicho ecológico de las especies. Los modelos estáticos de vegetación asumen por defecto que la vegetación está en

equilibrio o casi-equilibrio con el clima. (Lenihan, 1993). Sin embargo, si este equilibrio es real o aparente es una cuestión muy discutida en la literatura ecológica de los últimos años (Gastón, 2003). Especialmente para las simulaciones de efectos potenciales del cambio climático, es clave distinguir los roles relativos del clima presente y pasado determinando la distribución de las especies, para mejorar la comprensión de las respuestas de los taxa ante cambios ambientales en el futuro (Araújo *et al.*, 2005a).

En este ámbito de discusión, la cuestión crítica es hasta que punto están en equilibrio con el clima las distribuciones actuales de las especies. Según Hutchinson (1957), *una especie está en equilibrio con el clima si aparece en todas las áreas climáticamente apropiadas y está ausente de todas las que no lo son*. Para estimar de modo preciso los límites de tolerancia de las especies al clima es necesario diseñar y llevar a cabo experimentos controlados, pero estos son caros y consumen mucho tiempo, lo que ha limitado el alcance de este tipo de estudios a muy pocas especies. Algunos autores, como Svenning y Skov (2004) han utilizado modelos de distribución basados en envueltas bioclimáticas, interpretando que la razón entre las distribuciones observadas y simuladas puede ser un indicador del estado de equilibrio de las distribuciones actuales con el clima.

*Una envuelta bioclimática delimita el espacio ecológico que puede habitar una especie según variables como temperatura y precipitaciones.*

Otros autores (Araújo *et al.*, 2005a) han realizado estudios con idéntico objetivo, encontrando que la relación entre la composición de especies de plantas y el clima es, en general, bastante alta. Este equilibrio distribución-clima indicaría que aplicar modelos de distribución actuales, o simulaciones de distribución futura, basados en variable climáticas, puede ofrecer resultados confiables en plantas.

## 1.6 FACTORES QUE AFECTAN A LA DISTRIBUCIÓN GEOGRÁFICA DE LAS PLANTAS

Se entiende como área de distribución de una especie el conjunto de lugares del espacio geográfico en los que puede detectarse su presencia (Soberón, 2007). En la literatura especializada se reconoce un conjunto de factores que in-

teraccionan dinámicamente con diferentes intensidades y a distintas escalas para producir la entidad fluida y compleja que es la distribución geográfica de una especie (Guisan y Zimmermann, 2000; Soberón y Peterson, 2005):

- Los gradientes ecológicos.
- El ambiente biótico.
- La capacidad de dispersión.
- Las perturbaciones.
- La capacidad evolutiva.
- La historia biogeográfica.

*El efecto antrópico sobre la distribución de las plantas quedaría incluido en el apartado de perturbaciones, que limitan el espacio geográfico disponible para las plantas.*

#### 1.6.1 *Los gradientes ecológicos*

La importancia del clima en la distribución de las especies ya fue reconocida por Humboldt (1807) y de~Candolle (1855). Desde entonces el clima, en combinación con otros factores ambientales, se ha utilizado para explicar patrones de distribución de plantas a lo largo de todo el mundo (Salisbury, 1926; Woodward, 1987). El análisis de gradientes climáticos y topográficos que afectan a las especies vegetales cobra mayor importancia a partir de los estudios de Whittaker (1965), que proporciona evidencias de que las especies presentan respuestas claras pero complejas respecto a distintos gradientes topográficos y climáticos. Basándose en los estudios de Whittaker, Austin (1980) propone una clasificación de gradientes que sigue vigente en el ámbito de los MDE (Guisan y Zimmermann, 2000):

- **Gradientes indirectos:** El término se refiere a los factores ambientales que no tienen influencia fisiológica directa en las plantas, pero están correlacionados con otras variables que si la tienen.

La elevación y la orientación son los ejemplos más aparentes: la elevación es un gradiente complejo porque muchas variables de importancia fisiológica como la temperatura, evaporación y disponibilidad de agua

están directamente correlacionadas con la elevación, pero esta correlación es local, y potencialmente distinta en diferentes áreas geográficas (Austin 1980). Los modelos calibrados con este tipo de gradientes no pueden extrapolarse en el espacio (Guisan y Zimmermann, 2000), porque las especies tienden a compensar diferencias regionales en las condiciones climáticas seleccionando micronichos similares mediante cambios en su posición topográfica (Walter, 1953).

*Esta clasificación de gradientes será muy importante para llegar a entender hasta que punto un MDE solo puede referirse a un subconjunto de dimensiones del nicho ecológico de una especie.*

- **Gradientes directos:** Son las variables ambientales que tienen un efecto fisiológico directo en la planta, sin tratarse de un recurso (materia o energía consumidas) por el que tenga que darse una explotación competitiva con otra especie (Austin, 1980). El *pH* es un ejemplo de este tipo de recurso, porque modifica la disponibilidad de varios nutrientes, como el fósforo y además tiene un efecto importante en la solubilidad del aluminio (Rorison, 1973). La temperatura es otro gradiente directo de gran importancia, y más sencillo de medir que el *pH*. Tiene gran influencia en el crecimiento, porque regula la velocidad de las reacciones enzimáticas, y lo detiene (o directamente mata individuos) por congelación en el límite inferior de tolerancia de la especie o desnaturalización de proteínas en el límite superior. Los modelos calibrados con gradientes directos pueden ser generales, y extrapolarse a otras áreas con los mismos rangos de valores de los gradientes.
- **Gradientes de recursos:** representan la disponibilidad de un recurso directamente consumido por la planta porque es esencial para su crecimiento y reproducción. La luz solar, el agua, el dióxido de carbono, y los minerales esenciales son los más importantes. Un gradiente de recursos muy importante, y que se puede medir tanto en campo como calcular sus valores potenciales en un SIG es la radiación solar. Es especialmente importante en los hábitats montañosos, ya que puede diferir de forma importante entre vertientes norte y sur (Austin, 1980).

*Los gradientes directos e indirectos representan lo que en ecología se denominan condiciones.*

### 1.6.2 *El ambiente biótico*

El ambiente biótico de una especie está constituido por las interacciones biológicas intra e interespecíficas, tanto las positivas (mutualismo, facilitación, etc.) como las negativas (competencia, parasitismo, depredación, etc) (Soberón, 2007). Para definir el ambiente biótico de una especie se requiere un conocimiento profundo de su biología, pero incluso estudios detallados pueden omitir aspectos fundamentales en este sentido.

*El gran olvidado en los MDE, por las dificultades que entraña su determinación, y transformación a un mapa digital en un SIG.*

### 1.6.3 *La capacidad de dispersión*

La mecánica y distancia de dispersión de los propágulos, los eventos estocásticos de dispersión a larga distancia, las barreras geográficas, y la disposición en el territorio del hábitat idóneo son los elementos a tener en cuenta para saber dónde puede aparecer una especie. En una localidad concreta pueden darse todas las condiciones apropiadas para el establecimiento de la especie, pero no podrá ser ocupada si ningún propágulo puede alcanzarla. Igualmente, y según la teoría de metapoblaciones, una dinámica fuente-sumidero

*Hasta el momento, la introducción de este factor en los MDE ha sido limitada. En el Capítulo ?? presento un ejemplo de introducción de esta variable para mejorar las predicciones de los MDE.*

en especies con gran capacidad de dispersión posibilitaría la aparición de individuos fuera de las condiciones idóneas (Pulliam, 2000). La dispersión es un proceso espacio-temporal dinámico, que debe interpretarse como una dimensión extrínseca del concepto de nicho (Araújo y Williams, 2000).

### 1.6.4 *Las perturbaciones*

Los eventos catastróficos de origen natural, como los incendios o las inundaciones periódicas tienen influencia en la distribución de determinadas especies (Guisan y Thuiller, 2005). Pero el factor más restrictivo para la distribución de las especies en ambientes antropizados es la perturbación humana expresada en la forma de un cambio de uso del suelo, que normalmente desemboca en la ocupación y des-

*Su carácter estocástico hace que las perturbaciones naturales hayan sido (por el momento) prácticamente ignoradas en los MDE.*



trucción de hábitats naturales, restringiendo la distribución geográfica de muchas especies.

#### 1.6.5 *La capacidad evolutiva*

*El dinamismo del nicho ecológico determina esta capacidad. Las grandes dificultades para medirla por falta de datos obligan a omitirla en los MDE.*

La capacidad evolutiva es la habilidad de las especies para adaptarse a nuevas condiciones. Este mecanismo depende directamente de la plasticidad genotípica que condiciona la dinámica del nicho ecológico de la especie (Soberón y Peterson, 2005). Solo tiene sentido al trabajar en un contexto espacio-temporal de la distribución de una especie. Es realmente difícil de medir, pero en el ámbito de los MDE se asume que su efecto es pequeño sobre cortos periodos de tiempo (Soberón y Peterson, 2005).

#### 1.6.6 *La historia biogeográfica*

La historia biogeográfica de una especie comienza con su diferenciación respecto a un antecesor. Y un factor muy importante que define la distribución actual de las especies es **dónde** y **cuándo** se diferenciaron. El dónde predetermina las barreras dispersivas, la superficie de hábitat apropiado disponible dentro de la distancia de dispersión, o la interacción positiva o negativa que se establecerá con especies preexistentes, entre otros. El cuándo determina el tiempo que ha tenido la especie para dispersarse, y alcanzar o no todas aquellas localidades con condiciones apropiadas para su establecimiento. Pero la historia biogeográfica de un taxon es muy difícil de reconstruir, aunque algunas posibilidades al respecto las proporcionan los yacimientos de paleopolen o el registro fósil. Este factor es difícil de incorporar como entrada a los MDE, pero sí que puede utilizarse para evaluar la certidumbre de proyecciones en el tiempo (hacia el pasado) cuándo existen suficientes datos.

Como puede apreciarse, existe una estrecha relación entre algunos de los factores enumerados y la concepción de nicho ecológico que se ha explorado en apartados anteriores, que muestra el vínculo que existe entre espacio ecológico requerido por la especie y espacio geográfico en el que vive. Pero lo cierto es que no se trata de una idea nueva, porque la relación entre los nichos ecológicos y las áreas de distribución de las especies ya la expresó el propio Grinnell a principios del siglo XX. A consecuencia de los múltiples requerimientos de su nicho ecológico, todas las especies tienen distribuciones más o menos restringidas (Brown, 1984). En principio, una especie podría encontrarse en todas aquellas localidades donde aparece la combinación adecuada de requerimientos bióticos y abióticos, pero la presencia de todas las dimensiones del nicho ecológico no implica la presencia de la especie en ese lugar, porque la dinámica espacial de las poblaciones está dirigida por la mecánica dispersiva de la especie y la disponibilidad de parches de hábitat apropiado. Por otra parte, una especie puede estar en un lugar en el que no se dan todos los requerimientos de su nicho, por un continuo aporte de individuos de localidades vecinas en las que sí se dan las condiciones idóneas.



## LOS MODELOS DE DISTRIBUCIÓN DE ESPECIES

---

### 2.1 MARCO DISCIPLINAR

Actualmente existen grandes problemas que actúan a escala global: la transmisión de enfermedades, las invasiones biológicas, los riesgos y desastres naturales, la extinción de especies, la destrucción de hábitats, los impactos potenciales del cambio climático, o la sobreexplotación de los recursos naturales. De muchos de ellos desconocemos causas, mecanismos y posibles consecuencias, y el alcance y dimensión de estos problemas es tal, que difícilmente será posible encontrar una solución sin un cambio de perspectiva y un salto cualitativo en el modo de enfrentarnos a ellos.

Este salto ya lo dio en su momento la biología molecular con la aparición de la **Bioinformática** (Hogeweg y Hesper, 1978). La visión de futuro de los investigadores en genética y proteómica les permitió asimilar los cambios tecnológicos (larga capacidad de almacenamiento digital, potencia de procesamiento, internet, *world wide web*, tecnología de bases de datos distribuidas) según ocurrían para explotarlos en la búsqueda de respuestas sobre cuestiones de gran calado, como el código de bases del genoma humano.

Un segundo escalón hacia este nuevo modo de resolver problemas lo ha supuesto la **e-Ciencia** (también adoptada por los investigadores en física de partículas), nacida para explotar los beneficios de la alta conectividad y la potencia de las nuevas redes distribuidas de recursos computacionales (computación GRID), que posibilitan el trabajo con enormes cantidades de datos, y mejoran las posibilidades de trabajo colaborativo gracias a los nuevos y rápidos canales de comunicación, como la videoconferencia.

Los avances tecnológicos en cuanto a plataformas de teledetección (satélites Landsat, Ikonos, QuickBird, sensores multispectrales aerotransportados, etc.) y los Sistemas de

*El término e-Ciencia fue acuñado por John Taylor (director de la Oficina de Ciencia y Tecnología del Reino Unido) en 1999.*

*La computación GRID es la utilización simultánea de varios ordenadores conectados en red para resolver un único problema.*

Información Geográfica están cambiando el modo en el que la información del medio ambiente se captura, almacena y distribuye (Krishtalka y Humphrey, 2000). Como resultado, hoy día la comunidad científica dispone de un enorme volumen de datos relativos a distintos aspectos del medio ambiente (clima, topografía, cubierta vegetal, composición geológica del terreno, etc), que son imprescindibles para entender el funcionamiento de nuestro mundo, y están preparando a las Ciencias de la Tierra para alcanzar un nuevo nivel de desarrollo.

### 2.1.1 *Ecoinformática*

En este contexto está madurando una disciplina que ya tiene cierta relevancia, la **Informática del Medio Ambiente** (*Environmental Informatics*), encargada de asimilar conocimiento tecnológico para adaptarlo y aplicarlo a las necesidades de la investigación en las Ciencias de la Tierra. Dentro de la Informática del Medio Ambiente existen distintas áreas de especialización, más o menos desarrolladas, relativas a distintos campos de conocimiento, y en ocasiones muy relacionadas entre sí. Probablemente la más aparente (la que nos ocupa) sea la **Ecoinformática** (*Ecoinformatics*, también llamada por algunos autores *Ecosystem Informatics*). Esta nueva disciplina, que fusiona ecología y tecnologías de la información, es un área de trabajo relativamente reciente (apenas 20 años), y aunque aún no tiene una entidad propia en las universidades europeas, sí está madurando rápidamente en Estados Unidos y América del Sur.

*Informática del Medio Ambiente y Ecoinformática pueden resultar términos sinónimos, aunque desde mi punto de vista, el primero engloba al segundo, junto a otras disciplinas como la Informática del Clima.*

La Ecoinformática, en su aspecto más ambicioso, puede describirse como un marco de trabajo interdisciplinar que integra las ciencias del medio ambiente y las tecnologías de la información para definir entidades y procesos naturales, a todos los niveles de complejidad de los ecosistemas, desde los genes hasta las redes ecológicas, utilizando un lenguaje común a seres humanos y computadoras<sup>1</sup>. Intenta facilitar la investigación del medio natural, desarrollando bases de datos de información ambiental y métodos de ac-

<sup>1</sup> <http://www.waite.adelaide.edu.au/ISEI/>

ceso, así como nuevos algoritmos que permitan combinar distintos conjuntos de datos ambientales para comprobar hipótesis ecológicas y realizar simulaciones y pronósticos (Hale y Hollister, 2009). El objetivo es doble: incrementar el conocimiento científico del medio y apoyar la toma de decisiones en relación a cuestiones tan importantes como la sostenibilidad ecológica, biodiversidad y cambio global.

La Ecoinformática se ocupa de:

- Integración de datos a través de las categorías y niveles de complejidad de los ecosistemas.
- Inferencia de procesos ecológicos a partir de patrones en los datos.
- Simulación adaptativa y predicción de comportamiento de los ecosistemas.
- Técnicas de computación inspiradas en la biología como la lógica difusa, los autómatas celulares, las redes neuronales artificiales y algoritmos evolutivos.

El desarrollo de la Ecoinformática depende de la evolución de las Tecnologías de la Información. Algunos de los avances en este campo que se consideran cruciales para la Ecoinformática son:

- Computación de alto rendimiento para un rápido acceso a datos y procesamiento, y gran capacidad de almacenamiento.
- Representación de datos orientada a objetos para facilitar la estandarización e integración de datos embebendo los metadatos y las operaciones en las estructuras de datos.
- Internet, para facilitar la compartición de datos y software.
- Sistemas de Información Geográfica y Teledetección para facilitar la adquisición, almacenamiento y explotación de información geográfica.
- Animación y otras herramientas para facilitar la visualización y simulación de datos.

*El desarrollo de estándares de metadatos que faciliten el procesamiento automático de la información en red (Web Semántica) también es una tecnología crucial para la Ecoinformática, pero apenas se está desarrollando el concepto.*

### 2.1.2 *Informática de la Biodiversidad*

Una de las áreas de la Ecoinformática más desarrolladas es la Informática de la Biodiversidad (*Biodiversity Informatics*). **Soberón y Peterson (2004)**, definen la Informática de la Biodiversidad como *la aplicación de las tecnologías de información a la gestión, exploración algorítmica, análisis e interpretación de grandes cantidades de datos primarios relativos a la vida, particularmente al nivel de especie*. Este nivel de trabajo, la especie, hace que para la Informática de la Biodiversidad sea fundamental una base sistemática robusta, quedando condicionando su desarrollo al de la propia Taxonomía.

Las herramientas fundamentales de esta disciplina ecoinformática son los **Modelos de Distribución de Especies**, que se están aplicando con éxito en distintos campos, como la exploración de la biodiversidad (**Soberón y Peterson, 2004**); el diseño de áreas protegidas (**Chen y Peterson, 2002**); la evaluación de daños potenciales en cultivos por plagas (**Sánchez-Cordero y Martínez-Meyer, 2000**) y la evaluación de posibles invasiones biológicas **Evangelista et al. (2008)** o aparición de enfermedades (**Peterson et al., 2003**), entre otras.

*En una sección posterior se tratarán en profundidad los fundamentos teóricos y prácticos de los MDE, ya que en este trabajo suponen la herramienta fundamental.*

Sin embargo el panorama futuro de la disciplina no es del todo esperanzador. Aún hoy, en 2009, algunos de los retos y problemas de la Informática de la Biodiversidad apuntados por **Soberón y Peterson (2004)**, siguen vigentes, y pueden comprometer la trayectoria ascendente de la disciplina. El más importante se refiere a la calidad de los datos. En un contexto de miles de gigabytes de información de biodiversidad, el control de calidad es crucial. Pero es muy frecuente que en las colecciones aparezcan especímenes mal identificados, o una taxonomía obsoleta. Las bases de datos sin mantenimiento de la nomenclatura pueden quedar inservibles en pocos años (**Gastón y Mound, 1993**). Una cuestión paralela es la ausencia o mala georreferenciación de los registros de presencia de las especies, especialmente los más antiguos. Además, los datos georreferenciados de presencia de las especies suelen tener sesgos significativos en el tiempo y en el espacio, debido a diferencias en el esfuerzo de muestreo. La solución de estos problemas es laboriosa, porque

resulta muy difícil implementar métodos automáticos de depuración, y siempre es necesaria una cierta intervención de especialistas en los taxones implicados, que son un recurso muy escaso. Otro problema, que en realidad debe percibirse como una fase de desarrollo inevitable y necesaria, es el referido a las herramientas de software. Éstas se encuentran en una fase temprana de integración y operatividad, pero es cierto que están recibiendo cada vez mayor apoyo institucional, y se están beneficiando de la aplicación de la filosofía del software libre (comunidad *OpenSource*), gracias a una comunicación ágil entre los desarrolladores de herramientas y los investigadores que las utilizan.

Por último, la ausencia de formación en este campo en los programas universitarios, y la pérdida de vocaciones en taxonomía. La combinación de biología, informática y geografía no suele darse en los entornos universitarios, y la taxonomía está perdiendo financiación por su escaso interés económico y social.

*Programas de código abierto como OpenModeller o GRASS dan alas al avance de la Ecoinformática.*

## 2.2 EL CONCEPTO DE MODELO

El mundo real es complejo y multidimensional. Tiene diferentes escalas espaciales, que abarcan desde el nanómetro hasta el año-luz cúbico. Tiene una escala temporal, que comprende el intervalo entre el pasado conocido y el futuro por conocer. Tiene componentes discretos, las partículas fundamentales, que se combinan entre sí, formando una amplia gama de entidades superiores, desde átomos a galaxias. Estos elementos interactúan entre sí, mediante procesos físicos o químicos, dando lugar a propiedades emergentes, como la vida, que añade nuevas dimensiones a nuestro mundo: desde la biogeoquímica global a las interacciones entre individuos. Y el ser humano, una nueva fuente de complejidad, capaz de modificar su entorno a voluntad, para adaptarlo a sus necesidades.

En este mundo complejo se plantean grandes cuestiones que son críticas para su supervivencia del hombre: el desarrollo sostenible, la conservación de las especies, los desastres naturales, las invasiones biológicas, desertización



y desertificación, y una larga lista de problemas similares, de mayor o menor calado. Para resolverlos, es necesario conocer todos los factores implicados y descomponerlos en sus unidades fundamentales.

Pero nuestros insuficientes recursos [tanto en inteligencia como en desarrollo tecnológico] nos privan de la posibilidad de analizar y entender el mundo tal y como es, por lo que el hecho de abordar un problema supone, en sí mismo, un problema.

*Las definiciones sobre el término **modelo** difieren entre autores, y no es el objetivo de este texto recoger la diversidad de opiniones al respecto.*

Un **modelo** es una herramienta que ayuda a afrontar la complejidad; es una construcción conceptual que describe de forma simplificada un sistema físico que existe en el mundo real, y ayuda a entender su funcionamiento porque ofrece una versión manejable del objeto, adaptada a nuestros recursos intelectuales y a nuestras herramientas de procesamiento y análisis de datos. Recurriendo a una definición más formal, un modelo es *una representación simplificada de la realidad, en la que aparecen algunas de sus propiedades* (Joly, 1988).

Para construir un modelo es necesario definir un conjunto de suposiciones previas y unas condiciones iniciales. Posteriormente se seleccionan una escala temporal y una escala espacial apropiadas, se desglosa el problema en factores conocidos y manejables, se omiten aquellos poco conocidos (atención, peligro) o los que se suponen con efectos despreciables, y se establecen las relaciones entre estos elementos.

*Para cada cuestión compleja, siempre existe una solución simple e incorrecta (Albert Einstein).*

Esta construcción inicial es un modelo conceptual. Si el problema en cuestión requiere para su solución de resultados numéricos, el modelo puede ganar complejidad añadiéndole los valores de las variables de entrada y las ecuaciones que las relacionarán con los resultados. Entonces, el modelo habrá crecido hasta convertirse en un modelo matemático o estadístico. En este grupo se incluyen aquellos modelos que utilizan como base información geográfica, ya sea cualitativa o cuantitativa.

Dentro del conjunto de posibles tipos de modelos, Leohle (1983) afirma que hay dos bien diferenciados: las **herramientas de cálculo** que son modelos empíricos, fruto de la observación, que solo informan de la configuración del mundo, y los **modelos teóricos**, que son mecanicistas, porque

predicen respuestas a partir de relaciones causales plausibles o comprobadas.

Independientemente de su naturaleza o tipo, cuanto más complejo es un modelo, más se acercan sus resultados a la realidad (menor error o incertidumbre en sus resultados), y más fiables son sus pronósticos, pero es menos manejable, por lo que es importante alcanzar un equilibrio entre complejidad, incertidumbre y operatividad.

Pero los modelos, por defecto, nunca son perfectos; están generados por mentes humanas que cometen errores u omisiones. Además, por su propia naturaleza, solo trabajan con un subconjunto de las propiedades reales del sistema. Como consecuencia, al trabajar con modelos, se debe asumir automáticamente una desviación de sus resultados con respecto a la realidad (incertidumbre).

El uso de modelos implica asumir que tienen limitaciones, por lo que sus resultados deben ser analizados según los objetivos inicialmente previstos y según la información utilizada para su puesta en marcha **Rogers y Johnson (1998)**.

Según **Varcoe (1978)**, los modelos, además de un interesante conjunto de ventajas, tienen ciertas limitaciones.

Entre las ventajas más aparentes:

- Permiten la integración de variables complejas en un marco normalizado que permite la comparación con otros modelos.
- La creación de modelos requiere el trabajo de grupos interdisciplinarios, favoreciendo el intercambio de información, la sinergia entre ideas y el trabajo en equipo.
- La simulación es un proceso relativamente rápido y barato (según su magnitud), teniendo en cuenta la potencia y precio actual de los recursos computacionales.
- Los modelos permiten simular el paso de grandes periodos de tiempo en cuestión de minutos.
- Los modelos pueden actualizarse para incorporar nuevo conocimiento o capacidades tecnológicas.
- La creación de un modelo siempre mejora la comprensión del sistema que se trata de modelar.

*Tomando esta clasificación como referencia, los MDE son herramientas de cálculo, aunque pueden incluir principios mecanicistas, como podrá verse más adelante.*

*Construye tu teoría tan simple como sea posible, pero no más simple (Albert Einstein).*

- Un modelo ya establecido puede proyectarse sobre un escenario diferente, permitiendo explorar situaciones basadas en la pregunta ¿que pasaría sí...?.

Las limitaciones generales más importantes son:

- Los resultados de un modelo deben analizarse críticamente, teniendo en cuenta los supuestos y coherencia del modelo.
- Los modelos requieren una evaluación, pero no existe una metodología estándar para llevarla a cabo.
- Muchos modelos requieren gran cantidad de datos muy difíciles de conseguir.
- Es relativamente sencillo hacer un modelo sobre cualquier aspecto del medio obviando información crítica, pero difícilmente se ajustará a la realidad con garantía de éxito.
- Un modelo tiene límites de aplicabilidad que no deben excederse.
- Los modelos generalmente se ajustan correctamente a la visión de la realidad que su autor tiene, que puede ser diferente a la de otro autor.

*Todos los modelos son incorrectos, pero algunos son útiles (George EP Box).*

### 2.3 DEFINICIÓN DE MODELO DE DISTRIBUCIÓN

En la literatura especializada no abundan las definiciones de MDE, y su significado suele darse por sobreentendido. A continuación propongo una definición:

Un MDE es una construcción numérica que define en el espacio ecológico las relaciones que existen entre la presencia de una especie y los valores de variables ambientales con influencia en su distribución. El resultado se expresa en el espacio geográfico como un mapa digital que representa la idoneidad del hábitat o la probabilidad de presencia de la especie (según el método utilizado para construir la relación).

*El espacio ecológico es un hipercubo cuyos ejes representan los valores de distintas variables ambientales a lo largo de todo el territorio bajo estudio.*

Según esta definición, hay cinco elementos a destacar en cualquier MDE: registros de presencia de la especie, variables ambientales almacenadas como mapas digitales, un algoritmo que analiza la relación entre ambas entradas, un modelo definido en el espacio ecológico, y la representación geográfica del resultado en forma de mapa.

Esta aparentemente simple estructura esconde una gran complejidad en cada paso de desarrollo del modelo. Los puntos de presencia pueden proceder de distintas fuentes (citas bibliográficas, trabajo de campo) con distintos niveles de precisión en su georreferenciación. Las variables ambientales pueden representarse a distintas resoluciones, y tener o no influencia real en la distribución de la especie. Además existen muchos tipos de algoritmos para generar MDE, con distintas lógicas internas que proporcionan resultados diferentes.

Cada parte del proceso necesita herramientas (programas informáticos) diferentes, que no siempre están integradas, dificultando el procedimiento, y obligando al investigador a tener una cierta destreza con la tecnología informática.

Esto solo es una vista previa de la complejidad inherente al proceso de modelado, y lo que a continuación se ofrece es una disección exhaustiva del proceso de modelado.

*Generalmente los MDE proporcionan resultados en una escala de idoneidad del hábitat de 0 a 100, o de probabilidad de presencia de 0 a 1.*

## 2.4 EL PROCESO DE MODELADO

La generación de MDE es un proceso computacionalmente complejo, que necesita gran variedad de datos, técnicas, paquetes de software y hardware apropiado. Para garantizar un buen resultado, es importante seguir un proceso como el que se expone a continuación.

1. Formulación teórica del modelo
2. Captura y preparación de la muestra de presencia
3. Preparación y selección de las variables ambientales
4. Análisis de viabilidad de los datos
5. Selección y parametrización de algoritmos predictivos

6. Calibrado
7. Proyección al espacio geográfico
8. Evaluación
9. Post-procesamiento y análisis
10. Aplicación

#### 2.4.1 *Formulación teórica del modelo*

La predicción de la distribución geográfica de una planta o una comunidad vegetal comienza con el desarrollo de un modelo conceptual basado en la teoría del nicho ecológico, el análisis de gradientes de la vegetación, el postulado del equilibrio de las especies con el clima y en la premisa de que la distribución de la vegetación puede predecirse a partir de la distribución espacial de variables ambientales que están correlacionadas o controlan directamente la distribución de plantas (Franklin, 1995). También es necesario en esta fase seleccionar los factores que afectan a la distribución de la especie que podrían ser incluidos en el modelo, como distancias de dispersión o interacciones con otras especies. Todas estas decisiones deben estar condicionadas a los objetivos del trabajo, porque no será lo mismo preparar un MDE para buscar nuevas localidades de presencia de una especie, que uno para apoyar el diseño de una reserva natural. La formulación teórica exige un conocimiento profundo de las técnicas de modelado, la biología del organismo, y una definición clara de los objetivos.

*Los MDE se aplican indistintamente a especies o comunidades. En esta tesis aplico con mayor frecuencia la primera opción, pero en el Capítulo 5 trato un caso basado en comunidades.*

#### 2.4.2 *Captura y preparación de la muestra de presencia*

Los registros de presencia son imprescindibles para generar modelos de distribución. Consisten en pares de coordenadas que ubican en el espacio geográfico localidades en las que se ha detectado la presencia de la especie. La disponibilidad de registros de presencia de calidad es la primera limitación ante la que se encuentra un investigador a la hora de abordar un trabajo de modelado. Tanto es así que en muchos trabajos

de modelado los registros de presencia son recolectados *ex profeso* por el propio equipo de investigación, cuando no ha sido posible encontrar datos de calidad apropiada en servicios públicos como GBIF<sup>2</sup>, o en otras bases de datos similares.

---

<sup>2</sup> <http://www.gbif.org/>



## Parte II

# APLICACIONES DE LOS MDE AL ESTUDIO DE EFECTOS DEL CAMBIO DE USOS DEL SUELO EN LA FLORA AMENAZADA





## INTRODUCCIÓN A LA PARTE II

---

El Cambio Global comprende dos grandes cuestiones, el Cambio de Usos del Suelo, y el Cambio del Clima. En esta sección la tesis se enfoca en la primera de ellas con un caso de estudio tan particular como emblemático, la expansión de los invernaderos en el sureste árido de la Península Ibérica. Y el motivo de tratar tal caso tiene una explicación sencilla. Durante varios años nuestro grupo de investigación ha trabajado con algunos terófitos de los pastizales mediterráneos, especialmente con *Linaria nigricans*. Y año tras año hemos observado como la construcción de nuevos invernaderos restaba progresivamente hábitat a la especie, limitando su tamaño poblacional. Por supuesto no es la única especie afectada, pues todas sus compañeras de hábitat se encuentran ante el mismo peligro.

Los dos trabajos que expongo a continuación tratan de ofrecer avances en el desarrollo de herramientas para revelar poblaciones de plantas amenazadas por la construcción potencial de nuevos invernaderos, y de mejorar la predicción de la expansión de invernaderos.

El texto del Capítulo 3 corresponde fielmente al artículo *Assessing extinction-risk of endangered plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses*, publicado en la revista *Biodiversity and Conservation* (Benito *et al.*, 2009). En él se desarrolla una metodología para localizar poblaciones (conocidas y potenciales) de *Linaria nigricans* amenazadas por la construcción de nuevos invernaderos. La técnica es novedosa porque modela la distribución potencial de invernaderos tratándolos como si fueran una especie invasora, y la cruza con la distribución potencial de la planta para localizar las áreas de peligro.

El Capítulo 4, titulado *Greenhouses, land use change, and predictive models: MaxEnt and Geomod working together* se publicó como un capítulo del libro *Modelling Environmental Dynamics: Advances in Geomatic Solutions* (Benito y Peñas,

*Aunque el artículo se ha publicado en 2009, su redacción se remonta a mediados de 2007. Algunos problemas editoriales lo han mantenido enterrado más de un año.*

2008). En este capítulo, se intenta ir un poco más lejos en la predicción de la invasión potencial hábitats naturales por invernaderos, combinando un simulador de cambio de usos del suelo con un MDE para simular la expansión de la superficie con invernaderos ante distintos escenarios.

Ambos trabajos tienen aplicaciones prácticas en aquellas áreas mediterráneas en las que existen conflictos entre la conservación del medio natural y el desarrollo económico basado en los cultivos bajo plástico.

## ASSESSING EXTINCTION RISK OF ENDANGERED PLANTS USING SPECIES DISTRIBUTION MODELS: A CASE STUDY OF HABITAT DEPLETION CAUSED BY THE SPREAD OF GREENHOUSES

---

### 3.1 ABSTRACT

The species distribution models (SDMs) are a useful tools for investigating rare and endangered species as well as the environmental variables affecting them. In this paper, we propose the application of SDMs to assess the extinction risk of plant species in relation to the spread of greenhouses in a Mediterranean landscape, where habitat depletion is one of the main causes of biodiversity loss. For this aim, we chose *Linaria nigricans*, a narrow endemic and threatened species, which inhabits areas suitable for greenhouses expansion, as a model species. For this purpose, presence records of greenhouses and the model species, a dataset of environmental variables, and different only-presence-based modelling algorithms (Bioclim, Domain, GARP, MaxEnt and ENFA) were used to build SDMs for *Linaria nigricans* as well as for greenhouses. To evaluate the models a modified approach of the area-under-curve ROC (AUC) was applied.

The results were different models for greenhouse suitability as well as for the habitat of *Linaria nigricans*. Combining the most accurate models, we generated an extinction-risk model of *Linaria nigricans* populations, which enabled us to assess the sustainability of the most threatened populations. Our results show that is possible to model greenhouses spreading as a *biological invasion*. The procedure explained and used in this work is quite novel, and offers an objective spatial criterion intended for the management of natural resources and for the conservation of the biodiversity in areas threatened by habitat-depletion processes as particular as greenhouse expansion.

*En esta trabajo los invernaderos se modelan como una especie biológica por sus requerimientos climáticos y topográficos.*

**Keywords:** Greenhouses, habitat fragmentation and depletion, *Linaria nigricans*, Mediterranean landscape, south-eastern Spain, threatened species.

### 3.2 INTRODUCTION

Land use change is among the main cause of threat for the biodiversity due to the serious consequences that it is provoking in natural habitats (habitat depletion, fragmentation, and destruction) (Soule, 1991; Sala *et al.*, 2000). In a context of global change, one priority is to identify areas with high extinction risk (Williams y Araújo, 2000), in order to model the potential responses of the species to incoming changes (Elston y Buckland, 1993), and to apply that knowledge in conservation policies.

The development of the Geographical Information Systems (GIS) and its use in research is improving the design of powerful spatial-analysis methods directly applicable to the conservation of biodiversity and land management. These methods include the species distribution models (SDMs), which can predict the potential distribution of the habitat suitable for a targeted species by analysing relationships between presence records and predictor variables stored as GIS coverages (Guisan y Zimmermann, 2000). Recent methodological progress, in terms of predictive algorithms and environmental predictors at a better spatial resolution, has made predictive models more reliable and applicable to model the distribution of rare and endangered species (Guisan y Thuiller, 2005; Preston *et al.*, 2006). SDMs are successfully applied to locate new populations of rare and threatened species (Guisan *et al.*, 2006), to select areas that guarantee the persistence of endangered species (Williams y Araújo, 2000), to predict spatial patterns of biological invasions (Thuiller *et al.*, 2005a; Ward, 2007), to identify locations with high extinction risk (Araújo *et al.*, 2002) and to evaluate potential effects of global change in species distribution (Thuiller *et al.*, 2005b). In fact, SDMs can be used in plant-conservation planning to introduce objective and quantitative criteria for decision making.

In 1999 an estimated 682000 ha were occupied by greenhouses throughout the world, especially in China (380000 ha), followed by Mediterranean countries (161300 ha in France Italy, Spain, Greece, Turkey, Morocco, and Algeria (Takakura y Fang, 2002). The spreading of greenhouses usually implies habitat depletion and an overexploitation of scarce natural resources such as water or soil. In fact, the spread of greenhouses is the main threat against the outstanding biodiversity in south-eastern Iberian peninsula (Martínez-Fernández y Esteve, 2004), where it frequently signifies the occupation of natural habitats of European interest (Anónimo, 1992), affecting communities of *Periploca*, *Ziziphus*, and *Maytenus* (Mota *et al.*, 1996), which harbour protected species (Anónimo, 2003), and many other plant species, like *Androcymbium gramineum*, *Narcissus tortifolius*, *Euzomodendron bourgaeum*, *Cynomorium coccineum* and *Linaria nigricans*.

The aim of this work is to propose a methodology based on SDMs capable of predicting, in spatial terms, the extinction risk associated to natural populations of species, in relation to the spread of greenhouses. The case study is the spread of greenhouses in south-eastern Iberian peninsula, and the model species chosen is *Linaria nigricans*, a rare endemic and threatened species of this area. For this aim, we have applied different species distribution models to select the most accurate ones; also, we have combined the best models to build the extinction-risk model to evaluate the sustainability of the populations of *Linaria nigricans* in relation to greenhouses spreading. This is intended as a framework that could lead to an objective spatial criterion for improving conservation and management plans.

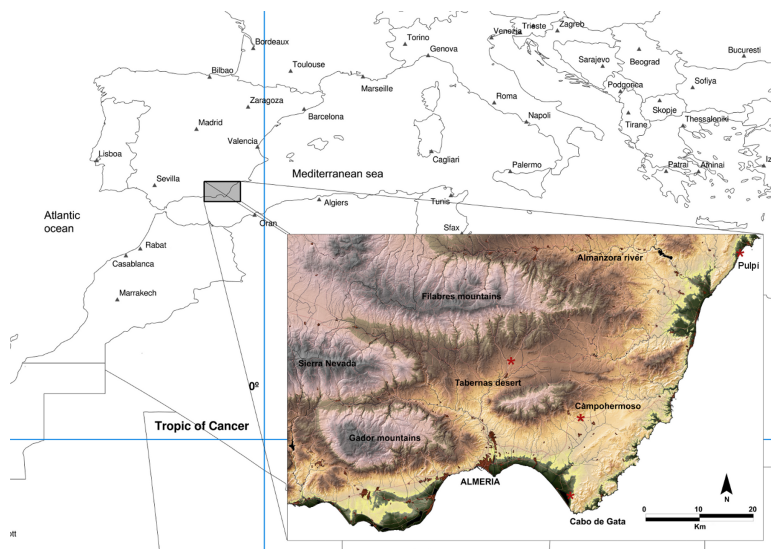
*La detección de zonas con alto riesgo de extinción de especies es uno de los objetivos principales de este trabajo.*

### 3.3 MATERIALS AND METHODS

#### 3.3.1 Study area, model species and greenhouses

The study area is located in south-eastern Iberian peninsula (Figura 1) (2.9° - 1.6° west and 36.6° - 37.5° north). The climate is Mediterranean, with rainfall of 200-300 mm (fre-

quently torrential), high evapotranspiration, and mean annual temperatures of 16-17°C, with warm, dry summers and temperate winters (Mota *et al.*, 2004). Geologically, post-orogenic sedimentary materials predominate, with calcareous or siliceous mountain and coastal dunes as well as Quaternary volcanism (Fernández, 1996). The landscape is dominated by a mosaic of shrubs and chamaephyte plant communities (with *Ziziphus lotus*, *Thymus* spp., *Helianthemum* spp., *Rosmarinus* spp., etc.), xerophytic grasslands with large patches of *Stipa tenacissima* (called *espartales*), and varied communities of ephemeral annual plants (Mota *et al.*, 2003).



**Figure 1:** Situation of the study area and geographical context. Marked localities in Tabernas, Pulpí, Campohermoso and Cabo de Gata, supports known populations of *Linaria nigricans*

As model species we have selected *Linaria nigricans*, because it characterizes diverse communities included in the Habitat Directive of the European Union (92/43/CEE; 2230 *Malcolmietalia dune grasslands* and 6120 *Xeric sand calcareous grasslands*, priority habitat, (Anónimo, 1992)), and is threatened by the spreading of greenhouses. *Linaria nigricans* is a short-lived therophyte, restricted to four localities of south-eastern Spain (see **Figura 1**: Campohermoso, Taber-

nas, Pulpí, and Salinas). It occupies flat areas with arid and sandy soils, and does not tolerate alterations in its habitat (e.g. compaction or changes in the soil properties). The plant has been catalogued as *Endangered* (EN) (Cabezudo *et al.*, 2005) and is protected by the Law of Flora and Fauna of Andalusia (Anónimo, 2003). *Linaria nigricans* inhabits annual grasslands with a great number of accompanying mesogean taxa, such as *Leysera leyseroides*, *Ifloga spicata*, *Am-mochloa palaestina*, *Lobularia lybica*, *Filago* spp., and other species with narrow biogeographical ranges (e.g. *Silene littorea* subsp. *adscendens*, also a locally endemic species; *Asphodelus tenuifolius*; *Eryngium ilicifolium*; *Ononis sicula*).

For more than 30 years, greenhouses have been spreading in SE Iberian peninsula, and are currently occupying a vast area of the plains near the coastline. This spread has been encouraged in recent years by the construction of new infrastructures intended to guarantee the water supply for the intensive agriculture in the region. In our framework, to model the potential distribution of greenhouses, they were considered to behave in a way comparable to a biological species in terms of topographical, climatic, and geographical requirements, with a great capacity of adaptation to relatively adverse conditions by virtue of technological innovations that improve productivity.

### 3.3.2 Sampling and environmental data

For two years, field surveys were conducted in the area occupied by *Linaria nigricans* to obtain presence records using GPS devices (ensuring at least 40 meters between two nearby presence records), resulting in 500 points. To work with equal sample sizes, 500 presence records of greenhouses were taken on a Landsat image (year 2000) by a random stratified sampling using the same minimum distance criterion. Each group of presence records was randomly split into two groups: 300 points as training dataset and 200 points as the evaluation dataset (Guisan y Zimmermann, 2000). The real area occupied by greenhouses was mapped by landcover analysis with neural network over a Landsat

*La expansión de invernaderos en la provincia de Almería es uno de los casos más espectaculares de destrucción irreversible de hábitats naturales de gran valor por su singularidad.*

*Gracias a estos esfuerzos de muestreo, Linaria nigricans es probablemente la especie amenazada de la que mejor se conoce su distribución.*



image (year 2000) by means of the MLP<sup>1</sup> modul included in Idrisi Andes (Eastman, 2006).

Topographical variables (as slope, aspect, and potential solar radiation) to be used in the modelling framework were derived from a digital-elevation model (20 m resolution). The climatic coverages of the period 1999-2005 (means, minimum, maximum, and range of temperature, air-moisture, and rainfall) were computed according to the interpolation techniques described by Ninyerola *et al.* (2000) using climatic records of the Andalusian Agroclimate Information Network. The coverages of distances from water resources (reservoirs and desalination plants), dams, roads, and urban centres were computed using distance modules of the Idrisi Andes software and land-use coverage as the data source. To avoid multicollinearity in the dataset, a selection process based on the Pearson's correlation index was applied, using Biomapper 3.0 (Hirzel *et al.*, 2004). Variables were eliminated sequentially until achieving a maximum correlation index lower than 0.70, leaving a total of 12 variables (see Table 1).

*Los invernaderos, además de cierta temperatura y radiación solar, deben estar cerca de vías de comunicación y fuentes de recursos hídricos, aunque la mayoría disponen de pozos propios.*

### 3.3.3 Species Distribution Models

SDMs are designed to compute the spatial distribution of the suitable habitat for a given species, using as inputs presence records and a set of predictor variables with influence on the distribution of the species, or correlated with other variables that do (Elith y Barry, 2006). Since different methods of habitat modelling offer different results from the same data (Thuiller, 2003; Segurado y Araújo, 2004; Elith *et al.*, 2006), we built models of *Linaria nigricans* and greenhouses using six different only-presence-based algorithms: Bioclim (Farber y Kadmon, 2003) Domain (Carpenter *et al.*, 1993), ENFA (Hirzel *et al.*, 2002), GARP (Peters y Stockwell, 1999), and MaxEnt (Phillips *et al.*, 2006). Each algorithm was calibrated using the settings recommended by the authors, the presence samples and the selected variables (see Table 1).

*La idea consiste en modelar la distribución usando varios algoritmos, y seleccionar para cada especie el más precisos según un test estadístico.*

---

<sup>1</sup> Multi-Layer Perceptron Classifier

Type	Variable	Units
Topography	Elevation	m
	Slope	degrees
	Direct radiation (yearly maximum)	Wh/m <sup>2</sup>
	Direct radiation (yearly average)	Wh/m <sup>2</sup>
Climate	Direct radiation (yearly range)	Wh/m <sup>2</sup>
	Rainfall (annual)	mm
	Temperature (yearly maximum)	°C
	Temperature (yearly minimum)	°C
	Temperature (yearly range)	°C
	Air moisture (yearly maximum)	%
Distances	Air moisture (yearly minimum)	%
	Air moisture (yearly range)	%
	To desalination plants	m
	To damns	m
	To human populations	m
	To any roads	m

**Table 1:** Environmental variables used to model the potential distribution of *L. nigricans* (all but Distances) and greenhouses (all variables).

### 3.3.4 Evaluation and selection of models

To test the model's accuracy, we applied the area-under-curve (AUC) ROC method (Fielding y Bell, 1997), widely used to evaluate distribution models (Arntzen, 2006; Elith et al., 2006). To work with only-presence datasets, absence records were replaced by 10000 samples of 200 completely random plots, which were then compared one by one with the evaluation sample (200 presence records) to compute 10000 AUC values.

For a given model, with the random selection of a presence record and a random plot, the mean of the 10000 AUC values (AUCm) expresses the probability that the suitability value of the presence record will be higher than the suitability value of the random plot.

This method is similar to the AUC test implemented in the MaxEnt software, and it provides a common frame of comparison between models. To detect significant differences in

*Aunque AUC presenta limitaciones importantes para evaluar modelos, su uso en este trabajo se ajusta a sus posibilidades, comparando modelos para la misma especie y mismo área de modelado.*

model performance between algorithms, AUC values were normalized using arcsine transformation to apply a repeated measures ANOVA test, using the algorithm as factor (Segurado y Araújo, 2004). Models with higher AUC<sub>m</sub> were selected as the most accurate ones. Differences in model performance between groups (*Linaria nigricans* and greenhouses) were explored by ANOVA test.

### 3.3.5 Postprocessing of distribution models

En esta sección asumo que en un territorio sin usos del suelo, el MDE proporciona la extensión del hábitat original de la especie. A eso es a lo que llamo POTENTIAL HABITAT. CURRENT HABITAT es POTENTIAL HABITAT pero restringido por los usos del suelo.

The *Linaria nigricans* model selected (called POTENTIAL L.n. HABITAT) was masked, assigning zero values to the land uses incompatible with the survival of the plant, resulting in the model CURRENT L.n. HABITAT. The same masking operation was applied to the model of the greenhouses (called POTENTIAL Greenhouses HABITAT), assigning zero values to cells in which there were greenhouses and other land uses not available for new greenhouses (occupied by roads, cities, etc.), resulting in the model CURRENT Greenhouses HABITAT, which shows the distribution of suitability to build new greenhouses (potential areas to build new greenhouses in the future).

### 3.3.6 Extinction-risk model

La concurrencia en una celda de altas idoneidades del hábitat para *L. nigricans* e invernaderos indica una cierta probabilidad de que una población que la ocupe sea destruida por la construcción de un invernadero.

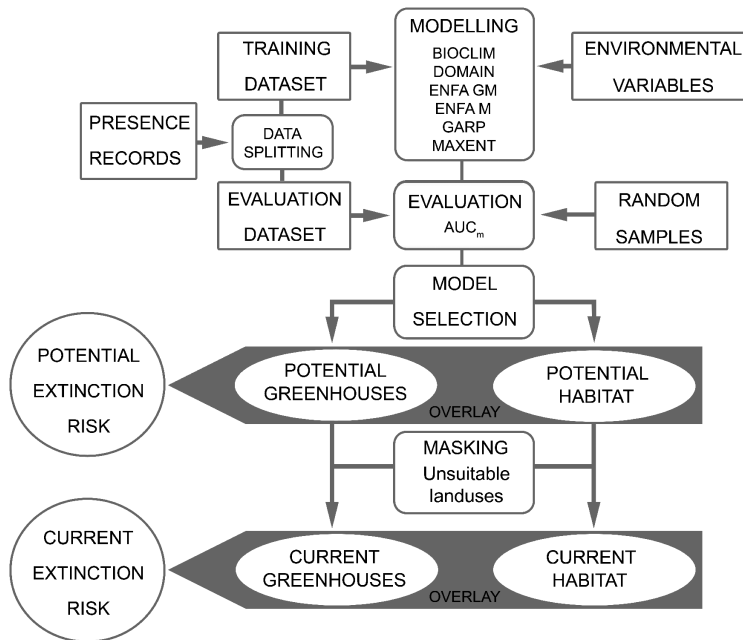
The extinction-risk model for *Linaria nigricans* populations is based on the idea that higher habitat suitability values of *Linaria nigricans* and greenhouses concur for the same cell, and thus the greater is the probability that in each cell the plant would exist and be destroyed by a future greenhouse construction.

To test our idea, we constructed a model called POTENTIAL EXTINCTION-RISK, by combining the POTENTIAL L.n. HABITAT and the POTENTIAL Greenhouses HABITAT coverages, multiplying them and scaling the result to values of 0 to 100 in order to provide a relative extinction-risk scale. We evaluated the success of the coverage POTENTIAL EXTINCTION-RISK by predicting the destruction of suitable habitat for *Linaria nigricans*, calculating the percentage

of cells really occupied by greenhouses (using the map of greenhouses obtained from the Landsat image analysis) for each extinction-risk level. The relationship between the invasion of greenhouses and extinction-risk level was tested by curve fitting in order to find a mathematical function describing the behaviour of the model.

Afterwards, we built a new model combining CURRENT L.n. HABITAT and CURRENT Greenhouses HABITAT in order to generate CURRENT EXTINCTION-RISK, which shows the areas with the highest possibilities of spatial conflict between the plant populations and the greenhouses in the future (see Figure 2). This model was analysed to establish the relative risk level of the known populations of *Linaria nigricans*.

*De POTENTIAL EXTINCTION-RISK solo evaluó su capacidad para localizar zonas de construcción de nuevos invernaderos, cuando realmente debería evaluarse usando puntos en los que se haya constatado la destrucción de una población de *L. nigricans* por un invernadero, pero no dispongo de datos suficientes para efectuarla.*

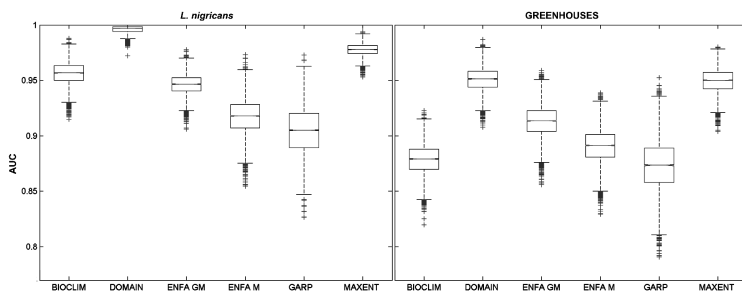


**Figure 2:** Flowchart showing the steps followed in the model construction. Rectangles indicate input data, rectangles with round corners indicate procedures, ellipses show intermediate results, and circles final results

### 3.4 RESULTS

The ANOVA tests confirmed significant differences between algorithms in terms of AUC values (*L.nigricans*:  $f=134102$ ,  $p=0.00$ ; greenhouses:  $f=58831$ ,  $p=0.00$ ). Domain was the best algorithm to model the distribution of *Linaria nigricans* ( $AUC_m = 0.996 \pm 0.00010$ ) and the greenhouses ( $AUC_m = 0.950 \pm 0.00003$ ) (see Figure 3). The ANOVA results of the comparison between groups shows a significantly better performance of the models corresponding to *Linaria nigricans* ( $f=372290$ ,  $p=0.00$ ).

*Al contrario que en otras comparativas entre algoritmos, en este caso Domain fué el que mejor modeló la distribución de L. nigricans y los invernaderos, aunque MaxEnt se le situó muy cerca.*



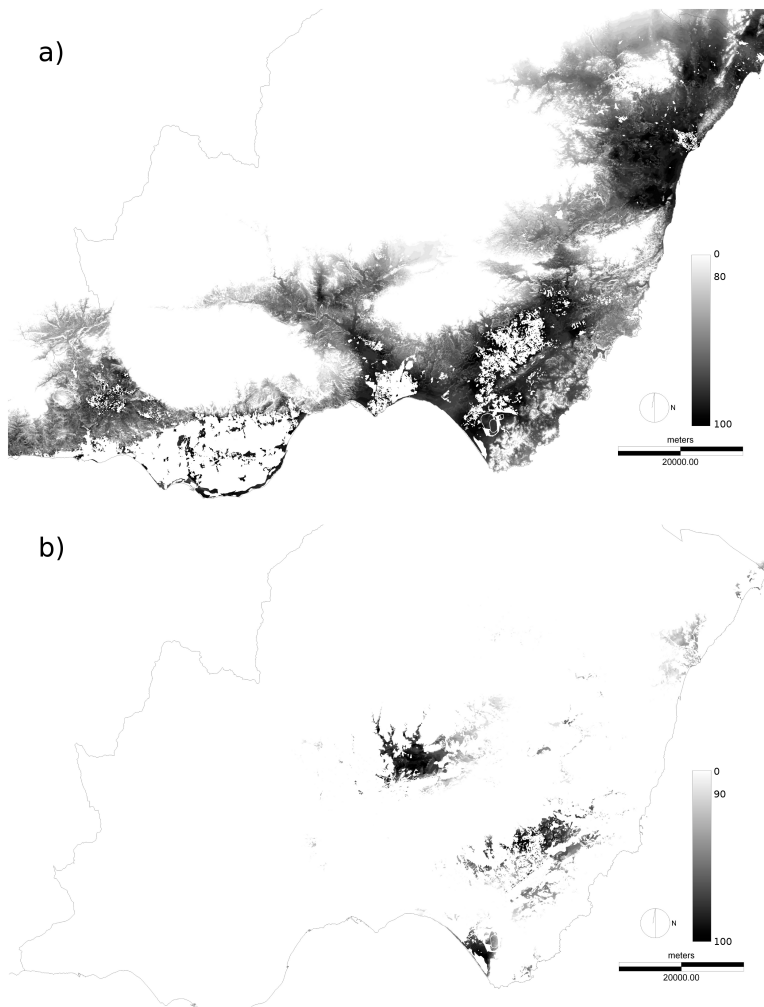
**Figure 3:** Results of model evaluation. The boxplot represents the mean (centre lines), standard error (knots), standard deviation (box), range (dot lines) and outliers (crosses) of the 10000 AUC values computed for each model

Figura 4 shows the coverage selected to build the EXTINCTION RISK models. Note that the POTENTIAL *L.n.* HABITAT coverage (Figure 4a) represents the potential distribution of the plant without taking into account the land uses, while the CURRENT *L.n.* HABITAT (Figure 4b) represents the current distribution, restricted by the land uses incompatible with the survival of the species populations.

*Según el ajuste de la curva, cuanto mayor es el nivel de riesgo de una celda, mayor es el porcentaje de celdas de ese nivel que realmente han sido ocupadas por invernaderos.*

The curve fitting analysis applied to the POTENTIAL EXTINCTION RISK model indicates a robust polynomial relationship ( $3^{rd}$  order;  $R^2 = 0.9655$ ;  $RMSE = 3.807$ ) between the relative extinction-risk level and the percentage of cells occupied by the greenhouses (Figura 5).

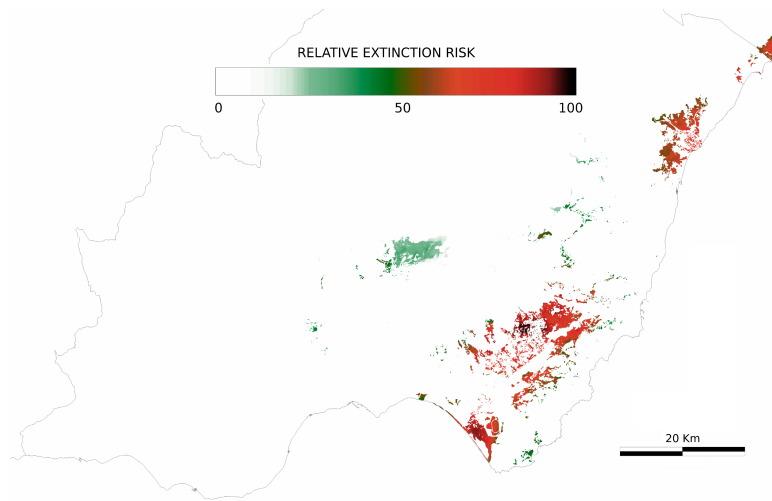
The model CURRENT EXTINCTION RISK (Figure 5, combination of CURRENT *L.n.* HABITAT and CURRENT Green-



**Figure 4:** SDMs for *L. nigricans* and greenhouses: **a)** CURRENT Greenhouses HABITAT; **b)** CURRENT *L.n.* HABITAT.

house HABITAT) shows the areas with the highest possibilities of spatial conflict between the plant populations and the greenhouses in the future; these areas determine the highest risk of destruction for the *Linaria nigricans* populations.

Of the 500 *Linaria nigricans* presence records available, 40% remained above the risk level 80, corresponding to a single population situated in Campohermoso (see Figure 1), and 25% between the risk levels 20 and 40 (population



**Figure 5:** CURRENT EXTINCTION-RISK for *L. nigricans* caused by greenhouses expansion. The areas with higher values indicate suitable habitat for *L. nigricans* with higher probability to be occupied by greenhouses.

of Tabernas; see Figure 1). The remaining 35% are not in danger from greenhouse construction because they are inhabit protected areas (Cabo de Gata-Nijar Natural Park) or other places not suitable for greenhouses.

Según el modelo de riesgo, un 40% de los puntos de presencia de *Linaria nigricans* están en peligro. Nuestros últimos muestreos en Campohermoso en el 2007 confirman que el proceso de destrucción del hábitat continúa.

### 3.5 DISCUSSION

#### 3.5.1 Modelling and model selection

Many works have compared the success of different modelling algorithms in order to identify those that offer the best results (Elith *et al.*, 2006; Phillips *et al.*, 2006; Segurado y Araújo, 2004). Although there are methods that generally function better than others, it is not possible to determine *a priori* which ones will be the most appropriate to work with a particular species and a given set of environmental variables. Therefore, taking into account the nature of our dataset, we used different only-presence-based algorithms in order to select the best model possible in each case. The differences in terms of the results given by the different

algorithms roughly coincide with those reported elsewhere (Loiselle *et al.*, 2003; Elith *et al.*, 2006; Ward, 2007), observing that Domain and Maxent amply surpass Bioclim and Garp in their AUC values.

But the evaluation of only presence-based models is still under discussion and there is no single method appropriate to assess them (Hirzel *et al.*, 2006). To evaluate only presence-based models using the AUC method, some authors have solved the lack of absences by generating random plots to be used in the confusion matrix as absences, calling them *pseudo-absences* (Zaniewski *et al.*, 2002; Engler *et al.*, 2004), or *random instances* (Phillips *et al.*, 2006). Our evaluation method follows these guidelines, but it is slightly different, because during the initial phase of our work, we confirmed that two models tested with two different single random samples could offer significantly different AUC values, changing the choice of the best model according to the spatial distribution of the random sample. We solved this inconsistency in model selection using 10000 different random samples to evaluate each model. This bootstrapping approach poses a problem in the possibility of generating random instances in suitable habitat areas, but working with large presence samples neutralizes the negative effect that this circumstance can provoke (Engler *et al.*, 2004).

The ANOVA test between groups shows that greenhouse models offered a worse fitting index than did those of *Linaria nigricans* (Figure 3), probably because the greenhouses have still not completely occupied their potential distribution area, so the model represents only the areas suitable for greenhouses, and not the really occupied areas. In addition, some greenhouses have a greater ecological tolerance range than the mean, favoured by technological innovations (heating, refrigeration, double- and triple-layered plastic, rigid roofs, etc.) implemented to achieve greater productivity. This group of unidentified outliers may add unexplained variance to the SDMs corresponding to the greenhouses. Another element that adds variation to the greenhouse model is the human factor, which acts in the form of territorial and urban planning, or in the form of decisions of landowners to use the space for greenhouse construction or not.

*Este mismo método de evaluación volverá a aparecer en el Capítulo 7.*

*Hoy sé que no es correcto comparar el ajuste de modelos para distintas especies según su ajuste AUC.*



One of the most noteworthy results of this work is that, despite these drawbacks, it is possible to determine the potential distribution of greenhouses, as a *biological invasion*, using modelling methods normally applied to organisms. This opens a new field of applicability for species-distribution models, as long as the factors determining the distribution of the entity to model are known and can be spatially represented in GIS coverages; our results show that is possible to model the spreading of greenhouses.

*Esta posibilidad de modelar la distribución de los invernaderos como una especie biológica fué la que dió la idea para el siguiente capítulo.*

The CURRENT L.n. HABITAT coverage represents the current habitat distribution of the plant restricted by land uses, including areas where the presence of *Linaria nigricans* was not confirmed in the field sampling. This can have at least two causes: **i)** a well-known behaviour of the presence-only based algorithms, which tends to overestimate the potential distribution area, because the models represent the distribution of the fundamental niche (not the real one) of the species (Zaniewski *et al.*, 2002); **ii)** the environmental variables do not cover all the factors that influence the distribution of the species (Guisan y Zimmermann, 2000; Costa *et al.*, 2008), and therefore the distribution of *Linaria nigricans* can be influenced by physiological factors or biotic interactions such as soil compactness or competitive exclusion, which cannot be reflected as GIS layers. Nevertheless, the model is useful as an objective spatial criterion for analysis in biodiversity conservation.

### 3.5.2 *Extinction-risk model*

The results of the curve-fitting analysis applied to the POTENTIAL EXTINCTION-RISK (Figure 5) are interpreted as a significant trend of the areas with greater relative extinction-risk to be really occupied by greenhouses, so this model is capable of predicting the destruction of *Linaria nigricans* populations caused by the construction of greenhouses. Though this result cannot be applied directly to the CURRENT EXTINCTION-RISK model, it would be expected to work similarly; that is, the areas with higher relative extinction-risk may have a higher probability of being

occupied by greenhouses. A visual comparison of the CURRENT EXTINCTION-RISK (Figure 5) with the CURRENT L.n. HABITAT (Figure 4b) shows that a large part of the potential distribution of the plant is situated in areas suitable for greenhouses in the future.

In the locality Campohermoso (see Figure 1), the growth of the surface area occupied by greenhouses since the 1960s is estimated at 100 ha per year (authors' unpublished data). Consequently, the habitat is being directly destroyed by greenhouse construction and degraded by intense fragmentation. In Tabernas the preservation of the species has been compatible with traditional cultivation techniques based on leaving the ground fallow; however, currently these traditional systems are being abandoned and some greenhouses are being established in the zone, although they do not yet directly affect the plant populations. The greenhouses constitute the main cause of the destruction of the habitat of *Linaria nigricans*. Up to 65% of the known populations of *Linaria nigricans* are at risk of extinction and thus require management and protection plans for habitats that take into account the spreading of the greenhouses, taking into account that the model species represents habitats of interest included in the Habitat Directive of the European Union (Anónimo, 1992).

*El abandono de los cultivos en Tabernas supone un gran problema para L. nigricans, que tiene problemas para desarrollar poblaciones en sustratos compactados por la falta de roturación.*

The success of the extinction-risk model depends on the rate of spreading of the greenhouses. Up to now, this spread could be limited by the scarcity of water resources, but the construction of new desalination plants and water channels may encourage continued expansion. If so, there is a major risk that numerous populations of *Linaria nigricans* will disappear, and thus the important accompanying species and habitat. To recalibrate the risk model, upgraded potential distribution models of greenhouses can be computed including information about new desalination plants in order to estimate changes in the risk of habitat destruction, especially in areas currently assigned low risk levels.

The building of extinction-risk models based on SDMs generates predictions based on solid and objective criteria. These predictions are useful for conservation plans because they enable us to identify potential trends of threatened

plant populations and establish the necessary corrective measures. This modelling framework offers an interesting tool applicable to Environmental Impact Assessments, for the management of natural resources and the conservation of biodiversity.

#### *Acknowledgments*

This research has been financed by the Spanish Ministerio de Education y Ciencia (REN2003-09427-Co2 project), and the Andalusian Consejería de Innovación, Ciencia y Tecnología de la Junta de Andalucía (RNM 1067 project). The authors are very grateful to N. Donadio, J. Cabello, C. Oyonarte, M. Piquer, D. Alcaraz, H. Schwarzeger, V. Vargas and J.L. Caparrós for their technical support and help during field work. We are also indebted to David Nesbitt for the English translation.

## GREENHOUSES, LAND USE CHANGE, AND PREDICTIVE MODELS: MAXENT AND GEOMOD WORKING TOGETHER

---

### ABSTRACT

We have developed a methodology which predicts the expansion of greenhouses and evaluates the results, combining a species distribution model (MaxEnt) and a simulator of land use change (Geomod). In the simulations, we take into account not only the effect of different environmental variables governing greenhouse expansion but also the spatial distribution of the error. The method has been tested on a region of SE Spain to establish future greenhouse-expansion scenarios. The results indicate that the combination of MaxEnt and Geomod improves the predictive capacity, as well as the functional interpretation of the land use change models. **Keywords:** Geomod, MaxEnt, land use, distribution model.

### 4.1 INTRODUCTION

In the context of global change, the study of land use change takes on great relevance because small changes at the local scale (plots of a few ha), added together, can exert an impact on the scale of the entire planet. An example is the deforestation of tropical jungles, which diminishes atmospheric carbon fixation, imposing long-term consequences on global climate (Dixon *et al.*, 1994). An analogous problem of emerging importance involves the expansion of greenhouses, a form of industrial agriculture that is developing on a grand scale in certain regions of the planet. In 1999, an estimated 682,000 ha were occupied by greenhouses throughout the world, especially in China (380,000 ha), followed by Mediterranean countries (161,300 ha in France Italy, Spain, Greece, Turkey, Morocco, and Algeria (Takakura y Fang, 2002).

The problems arising from the spread of greenhouses are directly related, on the one hand, to the natural resources available in the affected region (biodiversity, natural habitats, water resources, etc.) and, on the other hand, to human resources (nearby populations). The construction of greenhouses covers the soil, depriving it of its ecological functions (evapotranspiration, infiltration of precipitation, supporting habitats, etc.), and it degrades the dynamics of natural habitats by fragmenting and destroying them. Greenhouse crops, though designed to make maximum use of irrigation, nevertheless demand huge quantities of water, altering the regime of aquifers. Other problems associated with greenhouses that can concern human health are plastic waste and organic debris contaminated by pesticides and fertilizers.

In the last two decades, the European food market has generated a high demand for fresh vegetables and fruits, triggering the uncontrolled proliferation of greenhouses in productive regions. The growth rate of greenhouses and the lack of a territorial management policy have wreaked havoc, inflicting grave environmental repercussions. In this context of uncontrolled land use change, management plans are indispensable for balanced regional development in which the economy and natural conservation are in balance.

Some GIS-based methods are useful to design and improve land use management plans, such as the land use and cover change simulations (LUCCs: cellular automata, Geomod or Markov chains), which experimentally replicate the transition between land uses (Pontius *et al.*, 2001; Jantz *et al.*, 2004; Pontius y Pacheco, 2004; Aguilera, 2006). Other applicable methods are the species distribution models (SDMs: Bioclim, GARP, MaxEnt, etc.), which provide knowledge on the potential distribution of targeted species and are increasingly in use for the design of conservation plans (Guisan y Zimmermann, 2000; Posillico *et al.*, 2004; Johnson y Gillingham, 2005).

In this paper, we propose a method to predict land use change based on the integration of SDMs and LUCCs. The main idea is to use MaxEnt to compute distribution models, and use them in Geomod as suitability maps to perform better land use change simulations.

The main objectives of this paper are:

- Compare simulations performed by Geomod used in stand alone mode with the combined simulations performed by Geomod and MaxEnt, to test the feasibility of integrating the two methods.
- Introduce Procrustes analysis as a tool to evaluate the spatial agreement between simulations and ground-truth information.
- Introduce an easy method to compute the spatial distribution of certainty in Geomod simulations in order to generate certainty maps for assessing simulation accuracy.
- Test the proposed method in the period 1987-2001 (using 1987 data to calibrate and 2001 data to evaluate) to perform simulations for the period 2001-2010, in order to provide and explore three future scenarios of spreading of greenhouses.

*El objetivo principal de este trabajo es desarrollar una metodología basada en la combinación de un MDE, MaxEnt, y un simulador de usos del suelo, GEOMOD, para mejorar el pronóstico de la expansión de invernaderos.*

#### 4.2 TEST AREA, DATA SETS AND TOOLS

##### *Test area*

The test area selected was the province of Almeria SE Spain), located between 3.14°E and 1.62°E longitude and 36.6°N and 37.46°N latitude. The surface area analysed is 7,171 km<sup>2</sup>. The climate is Mediterranean, with rainfall of 200-300 mm and means annual temperatures of 16-17°C. Geologically, post-orogenic sedimentary materials predominate, and the landscape is dominated by a mosaic of chamaephyte plant communities, xerophytic grasslands and varied communities of annual plants. Greenhouses have been spreading in the area since 1960, occupying around 37,000 ha in 2001.

## Data sets

### *Environmental and geographical variables*

From a digital elevation model<sup>1</sup> of 20 m resolution a total of 11 topographical variables were derived: elevation, slope, northness, southness, eastness, westness (in gradient from 100 to 0), direct solar radiation (mean, minimum and maximum computed by the Solar Analyst extension for ArcView 3.2), topographical wetness index (TWI) and sediment transport index (STI), both computed in ILWIS 3.4<sup>2</sup> Open using the Flow\_indices script<sup>3</sup>.

From road maps from 1987, 2000 and 2006 (2006 map includes roads under construction expected for 2010), we mapped the distance to roads of first, second, third, and fourth order (motorways and national highways, regional roads, provincial roads and local roads, respectively) for years 1987, 2001 and 2010. For each year, an *accessibility index* coverage was built, computed by a weighted mean of the distance coverages. Weights were: 1 for first order roads, 0.75 for second order roads, 0.50 for third order roads and 0.25 for fourth order roads. The weighted sum was scaled into values from 0 to 100 using the module Stretch of the Idrisi Andes software<sup>4</sup>.

Coverages of distances to water resources in years 1987 and 2001 were drawn using the cartography of water infrastructures of the regional government of Andalusia. Distances to water resources in 2010 were computed using the locations of future desalination plants projected by the Water Plan of the Ministry of the Environment of Spain. Areas not suitable for greenhouses (towns, lakes, natural parks) were masked in the datasets so as to be excluded from the analysis.

To avoid high correlation between variables in the dataset, we used Biomapper 3.0<sup>5</sup> (Hirzel *et al.*, 2006), which com-

---

<sup>1</sup> Provided by the Environmental Information Network of the Andalusian Regional Government

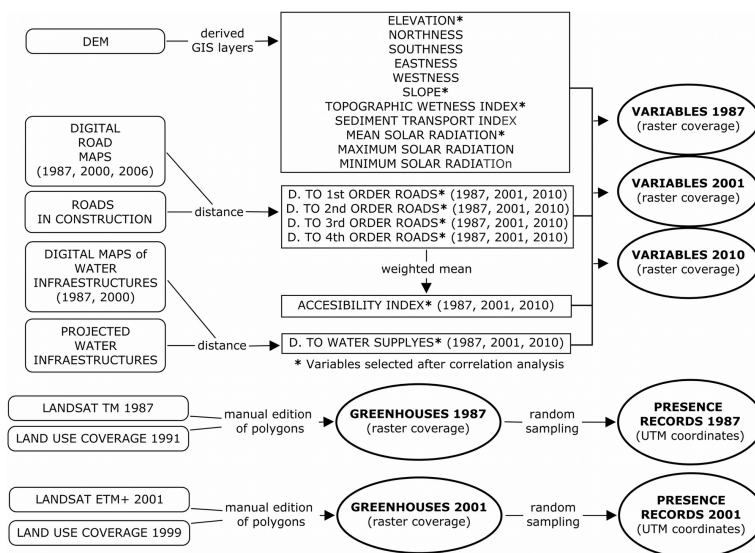
<sup>2</sup> <http://52north.org/>

<sup>3</sup> Available at <http://spatial-analyst.net>

<sup>4</sup> <http://www.clarklabs.org/>

<sup>5</sup> <http://www2.unil.ch/biomapper/>

putes UPGMA<sup>6</sup> trees using Pearson's correlation index as the distance between variables. With 0.75 being selected as the maximum correlation threshold, from each group of highly correlated variables, one was retained. The remaining variables (elevation, slope, topographical-wetness index, mean solar radiation, accessibility index, distance to water resources, and the distances to roads of first, second, third, and fourth order) were used to compose three data sets corresponding to the years 1987, 2001 and 2010, which had in common topographic variables but differed in the values of the distance variables (Figure 1).



**Figure 1:** Flowchart I. Making of the environmental datasets and presence samples

### Greenhouse coverages and presence records

For the calibration and evaluation of MaxEnt suitability maps and Geomod simulations, presence records and greenhouse coverages are needed. For calculating greenhouse coverages for 1987 and 2001, digital land use maps from the years 1991 and 1999 (stored as polygon layers) were

*Parece un baile de fechas confuso, pero se explica por asincronía en las distintas fuentes de datos. Los años a tener en cuenta son 1987, 2001 y 2010, y para cada intervalo trabajo con los datos más próximos temporalmente.*

6 Unweighted Pair Group Method with Arithmetic mean



manually corrected using Landsat images (RGB composites) of years 1987 and 2001 as reference. The resulting polygon layers were rasterized to determine greenhouse (absence/presence) Boolean coverages. From these coverages, 340 and 471 presence records of greenhouses were collected, respectively, by random sampling, establishing a minimum-distance criterion (at least 1000 m between records) in order to avoid spatial autocorrelation effects from using samples too close together (pseudoreplication).

Although the initial resolution of both data sets (variables and greenhouse coverage) was 20 m, the process that we wished to model occurred at a lower resolution, related to the dimensions of the greenhouses. A study of the mean area of the greenhouses indicated that a pixel size of 80 × 80 m would properly represent the land use change, and therefore all the spatial data were rescaled to this resolution using the module Contract included in Idrisi Andes (pixel aggregation for continuous data and pixel thinning for categorical data).

#### 4.3 MODELLING AND EVALUATION TOOLS

##### *MaxEnt*

MaxEnt (Maximum Entropy), a general purpose method for making predictions from incomplete information, has recently been applied to modelling biological species distribution (Phillips *et al.*, 2006). Successful tests have demonstrated that its results are among the best possible within the broad set of algorithms for distribution modelling (Elith *et al.*, 2006).

The algorithm needs a sample of presence records of the organism and a set of environmental variables of the entire study area to compute the distribution model. The environmental variables and functions representing the interactions among them are called *features*, from which an approximation to the ecological niche of the species is defined. Using presences, features, and a background sample (locations taken randomly from the entire study area) MaxEnt searches iteratively for the probability distribution of the maximum

entropy (the closest to uniform), but subject to one condition: the expected value for each feature under the estimated distribution matches its empirical average (calculated from the values of the feature in the presence records). The probability is computed in terms of *gain* (log of the number of grid cells minus the average of the negative log probabilities of the sample locations), which starts at zero and increases in each iteration, until differences between iterations fall below a predefined *convergence threshold*, or the *maximum iterations* number is reached (Phillips *et al.*, 2006).

The probability distribution is projected onto the geographical space, resulting in a distribution model with a range of values of between 0 and 100, which expresses in relative terms the suitability of the habitat for the species (suitability map). MaxEnt can also project the model over variables representing a different time, to explore simulated past or future scenarios. In order to provide a better understanding in the relationships between variables and presence records, MaxEnt performs a Jackknife test to measure variable importance, and plots the log response curves for each variable.

Greenhouses require a combination of environmental variables (temperature, solar radiation, etc.) and geographical ones (distance to roads, water resources) that influence its productivity. These requirements determine the greenhouse construction site selection in the same way as a biological species selects an appropriate habitat. This quasi-biological behaviour permits the application of MaxEnt to calculate the potential distribution of greenhouses, using the same method as applied to biological species. According to this idea, high suitability values in a greenhouse MaxEnt distribution model indicate areas adapted for the construction of greenhouses.

### *Geomod*

Geomod (Pontius *et al.*, 2001) is a land use change simulator implemented in Idrisi Andes. It simulates the land use change between two categories (e.g. from unoccupied to

*Los requerimientos apropiados para la construcción de invernaderos rentables pueden asimilarse a los que necesita una especie para desarrollar poblaciones viables. Como en el capítulo anterior, trato los invernaderos como una especie biológica.*

*GEOMOD es un simulador de cambio de usos del suelo basado en reglas de vecindad y un mapa de idoneidad para el uso del suelo*

occupied by greenhouses) using as start-up information the beginning and ending time of the simulation, a coverage with the initial state of the two categories, the land area changing in use (indicated by the number of cells), a series of environmental variables from which a suitability map is drawn (determining the areas most prone to use change), and a stratification map (enabling the area to be divided into regions that behave differently).

The simulation is based on certain decision rules:

- Land use change is simulated in only one direction, from occupied to unoccupied or vice versa, but not both simultaneously. If a stratification map is used, Geomod can simulate changes in different directions for different strata.
- A neighbourhood rule should be defined: in the constrained mode, a radius is established for the edge of the initial use patches within which Geomod will search for the areas prone to change. In the unconstrained mode, it searches for transition areas without restrictions on the radius, throughout the entire territory being analysed.
- The suitability map for land use change. Geomod computes a suitability map from a set of environmental variables (that influence land use change) and a coverage of the initial state of land use. The computing method reclassifies each variable into categories, assigning to every new category the value of the percentage of cells occupied by the land use towards which the change is going to be simulated. Finally, a weighted sum of the reclassified variables is used to compute the suitability map. The weighting factor may be equal for all the variables or defined for each one by the user. The values of the suitability map are called *lubrication values*: larger lubrication values implies high suitability for land use change (for more details, see [Pontius y Chen \(2006\)](#)).

### *Procrustes analysis*

The Procrustes analysis (Jackson, 1995) compares the fit between different matrices by linear transformation (translation, orthogonal rotation, and scaling) of the cells (landmarks) of a comparison matrix (e.g. a greenhouse occupation simulation) in order to achieve the best possible fit with a reference matrix (e.g. real greenhouse coverage). Once the comparison matrix is modified (if necessary), the algorithm measures the residual vectors between equivalent landmarks (euclidean distances between equivalent cells of the matrices). The goodness of fit between the two matrix is the sum of the squared errors detected ( $\mathbf{m}^2$ ). This value is standardized between 0 and 1, and the values close to 0 indicates the maximum similarity between the comparison and the reference matrix.

*En la publicación original hay un error, y se omite la explicación del análisis de Procrustes, estando en su lugar la de sensibilidad.*

### *Sensitivity*

Sensitivity (**S**) is the conditional probability that a presence cell in the reference image is predicted correctly in a simulation. It can be calculated from the confusion matrix provided by de Crosstab module of Idrisi Andes, dividing the true presences (correctly simulated cells) by the sum of true presences and false presences (incorrectly simulated presence cells). The result (the true positive fraction) is a measure of agreement between a simulation and a reference image in terms of quantity, without bearing spatial differences in mind. We use this additional evaluation measure to support the results of the Procrustes analysis.

## 4.4 METHODOLOGY AND PRACTICAL APPLICATION

Two simulation phases were executed using Geomod: to test the performance of MaxEnt suitability maps and select the decision rules that best represent the spreading of greenhouses, nine simulations using different combinations of decision rules were performed and evaluated for the interval 1987-2001 (1987 data to calibrate and 2001 data to evaluate simulations). Then, using the selected rules, Geo-

*La idea es generar simulaciones con GEOMOD utilizando distintas combinaciones de reglas y usando como base distintos mapas de idoneidad para invernaderos, unos generados con MaxEnt y otros con el propio GEOMOD, para compararlas, y seleccionar la mejor combinación para la simulación final.*

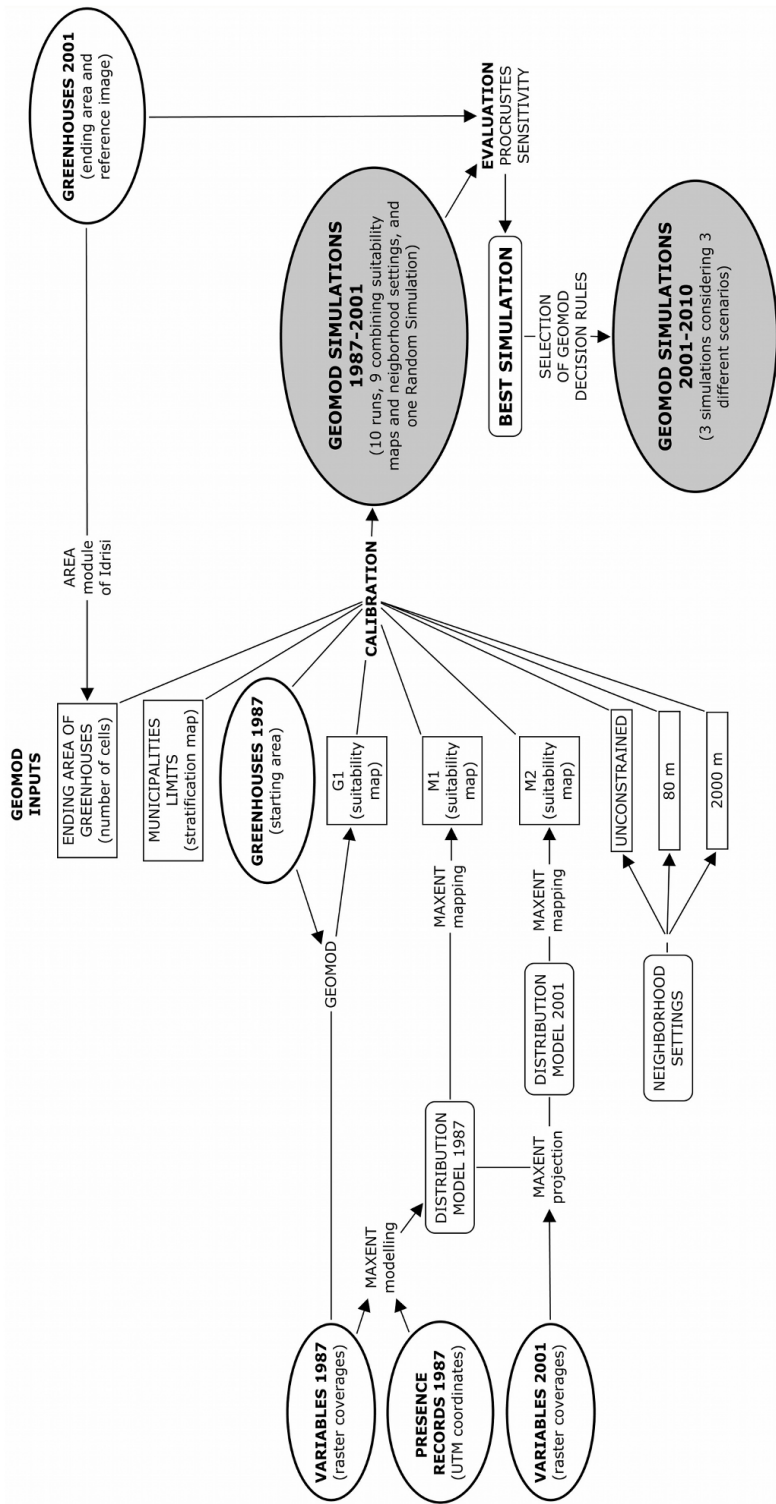
mod simulations considering three different land use change scenarios were performed for the interval 2001-2010.

### *Simulations 1987-2001*

The aim is to select the decision rules available to calibrate Geomod simulations that best describe the spreading of greenhouses in the study area. Suitability maps computed by MaxEnt and Geomod, and different neighbourhood rules were combined in nine performed simulations:

- Three suitability maps were used:
  - **M1**: computed in MaxEnt model, calibrated with the training sample and variables of 1987.
  - **M2**: computed in MaxEnt, model calibrated with the training sample and variables of 1987 and projected over variables of 2001 (using the Projection feature available in the software). Suitability maps computed with MaxEnt were calibrated using the default settings (Phillips *et al.*, 2006).
  - **G1**: computed in Geomod with the greenhouse's coverage and variables of 1987 (using the same weighting factor for all variables).
- Neighbourhood: settings used were **80 m** (1 cell around), **2000 m** (25 cells around) and **unconstrained**.

The simulations were calibrated setting the starting time at 1987, initial area of the greenhouses coverage of 1987 (38,743 cells, 24,795 ha.), ending time at 2001 and final area of the greenhouse coverage of 2001 (58,097 cells, 37,182 ha). All simulations were stratified by municipality limits, representing the diversity of land-management policies in different towns. An extra simulation (unstratified, without suitability map and unconstrained neighbourhood) was performed in order to simulate the random spreading of greenhouses, calling this the Random Simulation (hereafter, **RS**). To clarify this explanation, see Figure 2.



**Figure 2:** Flowchart II. Steps followed to select the suitability map and the neighbourhood setting most appropriate to simulate the spreading of greenhouses

### *Evaluation and spatial certainty of the simulations*

It is important to consider that, on comparing a simulation with the reference image, both share the entire area occupied by greenhouses at the starting time (1987). Consequently, any comparison index that we apply will interpret an inflated degree of agreement between the simulation and the reference image. To avoid this inflation, we eliminated (in all the simulations, the **RS**, and the reference image) the area corresponding to greenhouses in 1987. Therefore, the evaluation took into account only the area of the new greenhouses.

Results were evaluated by Procrustes analysis and sensitivity using the greenhouse coverage of 2001 as the reference image. The simulation with the least  $m^2$  and greatest **S** with respect to the reference image will determine the decision rules that best describe the spreading of greenhouses. Results were tested separately for Procrustes analysis and sensitivity by factorial ANOVA, establishing a *suitability map* (levels: **G1**, **M1** and **M2**) and *neighbourhood* (levels: 80 m, 2000 m, and unconstrained) as categorical predictors. The relationship between  $m^2$  and **S** were assessed by linear regression.

*Las medidas de evaluación escogidas (S y  $m^2$ ) no informan sobre la distribución geográfica de los errores. Para conocerla, analicé (para la mejor simulación) la relación entre el mapa de idoneidad, el resultado de la simulación y la cobertura real de invernaderos para hallar una ecuación capaz de explicarla.*

Usually, when evaluating a simulation by calculating its sensitivity, a homogeneous spatial distribution of certainty must be considered, assuming that all the simulated cells have the same likelihood of being correctly classified. In the real world, if greenhouses are constructed preferably in areas of high suitability (according to the suitability maps) because it favours greenhouse productivity, and Geomod selects as a priority these areas to simulate land use change to greenhouses construction, we can assume that the certainty of the simulation will vary according to the values of the suitability map. Following this reasoning, in the areas of greatest suitability, the probability of finding cells where the presence of new greenhouses has been correctly simulated is higher than in the areas of lower suitability. To test this idea, a joint analysis was made of the best simulation, its suitability map and the reference image (coverage of greenhouses in 2001), in order to:

1. Describe graphically the relationship between the suitability map and the total amount of hits (correctly simulated cells) and errors (incorrect simulated cells) in the simulation, plotting the number of hits and errors against suitability values
2. Find a suitability-certainty function that relates each suitability value to a given probability for a cell to be correctly simulated, which is useful to compute a simulation certainty map.

For this, the percentage of correctly simulated cells was plotted against suitability. The plot represents the specific behaviour of the best simulation, but we are looking for a more general function, capable of predicting approximately the behaviour of different simulations. With this aim, the data was smoothed by means of a moving average (using 25 as span size), and analysed by a polynomial-curve fitting using Octave.

#### *Simulations 2001-2010*

The spreading of greenhouses in the study area has been continuous from 1954, and today the construction of greenhouses is booming, due to the construction of new infrastructures oriented to increase water supply. But the greenhouses are involved in a dynamic market, and the profitability of the crops depends on multiple economic and social factors difficult to predict. Another emerging factor adding uncertainty in the last years is the competition with other Mediterranean countries with cheaper production. To confront this uncertainty we propose three simple scenarios of spreading of greenhouses for the period 2001-2010:

- **A:** Linear greenhouse area growth with the trend identified for 1987-2001.
- **B:** Accelerated growth (20%) over the linear trend due to increased demand.

*La idea es simular la distribución de invernaderos en 2010 (un plazo aceptable para una evaluación a posteriori) a partir de datos de 2001 según tres escenarios simples.*



- **C**: Slowed growth (20%) under the trend due to increased competition from countries with cheaper production (e.g. Morocco and Algeria).

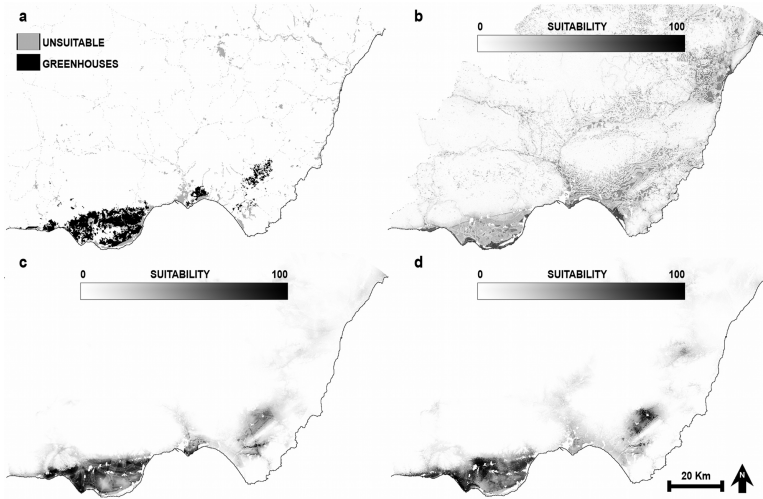
To be used a suitability map in the projections, a MaxEnt **M3**) distribution model was calibrated using the training sample and variables of 2001, and projected over the variables of 2010. Differences in suitability between **M2** and **M3** were computed by map algebra. Using the 2001 greenhouse coverage as the starting image, the suitability map **M3**, and the projected areas in the different scenarios, three Geomod simulations for the period 2001-2010 were performed. To assess the spatial certainty of the simulations, the computed suitability-certainty function was applied to the **M3** suitability map (replacing **M2** values by **M3** values and dividing the result by 100 to translate percent values into probabilities).

#### 4.5 RESULTS

##### *Simulations 1987-2001*

Figure 3 shows the coverage of greenhouses (1987) and the suitability maps used as inputs to calibrate Geomod simulations. Figure 4 summarizes the influence of the modelling variables in the MaxEnt distribution model (suitability maps **M1** and **M2**) and shows the log-response curves of the most relevant variables.

The results of the Procrustes analysis and sensitivity for the nine performed simulations are shown in Figure 5. All the simulations performed better than the **RS** ( $m^2=0.042$ ;  $S=0.021$ ), the results of which are not shown in Figure 5 due to problems of scale. Two simulations calibrated with the suitability map **M2** worked better than the remaining ones: the unconstrained neighbourhood simulation ( $m^2=0.006$ ;  $S=0.4060$ ), selected as the best simulation and the 2,000 m neighbourhood simulation ( $m^2=0.0063$ ;  $S=0.405$ ). Factorial ANOVA test found significant differences in simulations performance between suitability maps, but not between neighbourhood rules (see Table 1 for a summary of factorial



**Figure 3:** Greenhouses in 1987 and suitability maps. **a)** Coverage of greenhouses in 1987 (starting time) used to calibrate simulations and non-suitable areas excluded from the analysis; **b)** **G<sub>1</sub>** Suitability map, computed by Geomod; **c)** and **d)** **M<sub>1</sub>** and **M<sub>2</sub>** suitability maps calibrated in MaxEnt using presence records and variables of 1987 **M<sub>1</sub>**, and projecting the model over variables of 2001 **M<sub>2</sub>**. Dark colours indicate high suitability for the construction of greenhouses.

ANOVA results). Procrustes ( $m^2$ ) and sensitivity (**S**) values were closely correlated (adjusted  $R^2=0.947$ ;  $p$ -level=0.000).

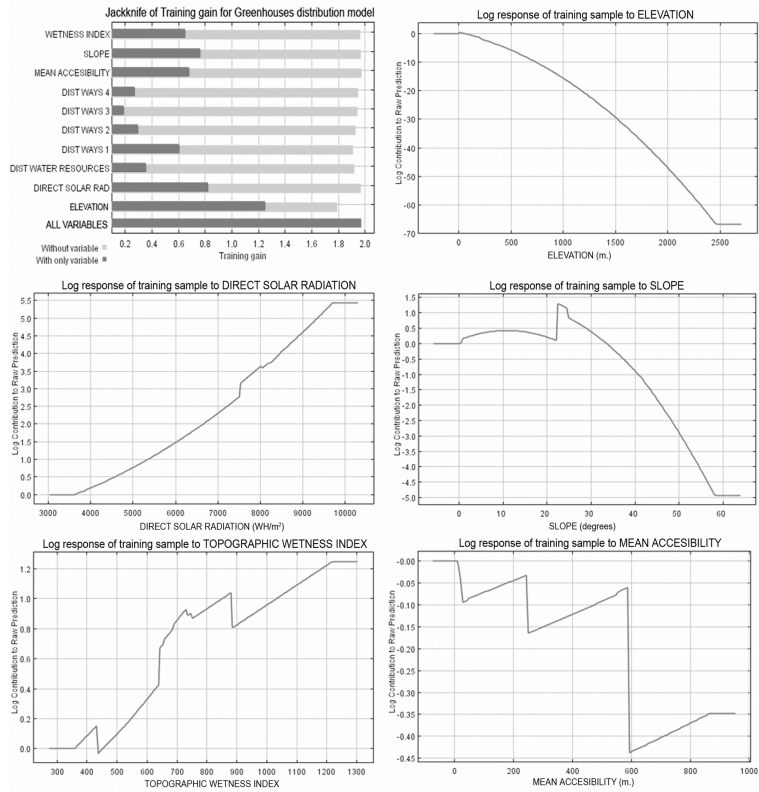
Unconstrained neighbourhood and MaxEnt suitability map (but replacing **M<sub>2</sub>** with **M<sub>3</sub>**) were the settings selected to calibrate and simulate the three future scenarios of land use change.

*Las mejores reglas para simular la expansión de invernaderos con GEOMOD son: distancia no restringida, y mapa de idoneidad calculado con MaxEnt.*

Dependent variable Categorical predictor	$m^2$		<b>S</b>	
	suitability	neighbourhood	suitability	neighbourhood
<b>F</b>	<b>9.132</b>	0.009	<b>29.158</b>	0.108
<b>p</b>	<b>0.032</b>	0.990	<b>0.004</b>	0.900

**Table 1:** Summary of results of factorial ANOVA. Significant values in bold.

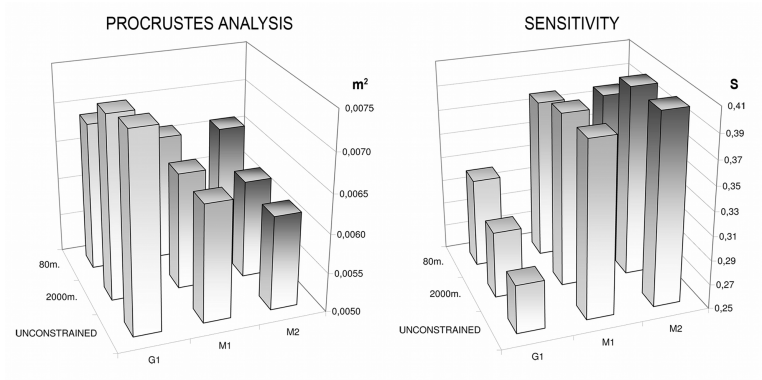
Figure 6 (top) shows the graphical analysis of correctly and incorrectly simulated cells of the best simulation. Figure



**Figure 4:** Jackknife test and response curves of MaxEnt distribution model. Bars plot: dark grey bars indicate model gain when computed with only the variable, and light grey is the model gain when computed with the other variables. Minor differences between the two bars indicate major importance of the variable in the model. Log-response curves: the five most important variables are shown. Values over 0 indicate suitable conditions for greenhouses, whereas the values below zero indicate unfavourable conditions.

6 (bottom) shows the polynomial relationship between the suitability values (of the M<sub>2</sub> suitability map) and the percentage of cells correctly predicted for that suitability value, calculated from the best simulation.

The equation 4.1 expresses the suitability-certainty function ( $R^2=0.99$ ;  $RMSE=1.24$ ). See curve in Figure 6 .



**Figure 5:** Evaluation results. Evaluation of 9 simulations performed in Geomod combining three neighbourhood rules (80 m, 2,000 m and unconstrained) and three suitability maps, one computed in Geomod (**G1**), and two computed in MaxEnt (**M1** and **M2**). Each bar corresponds to a performed simulation. The reference image is the real coverage of greenhouses in 2001 (ending time in the simulations). Low values of  $m^2$  and higher values of  $S$  are indicative of a good agreement between a simulation and the reference image. The best simulation was performed with an unconstrained neighbourhood and the suitability map **M2**.

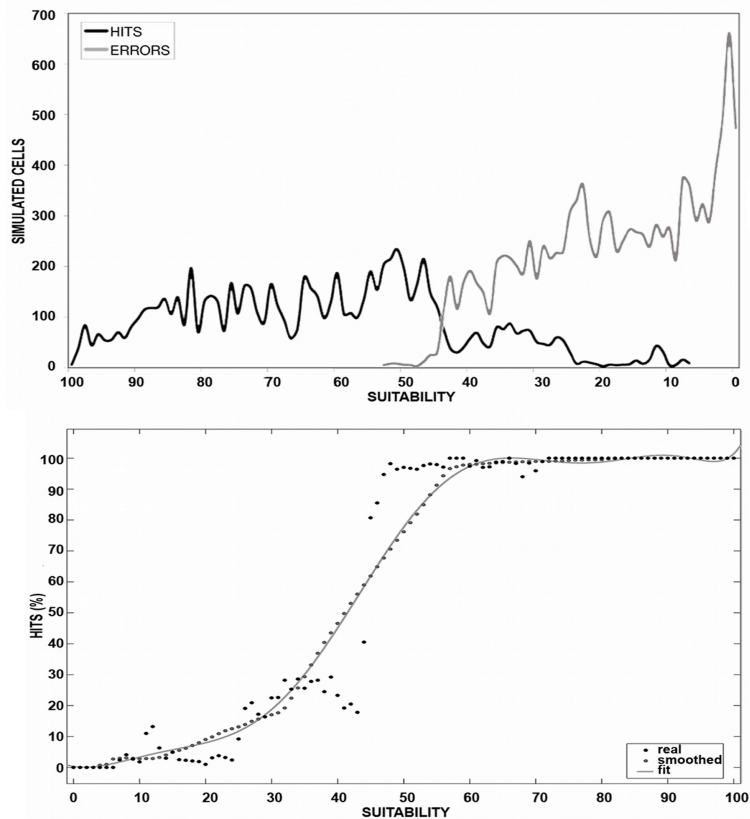
%HITS =

$$\begin{aligned}
 & 1.186 \cdot 10^{-13} \cdot M2^9 - 4.729 \cdot 10^{-11} \cdot M2^8 + \\
 & 7.417 \cdot 10^{-9} \cdot M2^7 - 5.681 \cdot 10^{-7} \cdot M2^6 + \\
 & 2.080 \cdot 10^{-5} \cdot M2^5 + M2^4 - 0.003 \cdot M2^3 + \\
 & 0.102 \cdot M2^2 - 0.311 \cdot M2 + 0.341
 \end{aligned}$$

(4.1)

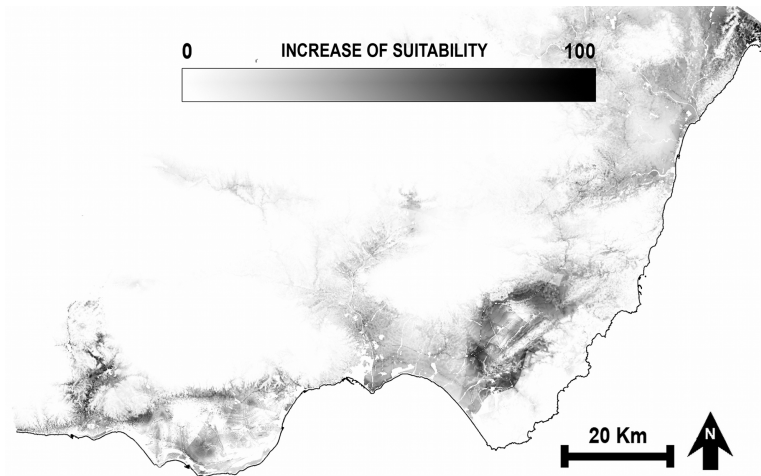
### *Simulations 2001-2010*

Figure 7 shows the differences in suitability between **M2** and **M3**. The construction of new infrastructures (roads and desalination plants) boosts the suitability for greenhouses



**Figure 6:** Certainty against suitability in the best simulation. Top: the plots describe the behaviour of the best simulation in terms of total amount of correctly (hits) and incorrectly (errors) predicted cells against suitability. Bottom: 9<sup>th</sup> order polynomial relationship (see equation 4.1) between the percentage of correctly predicted cells and suitability (**M<sub>2</sub>** suitability map). The black plots represent the real data, and the grey dots the smoothed data. The curve represents the curve fitted to the smoothed data

in areas that already were fulfilling suitable topographic conditions. The relative importance of the variables and the response curves in the **M<sub>3</sub>** distribution model were similar to those of the **M<sub>2</sub>** distribution model, confirming that both models had close similarities.



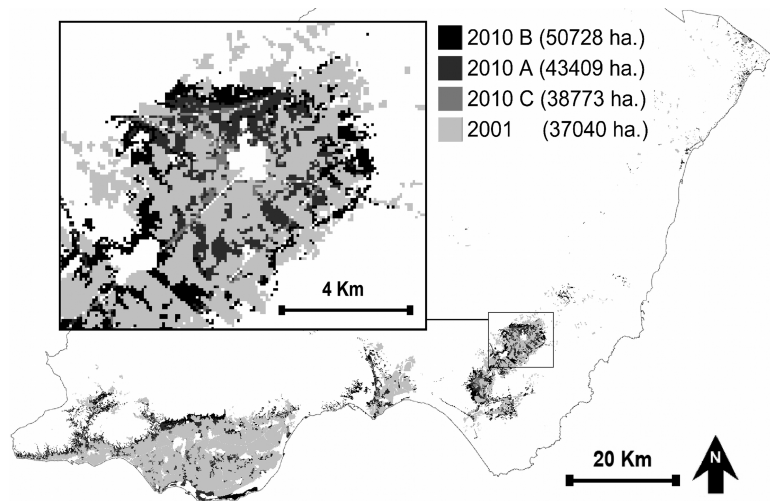
**Figure 7:** Differences in suitability between  $M_2$  and  $M_3$ . Differences in suitability between  $M_2$  and  $M_3$  were computed as  $M_3 - M_2$  in the raster calculator of Idrisi Andes. Higher values are indicative of new suitable areas for the spreading of greenhouses.

**Figura 8** shows the simulations corresponding to the proposed scenarios **A**, **B**, and **C**, compared to the real greenhouse-occupied area in 2001.

#### 4.6 VALIDATION OF THE RESULTS AND DISCUSSION

##### *Simulations 1987-2001*

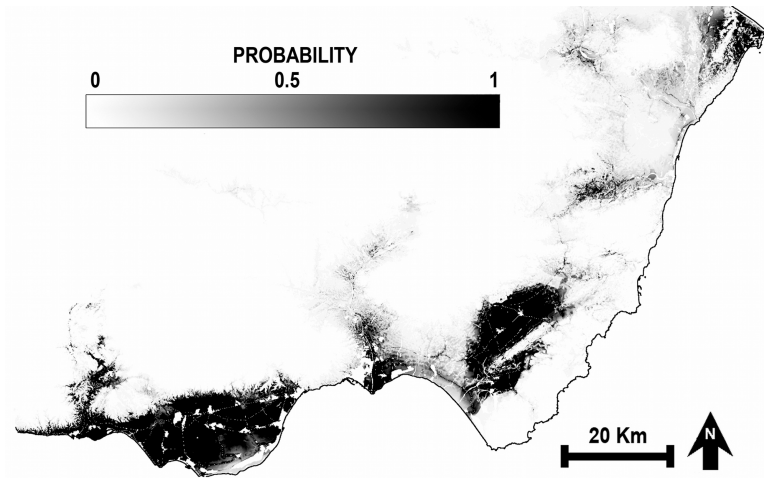
The ANOVA analysis of the evaluation values indicate that the simulations performed with MaxEnt suitability maps work better than those performed with Geomod ones (in terms of  $S$  and  $m^2$ ). While MaxEnt uses scattered presence points as input, Geomod uses the complete land use coverage. If, for reasons other than the suitability of the territory, there is a great concentration of greenhouses in a given area (e.g. historical causes), the combination of variables present at this site takes on relatively high importance with respect to the other combinations of variables in the rest of the territory, resulting in a misleading suitability map. The same can happen in the MaxEnt models when the pres-



**Figure 8:** Simulations of scenarios A, B, and C. Scenario B is the sum of 2010 B, 2010 A, 2010 C, and 2001 occupied areas. Scenario A is the sum of 2010 A, 2010 C, and 2001 occupied areas, etc. The zoomed area is a detail of Campohermoso, a locality with an intense growth of the area occupied by greenhouses in recent years.

ence records provided as input were very close together, but the initial treatment that we applied (minimum distance between presence points greater than 1,000 m) diminish any possibility that aggregation effects could affect the quality of the models. Another difference between the two methods to calculate suitability maps is based on the relative weight given to the variables. Geomod does not use any algorithm to compute weights, and they have to be established by the user (by subjective criteria, or criteria based on previous statistical analysis).

MaxEnt includes a Jackknife test, which automatically computes the relative contribution of every variable to the model. Another advantage of the MaxEnt algorithm successfully explored in this paper is the *projection feature*, capable of projecting a distribution model over variables with values expected for the future. It is a useful tool to explore alternative land use change scenarios bearing in mind expected changes in the values of the variables (accordingly to



**Figure 9:** Certainty map for simulations of scenarios A, B, and C. The certainty map computed from equation 4.1 applied over the  $M_3$  suitability map. The values indicate the probability that a cell with presence of greenhouses simulated by Geomod will really be occupied in 2010.

known information, such as projected roads). Our results support the idea that the MaxEnt algorithm can generate useful suitability maps applicable to Geomod simulations, outperforming the results given by a stand-alone use of Geomod.

The results for the importance of the variables in the suitability maps computed by MaxEnt (Figure 4) and of their response curves (Figure 5) indicate that the fundamental factors influencing greenhouse distribution in the study area are related to topography and distance to roads. The open plains (which coincide with the areas having high indices of topographical moisture) at low altitudes have the temperature, slope, and solar radiation appropriate for greenhouse operation. The factors related to the distances to roads do not appear to be limiting, although the longest distance to firstorder roads (motorways) are related to a lower presence of greenhouses. The variable *accessibility index*, the fourth in importance, accurately summarizes the distances to different types of roads, and it is useful to predict the spreading of greenhouses. Interpretation problems arise with the variable



Distance to Water Resources, because the great majority of greenhouses do not depend on centralized resources such as reservoirs or desalination plants, but rather use their own wells, which pump water from aquifers. This variable is the only one that has lost relevance over time among the 1987-2001 and 2001-2010 models, since desalination of sea water has proliferated on the Almería coast. Even so, the low gain shown by all the models indicates that the contribution of desalination does not significantly affect the results.

In the evaluation of the simulations, we considered two sides: agreement in number of predicted cells, expressed in terms of **S**, and spatial agreement, tested by Procrustes analysis and expressed in terms of  $m^2$ ). Both measures were correlated, but not perfectly because, for two simulations with the same sensitivity (compared with a reference image), there may be differences in the location of the errors detected by the Procrustes analysis. Procrustes analysis is a quick and simple way to assess spatial agreement between simulations and real land use coverages.

The analysis of hits and errors of the best simulation (Figure 6, top) shows that it works better in the section of higher values of the **M2** suitability map (especially in the range 100-50), and the errors increase when suitability trends toward zero. When the percentage of hits against suitability values is smoothed by a moving average (Figure 6, bottom), the pattern remains quite clear, making it possible to find, by means of curve fitting, a function (9<sup>th</sup>-order polynomial, see equation 4.1) describing the behaviour of the simulation.

#### *Simulations 2001-2010*

The construction of new roads and desalination plants can increase the area suitable for greenhouses, as shown in the map of differences between **M2** and **M3** suitability maps (Figure 7). Apart from the increase of suitable area, both models show identical behaviour regarding the relative contribution and the response curve that every variable presents. During the studied periods 1987-2001 and 2001-2010, the

relationships between the presence of greenhouses and the variables that influence their distribution did not significantly change.

Geomod is designed to predict the locations of land use change, not the quantity of area that changed. Therefore, the validity of the simulations is based on a solid interpretation of the data for surface-area growth of greenhouses. Using only the two available sets of temporal data (1987 and 2001), we used a linear estimation to calculate the amount of occupied area in the scenarios **A**, **B** and **C**, but it would be more appropriate to use data from temporal series with a greater number of control points. The problem arises when inflexion points are foreseen in the growth curves of the occupied area, a possibility in the study areas because there is a high degree of saturation (a large area of land that can be occupied by the greenhouses is already occupied) and the resources supporting the greenhouse industry (mainly soil and water) are being depleted. Thus, to our knowledge, scenarios **A** and **C** are probably the closest to reality (see Figure 8).

The certainty map (Figure 9) can be useful to assess the expected accuracy of the simulations when it is not possible to validate them with ground-truth information. Nevertheless, the function used to compute the certainty map of the simulations 2001-2010 has been calculated for a simulation performed for the period 1987-2001 and the suitability map **M2**, there are at least two ideas that may justify its application:

- **M2** and **M3** distribution models are quite similar, and therefore a significant behavioural change in the suitability-certainty function between models is not expected.
- The smoothing of the data by moving the average prior to the curve fitting removes bias and generalizes the function, allowing its application to other simulations performed under the same conditions.

However, it is important to bear in mind the limitations of this application: the function does not take into account

the effect of the area that will predictably undergo land use change. This effect is important because it tends to increase the percentage of correctly predicted cells of the simulation and can influence the relationship between certainty and suitability, altering the shape of the curve and changing the coefficients of the function. This can lead to an erroneous interpretation of the probability values of the certainty map. It would be useful to make an in-depth study of the relationship between suitability and certainty for different simulations to find an equation that can function in a general way in order to associate each simulating cell with a particular certainty value.

#### 4.7 CONCLUSIONS AND PERSPECTIVES

##### *Conclusions*

The combined use of MaxEnt and Geomod provide a series of significant advantages with respect to the stand-alone use of Geomod in land use change simulations:

- Geomod simulations using MaxEnt distribution models as suitability maps significantly outperform simulations calibrated with suitability maps computed by Geomod. In addition, the *projection* feature of MaxEnt makes it possible to explore alternative scenarios by changing the values of the variables used to calibrate the model.
- It is possible to predict accurately the spreading of greenhouses using only topographic variables and distances to roads. In this sense, the proposed *accessibility index* is a useful variable that summarizes distances to different types of roads.
- Relationships between greenhouses and variables are stable in time for the periods studied, allowing the exploration of future scenarios.
- Procrustes analysis is a powerful tool to assess spatial similarity between simulations and ground-truth infor-

mation, and provides a simple and easily interpretable measure of agreement ( $m^2$ ).

- The certainty of Geomod simulations is not spatially uniform. There is a strong relationship between the amount of correctly simulated cells and suitability. This relationship is useful to compute certainty maps to assess the spatial accuracy of simulations.
- The proposed methodology can be applied to territorial management of areas in which greenhouse expansion can represent an environmental problem, as in the Mediterranean countries mentioned in the present work. From the simulations, it is possible to identify the hotspots on which to focus environmental management and conservation efforts.

### *Perspectives*

In the context of global change, the studies on land use change are becoming as relevant as those related to climatic change. Though we have specifically oriented tools, the complexity of the web of factors affecting land use is such that it is difficult to develop highly accurate techniques. It is necessary to continue to delve into the possibilities offered by geographic information technology to formulate predictions that enable us to face coming changes.

The present study seeks to combine two different perspectives: ecological species-distribution models used in biology, and land use change models used in geography. Both approaches combined can generate powerful tools to analyse our changing world and to explore alternative scenarios.

Although the analysis proposed for land use change is applied only to greenhouses, it has other potential applications (perhaps irrigation, urbanization, tourist facilities, etc.). This implies another alternative to the different use of change models currently being used.

## ACKNOWLEDGEMENTS

This research is supported by the Spanish Education and Science Ministry (REN 2003-09427-Co2 project) and the Andalusian Innovation, Science and Industry Regional Ministry (Consejería de Innovación, Ciencia y Tecnología de la Junta de Andalucía) (RNM 1067 project). We are indebted to David Nesbitt for the english translation. We are also grateful to Aitor Gastón and two anonymous referees, whose valuable commentaries improved this paper.

Parte III

PRONÓSTICO DE EFECTOS  
POTENCIALES DEL CAMBIO DEL  
CLIMA EN LA FLORA



## INTRODUCCIÓN A LA PARTE III

---

Fuera del ámbito científico, el Cambio Climático es un tema controvertido, con un color político que no beneficia a nadie. Dentro de la comunidad científica apenas hay rastros de este debate estéril, porque se considera probado que el clima está cambiando, y continuará haciéndolo los próximos decenios. Son muchos los investigadores a lo largo del mundo que recogen pruebas de la respuesta de los organismos a estos cambios: adelantos en la fenología de las especies, migraciones altitudinales y latitudinales, y pérdidas (y ganancias) de eficacia biológica son algunas de las huellas que la nueva situación nos está dejando.

Estas pruebas dibujan una perspectiva futura de la biodiversidad muy distinta de la que tenemos hoy en día, con cambios en la riqueza de especies, extinciones locales, cambios en la distribución geográfica, desaparición de comunidades, dominancia de nuevas especies, y tal vez algunas otras que no llegamos a imaginar.

Un pronóstico preciso de los cambios que acechan a la biodiversidad del planeta (ya bastante maltrecha) sería una gran herramienta de trabajo, porque nos daría la oportunidad de conocer y decidir lo que podemos salvar y como hacerlo. En esta sección se exploran y tratan de mejorar las herramientas de pronóstico, aplicando sus resultados a distintas especies vegetales de hábitats frágiles de Andalucía, tomada como exponente de las regiones mediterráneas tan ricas en diversidad biológica.

Esta tercera parte está compuesta de tres capítulos, que conforman un arco ascendente en cuanto a complejidad de las técnicas empleadas, área de trabajo y número de especies tratadas.

En el Capítulo 5 trata de obtener una vista previa (a alta resolución espacial) del posible comportamiento de cuatro formaciones vegetales de gran interés en el Parque Nacional de Sierra Nevada (encinares, robledales, enebrales y pastizales psicroxerófilos). La técnica utilizada es sencilla, casi



de corte artesanal, pero los resultados revelan unas tendencias de cambio preocupantes. En este capítulo se ofrecen dos novedades: una técnica (Diferencial de Idoneidad) para localizar frentes de avance y retroceso de las poblaciones vegetales útiles para la gestión proactiva de las formaciones vegetales, y un método para dilucidar automáticamente el patrón de sustitución potencial entre formaciones vegetales. Cronológicamente, este ha sido el primer intento conocido de pronosticar el efecto potencial del cambio del clima en una alta montaña mediterránea usando alta resolución espacial.

Los procedimientos de pronóstico como el desarrollado en el Capítulo 5 son muy complejos, con muchos pasos a llevar a cabo casi manualmente por el investigador, lo que limita el número de datos que se pueden manejar en periodos de tiempo razonables. Tratando de superar esta limitación, y entrando de lleno en el mundo de la Econinformática, en el Capítulo 6 se diseña y desarrolla un flujo de trabajo para automatizar completamente el proceso de predicción, facilitando el trabajo con grandes números de especies. Para probar su funcionamiento, se aplica a distintas especies amenazadas de dos hábitats frágiles emblemáticos de Andalucía Oriental: el altiplano árido de Guadix-Baza, y Sierra Nevada, con el objetivo de conocer los distintos efectos que el cambio del clima puede ocasionar a hábitats distintos.

En el Capítulo 6 se añade al flujo de trabajo desarrollado anteriormente el factor migración, que no había sido tenido en cuenta, para tratar de pronosticar los patrones de cambio de la biodiversidad vegetal de Andalucía. Se trata de un trabajo de modelado de gran envergadura, que involucra 200 especies, algunas importantes por su estado de amenaza, y otras, por su dominancia dentro de los ecosistemas.

## SIMULATING POTENTIAL EFFECTS OF CLIMATIC WARMING IN VEGETATION DYNAMICS OF MEDITERRANEAN-ALPINE ECOSYSTEMS

---

### ABSTRACT

In this work, we study an isolated high-mountain (Sierra Nevada, SE Iberian Peninsula) to identify the potential trends in the habitat-suitability of four vegetation types (occupying different altitudinal belts) that might result from a sudden climatic shift.

We used topographical variables and downscaled climate warming simulations to build a high-resolution spatial database (10m) according to four different climate warming scenarios for the 21<sup>st</sup> century. The spatial changes in the suitable habitat were simulated using a species distribution model, in order to analyse elevational shift and potential habitat loss of the plant formations. Thus, the advancing and receding fronts of known occurrence locations were computed by introducing a new concept named *differential suitability*, and potential patterns of substitution among the plant formations were established.

The average mean temperature trend show an increase of 4.8°C, which will induce the vertical shift of the suitable habitat for the vegetation types at an average rate of 11.57 m/year. According to the simulations, the suitable habitat for the higher (and valuable) vegetation types may disappear before the middle of the century. The other vegetation types shows moderate to drastic suitable habitat loss depending on the considered scenario. Climate warming should provoke a strong substitution dynamics between vegetation types, increasing spatial competition between both of them.

In this study, we introduce the *differential suitability* applications into the analysis of potential impact of climate change, forest management and environmental monitoring,

and discuss the limitations and uncertainties of the simulations.

**Keywords:** simulation, MaxEnt, high mountain, vegetation types, elevational shift.

## 5.1 INTRODUCTION

Climatic monitoring during the 19<sup>th</sup> and 20<sup>th</sup> centuries, and the reconstruction of the climate over the last millennium, indicate that the earth's climate is warming (Mann, 2007), and this effect has been increased during the last quarter of the 20<sup>th</sup> century (Jones *et al.*, 1999). Climate change impacts on the biota are known to act on four levels: metabolism, phenology, evolution, and spatial distribution (Hughes, 2000). Evidence also supports the idea that climatic warming has already triggered species distribution shifts in many parts of the world (Thuiller *et al.*, 2005b) and is forcing a consistent temperature-related shift in species ranging from molluscs to mammals and from grasses to trees (Root *et al.*, 2003).

In high mountains the effects of global warming with regard to the biota are amplified. Several works in mountains across the world have reported an increase in the annual average temperature ranging between 0.5-0.8°C, as well as the forced response in the altitudinal migration of plant species of 7-30m in altitude in New Zealand (Wardle y Coleman, 1992) and even 100m in northern Europe (L., 2002), or in Alps, where resampling of vegetation in some areas has shown a significant increase in the number of shrub and herbaceous species in alpine ecosystems (Grabherr *et al.*, 1994). As a result of the former, it is expected that such fast and persistent warming will reduce biodiversity due to biologic restrictions in terms of dispersal (vertical shift), or narrow tolerance range to global warming (Theurillat y Guisan, 2001).

To assess the potential trends of climatic warming, two important tools have been developed: the Special Report on Emission Scenarios and the Atmospheric and Oceanic Global Circulation Models (SRES scenarios and AOGCMs

hereafter). The IPCC<sup>1</sup> has designed four story lines (A1, A2, B1, and B2) that represent different economic and industrial trends over 21<sup>th</sup> century. The story lines provide data to build the different SRES scenarios, which constitute a geographical interpretation of greenhouse-gas emissions. The AOGCMs, which are physics-based models designed to study climate behaviour, have succeeded in accurately replicating the climate during the two last millennia (Mann, 2007), and are currently used to simulate climate change throughout the 21<sup>th</sup> century, using the SRES scenarios as inputs (IPCC, 2001). However, the spatial resolution of the AOGCMs does not reproduce the dependence between air temperature on the surface and terrain elevation (Giorgi *et al.*, 1997). Therefore, these models prove inadequate to generate simulations that are useful at small scales. Nevertheless, the recently developed downscaling methods of AOGCMs enable simulations to be adapted at higher resolutions, making these techniques useful to determine impact at both regional and local scale (Murphy, 2000).

Species-distribution models and climatic simulations should be combined to study the Global Warming effects on the geographical distribution of the biodiversity. The evaluation of extinction risk (Araújo *et al.*, 2006), the designing of migration corridors (Williams *et al.*, 2005), or the testing of the effectiveness of reserve networks (Araújo *et al.*, 2004) are some of the applications for species distribution models in the context of climate warming.

However, these kinds of studies have been usually applied in extensive areas (Fitzpatrick *et al.*, 2008; Benito-Garzón *et al.*, 2008), using coarse resolutions, which are not useful to be applied at small scales in the intricate topography of high-mountains. Downscaled climatic simulations are needed to appraise climatic-warming effects on mountain habitats, but high quality data are still scarce.

Until the work of Nogués-Bravo y Araújo (2008), no detailed prognosis was available on the potential trends of temperature in the mountains of the Mediterranean Basin. The importance to undertake these studies lies in the fact

---

<sup>1</sup> Intergovernmental Panel on Climate Change: <http://www.ipcc.ch>

that Mediterranean-type ecosystems have high biological diversity (Médail y Quèzel, 1999) and mediterranean high mountains are one of the most vulnerable to global warming (Grabherr *et al.*, 1994; Nogués-Bravo y Araújo, 2008). To date, no high-resolution simulation is available to analyse potential changes in the distribution of plant biodiversity in the mountains of the Mediterranean Basin within the context of global warming. In this sense, we have selected Sierra Nevada (SE Iberian Peninsula) as a case study in order to achieve a more precise perspective on the responses of the vegetation under different climate warming scenarios.

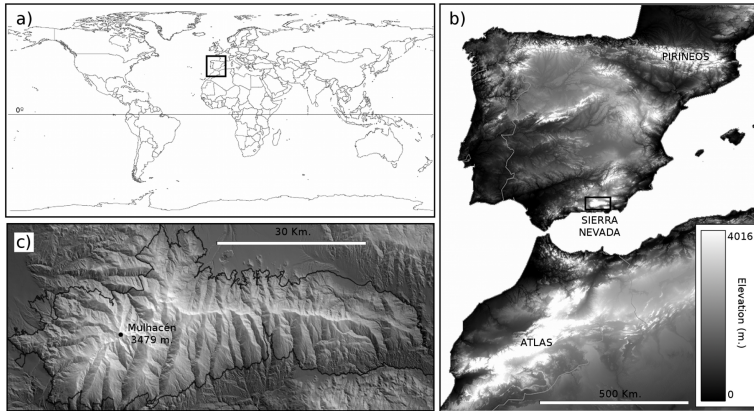
The present work seeks to help develop high-resolution tools to evaluate potential changes facing biodiversity in mountain systems due to climate warming and estimate qualitatively and quantitatively the potential impact of such changes in a key Mediterranean mountain.

## 5.2 MATERIALS AND METHODS

### *Study area*

Sierra Nevada is a high-mountain (reaching 3482 m a.s.l.) in SE Iberian Peninsula (37 °N, 3 °W). With a surface of 2,100 km<sup>2</sup>, it has a complex orography and soil composition. The average annual rainfall is highly irregular, with values ranging between 350 mm and 1,200 mm per year, depending mostly on altitude. The average temperatures are below 0°C during winter with a snow cover that can persist up to eight months in the highest areas (occasionally up to ten months in small, sheltered areas). Sierra Nevada is an isolated high-mountain, more than 700 kilometers from other comparable mountain ranges, such as the Pyrenees (N Iberian Peninsula) in the north or the High and Middle Atlas (N Morocco) in the south. This mountain marks the southernmost limit of the influence of the Quaternary glaciations in Europe, when it was covered with glaciers only in areas above 2500 m, while large areas of it remained free of permanent ice. All this helped make Sierra Nevada a refuge for many plant species during glacial ages (Blanca *et al.*, 1998) as well as

for isolated populations that have evolved under particular conditions (e.g. soil type or isolated summit areas). This has encouraged speciation, resulting in a high number (c. 80) of endemic species (Lorite *et al.*, 2007), and 123 species being threatened by different factors (Blanca *et al.*, 1998). The study area covers about 1723 km<sup>2</sup> (see Figura 1).



**Figure 1: Geographical context of Sierra Nevada range.** a) global position; b) location in the Iberian Peninsula, showing Atlas in the south, and Pyrenees in the north (both with more than 2,500 m a.s.l.); c) orography and limits of study area in Sierra Nevada mountain.

### *Vegetation types*

The four main vegetation types of the mountain range were selected to simulate potential shift of suitable habitat: Holm-oak, Pyrenean-oak, Juniper-genista thickets, and perennial psicroxerophylous-Mediterranean Alpine pastures. All were selected according to criteria such as: ecological importance, singularity, occupation area, appearance in other Mediterranean and European Alpine areas (Lorite *et al.*, 2007). Below, details are provided on selected vegetation types (Lorite *et al.*, 2008): **i)** Holm/oak communities, dominated by the tree species *Quercus ilex* subsp. *ballota*, have a widespread potential distribution in flat, low and medium mountain areas, reaching 2,000 m a.s.l., with about 8,800ha. **ii)** Pyrenean/oak

communities, formed by the marcescent tree species *Quercus pyrenaica*, reaches its southernmost European limit in Sierra Nevada, ranging between 1,100-2,000 m a.s.l. with an occupation area of about 2000 ha. **iii**) Juniper-genista thicket patches, composed of: *Juniperus communis*, *Juniperus sabina*, *Genista versicolor*, and *Cytisus galianoi* occupy some 29,000 hectares in an altitudinal belt between 1,800 and 3,100 m. **iv**) Alpine perennial psicroxerophyllous pastures cover about 2,900 hectares, over 2,700-3,000 m a.s.l. (depending on sun exposure). This perennial pastures contain most of the endemic species as well as Arctic-Alpine disjunctions of Sierra Nevada (Blanca *et al.*, 1998). The main species are: *Festuca clementei*, *Festuca pseudoeskia* and *Erigeron frigidus*. The selected plant communities occupy different altitudinal belts or exposures and represent all the potential vegetation types in this mountain.

#### *Climatic-warming scenarios and modelling approach*

Current climate records (1990-2005) and future downscaled simulations (2011-2100) were provided by the FIC<sup>2</sup> and the AEMET<sup>3</sup>. The FIC has developed the *FIC Analogues Method* (Brunet *et al.*, 2007), a statistical downscaling method of AOGCMs that simulates future diary records of minimum and maximum temperatures (and rainfall, but we reject them in the early stages of this research because it was not possible to reach an accurate rainfall spatial model for the mountain range) for every weather station of the Spanish meteorological network. The simulations provided by the FIC combines two SRES scenarios (**A2** and **B2**) and two AOGCMs: ECHAM4 ((Roeckner *et al.*, 1996) and CGCM2 (Flato y Boer, 2001), resulting in four climate warming scenarios (**ECHAM4-A2**, **ECHAM4-B2**, **CGCM2-A2**, and **CGCM2-B2**) for the period 2011-2100.

To build high-spatial-resolution temperature maps for each climate warming scenario, we processed current observational and future simulated temperature records (from

<sup>2</sup> Foundation for Climate Research: <http://www.ficlima.org>

<sup>3</sup> State Agency of Meteorology, Government of Spain: [http://www.aemet.es/es/elclima/cambio\\_climat/escenarios](http://www.aemet.es/es/elclima/cambio_climat/escenarios)

forty three thermometric stations 50 kilometers around Sierra Nevada, with an elevational range between 300 and 2,860 meters) in the R software (R-Team, 2009) to compute yearly mean temperatures for six 15-year intervals (2011-2025, 2026-2040, 2041-2055, 2056-2070, 2071-2085, and 2086-2100). To convert point temperature records into continuous digital maps, we applied a multiple-regression method with residual correction (Ninyerola *et al.*, 2000). As temperature predictors, we derived topographic and geographic variables (elevation, latitude, longitude, E-W and N-S gradients, and distance to the coastline) from a high-resolution elevation (10 m) model in the GRASS GIS environment (Team, 2009). We analysed the resulting temperature maps to measure the future temperature trends expected in Sierra Nevada by the end of the century. To represent the microtopographical complexity of the range and use them as plant distribution predictors at small scales, we derived three topographical variables from the digital elevation model: wetness index, yearly mean solar radiation and slope.

Maxent (Phillips *et al.*, 2006; Phillips y Dudík, 2008) is an algorithm designed to model the geographic distributions of species. Based on artificial intelligence and the maximum-entropy condition (Jaynes, 1957), it provides accurate results compared with a broad range of species-distribution models (Elith *et al.*, 2006). Maxent computes the species-distribution models, finding the maximum-entropy equation (close to uniform) that explains the relationships between species-presence records and the values of the environmental variables used as predictors. The resulting equation is geographically projected as an habitat-suitability map if the output is set to *cumulative* (Phillips y Dudík, 2008). The model should be transferred into space or time (for example, to future simulated conditions); this approach has been successfully evaluated (Hijmans y Graham, 2006) and applied to plant species (Fitzpatrick *et al.*, 2008).

The current habitat-suitability models (period 1990-2005) were calibrated using as environmental variables the mean annual temperature (1990-2005 period), and the three topographic factors. The amount of presence records used to calibrate the models were: Holm oak, 3576 points; Pyrenean



oak, 698 points; Juniper-Genista thickets, 9143 points; Alpine psicroxerophyllous-Mediterranean perennial pastures, 952 points. Maxent settings were: iterations, 2000; convergence threshold, 0.00001; background points, 10000. The calibrated models were then projected onto the four scenarios, using the corresponding mean annual temperature simulation for each time interval, to compound the simulated time series for every scenario.

Model evaluation was performed by the AUC test implemented in MaxEnt, using random data-splitting (25% of presence records as the evaluation dataset). To simplify the spatial analysis, we converted the habitat-suitability models from continuous (the cumulative MaxEnt output, in the range of habitat-suitability from 0 to 100) to boolean (1 and 0 as suitable and non-suitable habitat respectively), using a specific threshold criteria for every simulation: for all the evaluation presence records, mean and standard deviation of habitat-suitability values were computed, and the threshold was established by the expression [mean – standard deviation]. All values above the threshold were recoded to 1, as areas of suitable habitat.

The four time series were analysed from two viewpoints:

1. Measurement of potential area and altitudinal shift: Based in the boolean simulations, the area and average elevation of the suitable habitat for every vegetation type were computed.
2. Analysis of habitat-suitability change. In this section, we introduce a simple term, the *Differential Suitability* (**DS** hereafter), as the difference between habitat-suitability values upon different time intervals. **DS** Values are in the range [100, -100]; positive and negative values denote habitat-suitability gain and loss, respectively.

We use the **DS** concept to:

- Locate potential advance and receding fronts, by assembling (by model averaging) the habitat-suitability models corresponding to scenarios based in CGCM2 and computing **DS** between periods 1990-2005 and

2011-2025 (first and second components of the time series). As advance fronts were considered to be those areas with positive DS (habitat-suitability for the vegetation type is potentially improving), whereas areas with negative DS were considered to be receding fronts. Additionally, the DS values were sampled and analysed for all presence records, to evaluate and compare the trends of the known localities for each vegetation type.

- Analyse potential replacement patterns between vegetation types. This analysis is based on this idea: taking into consideration the vegetation types **a** and **b**, and the corresponding habitat-suitability models, if in a locality with presence of **a** the suitability for **a** diminishes, simultaneously to the increase in suitability for **b**, then the ecological conditions are favouring **b** to replace **a**. Assembling (by model averaging) the simulations corresponding to scenarios based in CGCM2 for the period 2005-2055, the suitability gained and lost for each vegetation type was sampled for all the presence records, and the results were summarized in a flowchart representing the potential replacement pattern for the first half of the 21<sup>th</sup> century.

### 5.3 RESULTS

The analysis of the four climatic-warming scenarios show that the different combinations of AOGCM and SRES scenarios cover a broad spectrum of possibilities. The global-circulation model ECHAM4 simulates the most severe temperature increases, between +7°C of the A2 scenario and +5.6°C of the B2. The simulations based on CGCM2 gave milder results, between +4.4°C of the A2 scenario and +2.3°C of the B2. An average of the four scenarios gives an increase of +4.8°C at the end of the 21st century (see Table 5.3).

The habitat-suitability models of the four plant formations (period 1990-2005) show robust fits according to their AUC values: Holm oak forest = 0.937; Pyrenean oak forest = 0.963; juniper-genista thicket = 0.973; pasture = 0.998. Figure

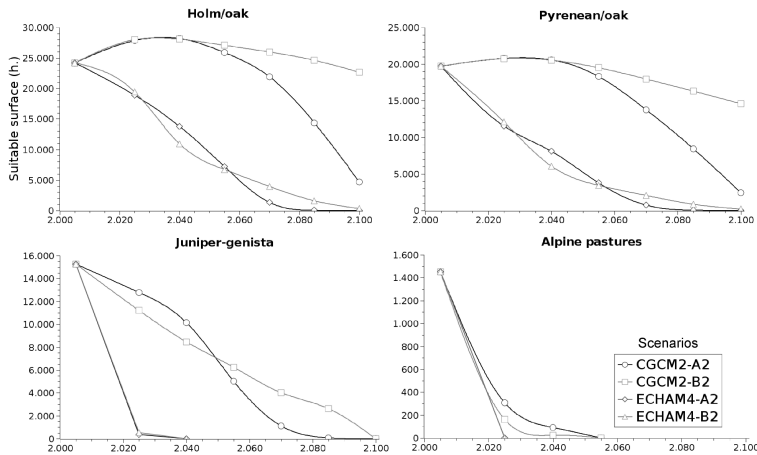
	A2 (°C)	B2 (°C)	Average (°C)
ECHAM4	7.0	5.6	6.3
CGCM2	4.4	2.3	3.3
Average	5.7	3.9	4.8

**Table 1: Potential increase of temperature in Sierra Nevada (2000-2100):** Comparison between combinations of two SRES scenarios (A2 and B2) and two AOCCMs (ECHAM4 and CGCM2).

2 presents the temporal series of the potential area, with sharp differences between the formations at lower altitudes (Holm oak forest and Pyrenean oak forests), and those at higher altitudes (juniper-genista thickets and pastures). The simulations based on the CGCM2 scenarios indicate a mean altitudinal shift of the suitable-habitat for all the formations of  $7.81 \pm 2.67$  m year<sup>-1</sup> (mean and standard deviation), with the maximum of the Holm oak forest (11.73 m year<sup>-1</sup>), and the minimum for the pasture (5.3 m year<sup>-1</sup>). The simulations based on ECHAM4 indicate far higher potential ascents, with a mean among the formations of  $15.33 \pm 2.47$  m year<sup>-1</sup>, in which the Holm oak forest also registers a maximum (17.93 m year<sup>-1</sup>) and the pasture the minimum (12.17 m year<sup>-1</sup>). The average ascent for all the formations and scenarios was  $11.57 \pm 4.61$  m year<sup>-1</sup>.

The DS analysis (see example in Figure 3) indicates that a large part of the presence records of the vegetation types appear in areas with a negative DS. The vegetation type with the most presence records in areas of negative DS was the pasture, in which 100% were situated at suitability-loss levels of -15 to -100. This was followed by the juniper-genista thicket, with 74% of the records between the levels -1 and -100, the Holm oak forest with 64% in the same range, and the Pyrenean oak forest with 62% of the records in negative values.

The analysis of the potential replacement among plant formations (Figure 4) indicates that the areas currently occupied by Pyrenean forests are becoming suitable for the Holm oak forest. The same circumstance is repeated in some areas occupied by the juniper-genista thickets, which are increas-



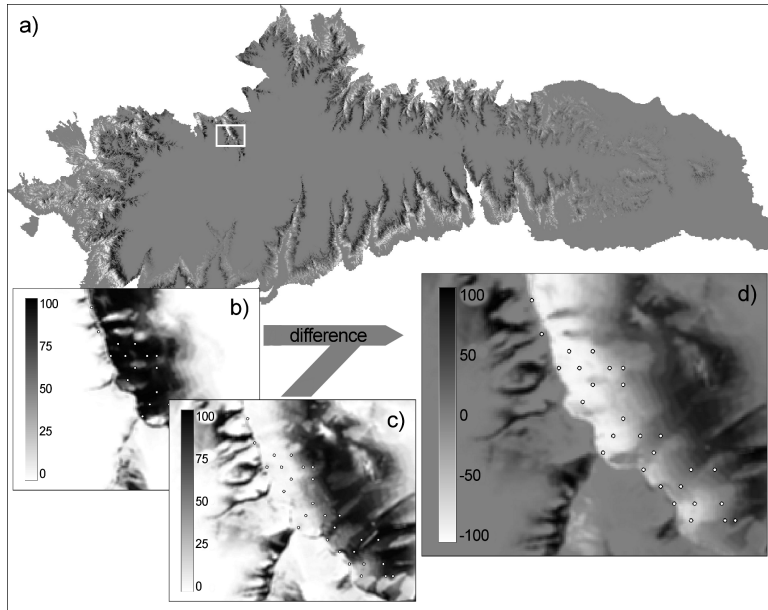
**Figure 2: Temporal analysis of the suitable-habitat surface.** Temporal evolution of the potential area of the different plant formations in four scenarios of climatic change.

ing in suitability for the Pyrenean oak forests. The territory currently occupied by the pasture is becoming suitable for the juniper-genista thicket, implying potential replacement by the latter.

## 5.4 DISCUSSION

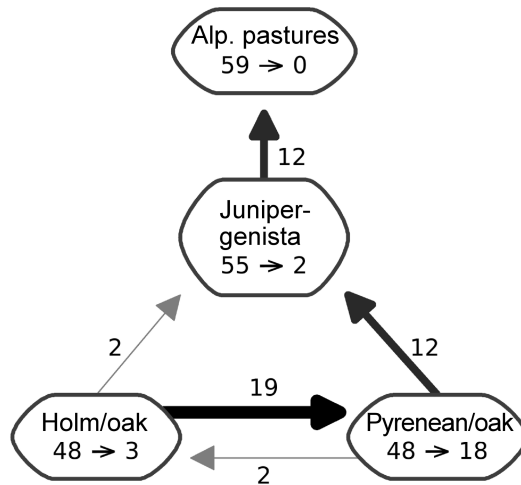
**Hulme y Sheard (1999)**, using an assembly of 7 AOGCMs, quantified the temperature increase on a cell of 300 x 200 km (which contained Sierra Nevada) at 5.8°C for the A2 scenario and 2.9°C for B2. More recently, **Nogués-Bravo y Araújo (2008)**, in a similar work but focused in Mediterranean mountains, quantify the possible range of temperature increase at between +1.6°C and +8.3°C. A comparison of these results with ours indicates that maps of future temperatures for the Sierra Nevada generated for the present work from the FIC downscaled simulations are consistent with current knowledge on the potential trends of climate change in the southern Iberian Peninsula.

A detailed analysis of the projected habitat-suitability models leads us to the conclusion that even the most optimistic simulations predict a potential loss of suitable habitat,



**Figure 3: Detail of models of differential suitability.** **a)** Model of differential suitability for Holm oak forest, representing potential fronts of advance (black) and recession (white) of the plant formation for the period 2005-2025. **b)** Suitability of the habitat for the Holm oak forest corresponding to the period 1990-2005. The dark tones indicate the appropriate habitat. The dots represent presence records of the plant formation. The area represented in **a)** with a blank window. **c)** Suitability of the habitat for the Holm oak forest corresponding to the period 2011-2025; **d)** detail of the differential-suitability model represented in **a)**. This map is the result of calculating the difference between the suitability maps **a)** and **b)**.

which may become total loss for high-mountain formations such as juniper-genista thickets and pastures. These results are consistent with those found in other similar studies (although applying coarser resolutions) in the Alps (Thuiller *et al.*, 2005b) and Australia (Fitzpatrick *et al.*, 2008). However, in the present work, we found a rapid decline in the potential area and a significantly high vertical migration rate, consistent with the results of studies such as that of Grabherr *et al.* (Grabherr *et al.*, 1994), according to which over the 20<sup>th</sup> century, 9 species of the Alps have migrated altitudinally between 70 and 360 m.



**Figure 4: General pattern of potential replacement among vegetation types.** Graphic summary of the calculations made on averages of the simulations **CGCM2-A2** and **CGCM2-B2** for the period 2005-2055. The numbers within each plant formation indicate the decline in suitability for that formation (average of the values of all the presence data of the formation). The arrows indicate that transference of suitability among plant formations. For example, according to the diagram, it is interpreted that in the areas of presence of the Pyrenean oak forest, the mean suitability increases by 19 units for the Holm oak forest.

In this scenario of such rapid change, the plant formations of Sierra Nevada have two options: *in situ* adaptation or migration to favourable conditions (Lamont y Connell, 1996). In this context, the case of the Holm oak forest and the Pyrenean oak forest should be analysed with attention, as both formations appear in zones of North Africa such as the Rif or the High and Middle Atlas (Morocco), under warmer conditions (8.6°C mean annual temperature at 2000 m of altitude in Sierra Nevada vs. 11.6°C in the High Atlas, according to WORLDCLIM<sup>4</sup> database; Hijmans *et al.* (2005), leading us to assume that both have wider ecological ranges than those found in Sierra Nevada. This circumstance, together with the long life cycles that characterize the key species of these formations, may buffer the negative effect of

<sup>4</sup> <http://www.worldclim.org/>

temperature change in order to maintain stable populations over long time periods (Ghalambor *et al.*, 2007). However, a question that has not been completely elucidated is key for interpreting the simulation results that is, the possibility of the species to adapt to rapid environmental changes (Pearman *et al.*, 2008). According to several works (Davis y Shaw, 2001; Prinzing *et al.*, 2001), the fossil record of some plant species indicate that they have maintained long periods of stability in their ecological requirements, while others have shown very rapid changes. The behaviour of the key species of these formations will determine the survival of many accompanying species, given that the former act as nurse plants for the latter, creating micro-niches appropriate for their regeneration (Castro *et al.*, 2004).

When there is no possibility of quick *in situ* adaptation, the alternative is to migrate to favourable conditions. In Sierra Nevada, due to its island-like ecological behaviour and topographical configuration, only upwards shift is possible. The analysis of our simulations indicates that vertical shift of plant populations should be very swift so that they can maintain themselves within an optimal habitat-suitability range. This circumstance presents some problems: first, intrinsic to the species, given that many species might encounter limitations when migrating upwards because their seed-dispersal abilities are limited, or due to natural or anthropogenic barriers (Theurillat y Guisan, 2001). Afterwards, migration will depend on the potential for establishment (survival of seedlings and juveniles), this being limited by factors such as herbivory, which lowers the recruitment rate. For example, on Sierra Nevada, the advance of the Holm oak formations into zones of thickets would be seriously limited by herbivory (Gómez *et al.*, 2003), while a similar circumstance occurs for the regeneration of *Juniperus communis* (García *et al.*, 1999). The extrinsic problems are a consequence of the topographical configuration. With the rise in altitude, the area of available habitat diminishes, the topography becomes more hostile (steeper slopes), and the soil loses the power to sustain shrub and tree species.

Simulations of climatic change are vulnerable to what is called the *cascade of uncertainties*, some inherited from

the AOGCMs and the SRES scenarios, while others are related to the concrete technique of downscaling (Mitchell y Hulme, 1999). To confront this uncertainty, Mitchell y Hulme (1999) recommend the use of different scenarios of climatic change and the application of assembly techniques, as used in the present work. On the other hand, the habitat-suitability models have a number of known problems, such as not considering the biotic interactions of the potential of species to adapt *in situ* to new conditions (Broennimann *et al.*, 2006). These questions lead to the establishment of spurious relationships between climate and species presence, and provokes errors on extrapolating these relationships over time (Fitzpatrick *et al.*, 2008). The distribution models also ignore the mechanisms that affect species demography, although it is assumed that the relationship between species presence and environmental conditions is a good substitute for demographic processes (Austin, 2007). In addition, the projections of the distribution models in time cannot be evaluated with empirical data (Araújo *et al.*, 2005a), and therefore the uncertainty associated with their results is difficult to estimate. An approach to this problem is the assembly of distribution models (Araújo y New, 2007), but it is a costly computational solution when the volume of the work is large, and it can give results that are difficult to interpret when there is little consensus between the different models, something that often occurs when using algorithms based on different logic systems.

Given these limitations, the simulations made in the present work cannot be understood strictly as predictions of future distributions, because there should be a lag between suitable-habitat shift and plant populations shift. But our results are of great usefulness to offer general responses about the potential effects of climatic change on the vegetation types studied in Sierra Nevada, and the results can be extrapolated to other formations and to other mountains of the Mediterranean basin.

Even in the most optimistic scenarios of climatic warming, Sierra Nevada runs a serious risk of increased climatic stress for the studied vegetation types. This risk, is greater in the high-mountain formations, which harbour the greatest



part of the communities and the endemic species of Sierra Nevada (Lorite *et al.*, 2007). Our options to act at the local level on phenomena such as global warming are scant, but these high-resolution simulations provide valuable information that can help in the design of management strategies to face potential changes due to climatic change. In particular, the DS method make it possible to identify the advancing and receding fronts of the plant formations with a high degree of detail. These areas which have a great probability of undergoing the first effects of climatic warming are the most appropriate ones for the establishment of monitoring points (Thomas y Franco, 2006), which would be useful to test the impact predicted by the models for the different formations. The DS method also make it possible to locate the areas in which to apply measures that will facilitate the transition of one formation or another, as for example the conditioning of pine forests for reforestation for future establishment of Holm oak forests or even the use of abandoned agricultural land in terraces (Douglas *et al.*, 1996).

The real challenge lies in the preservation of biodiversity of the mountain summits, since the formations living there lack areas for expansion and will be subjected to great pressure, both by the degradation of the conditions appropriate to each formation as well as by the arrival of new competitive species from lower altitudes. In this context, it is a priority to analyse the development of the changes in these formations, as well as to investigate the possibilities of conservation in surroundings that are becoming hostile.

### *Aknowledgements*

The present work was financed by a research project of the Council of Innovation, Science and Enterprise of the Government of Andalusia, (Spain), called "Conservation of endemic and threatened flora in fragile habitats: sub-deserts and high mountains of Andalucía" (P05-RNM-01067). Also, we wish to show our appreciation to the research staff of the Foundation for Climate Research and the State Agency of Meteorology of the Government of Spain for preparing

the data required to generate the regionalized scenarios of climatic change and making them publicly available.



## FLUJO DE TRABAJO AUTOMATIZADO PARA LA SIMULACIÓN DE CAMBIOS EN LA DISTRIBUCIÓN GEOGRÁFICA DE LAS PLANTA

---

### 6.1 INTRODUCCIÓN

Los MDEs tienen una serie de aplicaciones prácticas que justifican sobradamente su interés: cartografía de especies en zonas de difícil acceso, localización de poblaciones de especies en peligro, análisis de fragmentación de hábitat, selección de reservas, estudios biogeográficos, y más recientemente, la simulación de los efectos potenciales del Cambio del Clima en la distribución geográfica de las especies.

Este tipo de simulaciones son computacionalmente muy complejas, y requieren gran cantidad de pasos. Aunque en estos últimos años el software de modelado ha evolucionado mucho y existen algunos paquetes de gran calidad como MaxEnt, OpenModeller o Biomapper, todos presentan alguna carencia que impide llevar a cabo el proceso completo de simulación de modo automatizado.

Entre las carencias más importantes está la omisión sistemática de una fase de gran importancia para garantizar la calidad de los resultados: la preparación de los datos de presencia de las especies, dejando esta completamente en manos del investigador, cuando se trata de un trabajo laborioso y apenas normalizado. Una segunda limitación consiste en la aplicación de un único algoritmo para generar las simulaciones, cuando actualmente empieza a aceptarse que la mejor aproximación es generar varios modelos de la misma especie con diferentes algoritmos y ensamblarlos para estudiar la tendencia central del conjunto (Araújo *et al.*, 2005b; Araújo y New, 2007; Marmion *et al.*, 2009).

Un MDE, calibrado con las condiciones climáticas actuales, puede proyectarse hacia el futuro (o el pasado), si se dispone de simulaciones climáticas para el periodo de proyección.

*En el Capítulo 2, apartado Aplicaciones (página 56) puedes ampliar información sobre aplicaciones de los MDE.*

Algunos de los programas de modelado directamente carecen de esta capacidad de proyección (Biomapper), pero otros como OpenModeller y MaxEnt si disponen de esta característica, pero no proporcionan herramientas para organizar y analizar los resultados de las proyecciones, o ensamblar los resultados de distintos algoritmos.

Cuando un proyecto de modelado implica a muchas especies, el esfuerzo necesario se multiplica, y cuando los MDE se van a proyectar sobre distintos escenarios de Cambio Climático, se vuelve directamente inabordable con los medios que un usuario normal de estos programas tiene a su disposición. En estas circunstancias la mejor alternativa ha sido recurrir a combinaciones de software de modelado, SIG y un entorno de programación (para un ejemplo, ver [Benito-Garzón 2006](#)), pero son necesarios ciertos conocimientos en lenguajes de programación y un tiempo considerable para poner a punto el sistema.

Recientemente las posibilidades técnicas para enfrentarnos a estas situaciones han mejorado gracias al desarrollo de los programas de diseño y ejecución de flujos de trabajo, que utilizan un lenguaje visual basado en diagramas de flujo para sustituir líneas de código, más asequible para usuarios sin experiencia. Como consecuencia de la emergencia de esta nueva tecnología, ahora es posible combinar programas SIG y programas de modelado mediante flujos de trabajo, para ejecutar, de modo automático, todas las tareas requeridas por un proyecto de modelado de cierta envergadura.

## 6.2 OBJETIVO

En este capítulo mi objetivo es describir un flujo de trabajo automatizado diseñado para simular cambios en la distribución de plantas a consecuencia del Cambio Climático, aplicando una técnica de ensamblado de modelos obtenidos mediante distintos algoritmos. Una herramienta de estas características ofrece al investigador la oportunidad de obviar todo el tedioso procesamiento de los datos, para dedicar su tiempo a lo que es realmente importante: la interpretación de los resultados.

### 6.3 ANTECEDENTES

Probablemente en nuestro país ha sido Marta Benito-Garzón (Benito-Garzón, 2006) la pionera en la proyección de modelos de distribución de plantas sobre distintos escenarios de cambio climático. En su tesis doctoral ha ofrecido una imagen retrospectiva y prospectiva sobre la distribución espacial de las especies arbóreas más importantes de la Península Ibérica. Su desarrollo se basa en software libre, utilizando GRASS para análisis SIG, R para generar los modelos de distribución (utilizando redes neuronales, árboles de regresión y el algoritmo Random Forest como algoritmos de modelado), y como cemento, para facilitar el flujo de datos entre los distintos paquetes y la automatización de tareas, distintos lenguajes como bash o perl.

Otro antecedente más directamente relacionado con este trabajo puede encontrarse en Pennington *et al.* (2007), en el que los autores presentan un flujo de trabajo diseñado sobre el software Kepler para ejecutar un proyectos de modelado de gran envergadura. Los autores utilizan un único algoritmo de modelado (GARP), aprovechando la información distribuida de la red EcoGrid (planeada para análisis ecológicos a gran escala en los EEUU) para generar sus pronósticos. Gran parte de este grupo de trabajo opera en el Centro de Supercomputación de San Diego<sup>1</sup>, por lo que resultaría ridículo comparar su trabajo con el expuesto en este capítulo, teniendo en cuenta los recursos humanos y técnicos de los que disponen.

Durante el desarrollo de este trabajo se ha liberado una nueva versión de BIOMOD (Thuiller, 2003; Thuiller *et al.*, 2009), una librería de R diseñada para modelado con múltiples algoritmos que ofrece un gran repertorio técnico para generar proyecciones sobre distintos escenarios de cambio climático y análisis de los resultados. Su uso ha estado bastante restringido hasta ahora, pero está adquiriendo una gran aceptación en la comunidad científica, y cada vez son más los trabajos que la utilizan. Los autores de esta librería son referentes mundiales en la disciplina MDE.

---

<sup>1</sup> <http://www.sdsc.edu/>

## 6.4 MATERIAL Y MÉTODOS

### 6.4.1 Herramientas informáticas

Las distintas etapas de procesamiento que exige el proyecto solicitan programas con distintas capacidades, que puedan integrarse, y cuyas funciones principales se pueden dividir, en cinco grupos: control de flujo, procesamiento SIG, modelado de distribución, cálculo numérico y composición de gráficos. Hay dos cuestiones claves a tener en cuenta: los programas deben ser interoperables (compartir información sin restricciones de formato), de código abierto (para poder hacer modificaciones según los requerimientos del proyecto), y coste *cerro*. Por suerte, la comunidad **Open Source** proporciona a los investigadores herramientas de gran calidad, como las que trataré a continuación, que permiten llevar a cabo proyectos de estas características.

#### *Kepler*

Kepler es un proyecto colaborativo, de código abierto (licencia BSD<sup>2</sup>), que pretende proporcionar un **entorno de modelado y resolución de problemas**. más concretamente, se trata de un sistema para diseñar y ejecutar flujos de trabajo científico, utilizando una representación visual de los procesos que implican.

Kepler se basa en un paradigma de modelado orientado a actores. Cada uno de los componentes que ejecuta una función dentro de un flujo de trabajo (por ejemplo, una entrada de datos) es un actor (ver Figura 1). El actor es la mínima unidad computacional de Kepler. Tiene puertos de entrada a través de los cuales se alimenta de datos, un núcleo de procesamiento que opera con ellos, y puertos de salida, que transmite los datos transformados a un nuevo actor, o directamente al usuario.

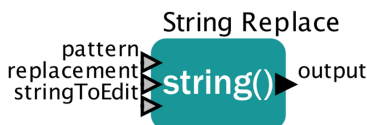
Kepler incluye actores con capacidades matemáticas, estadísticas, procesamiento de señales, entrada, manipulación y salida de datos, conexiones con otros programas como R, MATLAB o GRASS. Los actores transmiten información de

---

<sup>2</sup> [http://en.wikipedia.org/wiki/BSD\\_licenses](http://en.wikipedia.org/wiki/BSD_licenses)

unos a otros mediante canales y relaciones (ramificaciones de los canales). Distintos actores pueden combinarse en un único actor compuesto, que adoptará las funcionalidades que le confieren los distintos actores que lo forman.

El **Director** controla la ejecución del flujo de trabajo, determinando en que modo los actores se comunican entre sí, cuántas veces se ejecuta el flujo (iteraciones), y si se ejecutan las operaciones secuencialmente o en paralelo.



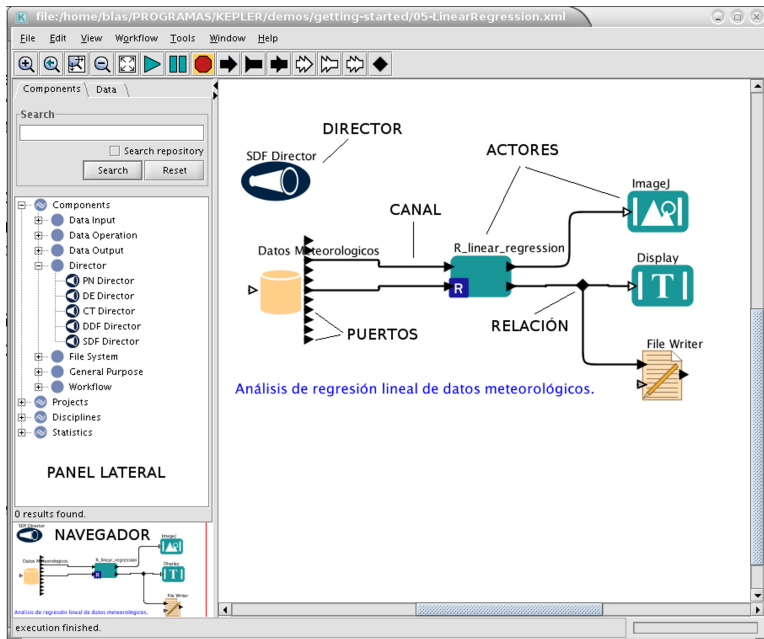
**Figura 1:** Actor típico de Kepler con sus puertos de entrada y salida. Se trata de un actor *String Replace* que reemplaza una cadena de caracteres por otra en un flujo de texto. La cadena a reemplazar se introduce a través del puerto *pattern*, el reemplazo a través del puerto *replacement*, el texto completo a través *stringToEdit*, y el texto resultante se devuelve a través de *output*.

Una función interesante de Kepler es la posibilidad de diseñar un *Actor Compuesto*, que permite anidar flujos de trabajo dentro de otros flujos de trabajo de nivel superior. Cada uno de estos niveles puede tener su propio Director, dependiendo de los requerimientos computacionales del nivel. En este proyecto se aprovecha esta característica para anidar procesos que ocurren un cierto número de veces dentro de un mismo flujo.

Un flujo de trabajo de Kepler (Figura 2) está formado por un Director, unos parámetros específicos del flujo, y una red de actores unidos entre sí por canales y relaciones (bifurcaciones de los canales).

En este proyecto Kepler ha sido la herramienta fundamental para diseñar y ejecutar el flujo de trabajo, llevar a cabo determinadas tareas dentro del mismo (escritura de ficheros, organización de datos, control de iteraciones), y llamadas a ejecución para los restantes programas.





**Figura 2:** Un flujo de trabajo en Kepler. El actor *Datos Meteorológicos* da entrada a una base de datos meteorológica. Cada columna de la base de datos es un puerto de salida en el actor. Los canales llevan la información hasta un actor que utiliza el lenguaje estadístico R para generar un modelo lineal de los datos recibidos. Los datos resultantes se escriben a un fichero de texto (actor *File Writer*), se muestran en una ventana en pantalla (actor *Display*) y se representan gráficamente en un diagrama de dispersión (actor *ImageJ*).

## GRASS

GRASS<sup>3</sup> es un programa especializado en el almacenamiento, análisis y explotación de información geográfica. Desarrollado inicialmente por el laboratorio CERL<sup>4</sup> desde 1982, y liberado con licencia GPL<sup>5</sup> en 1999, GRASS es actualmente un proyecto oficial de OSGeo<sup>6</sup>. GRASS tiene una gran capacidad de procesamiento raster y vectorial, y herramientas para la visualización de datos geográficos en dos y tres di-

3 Geographic Resources Analysis Support System: <http://grass.itc.it>

4 U.S. Army Construction Engineering Research Laboratories

5 [http://en.wikipedia.org/wiki/GNU\\_General\\_Public\\_License](http://en.wikipedia.org/wiki/GNU_General_Public_License)

6 Open Source Geospatial Foundation: <http://www.osgeo.org/>

mensiones. Pero sin duda, la característica más interesante, que determina su adopción como plataforma SIG para este proyecto, es su potente lenguaje de programación, que se integra perfectamente con el lenguaje bash en sistemas unix/linux. Esta integración permite ejecutar tareas de modo automático fácilmente a través de *batch-jobs* (secuencias de órdenes), que permiten automatizar cualquier tarea. Otras características interesantes son su gran estabilidad y velocidad trabajando con raster de gran tamaño, y la diversidad de formatos de importación y exportación mediante las librerías GDAL-OGR<sup>7</sup>.

### *OpenModeller*

OpenModeller<sup>8</sup> es un entorno amigable de modelado de distribución de especies, actualmente desarrollado por la entidad brasileña CRIA<sup>9</sup> con la colaboración de otras instituciones. Está liberado bajo licencia GPL, y es uno de los programas de su género más desarrollado, ya que contempla gran parte de los procesos necesarios para generar modelos de distribución. Permite la utilización de distintos algoritmos para generar los modelos (unos 10, dependiendo de la versión), la evaluación estadística de los resultados, y la exportación de los modelos a distintos formatos. Desde el principio de su desarrollo está pensado para ser multiplataforma, y está dotado de herramientas que permiten el acceso al programa a través de la línea de comandos, facilitando su explotación mediante scripts.

### *MaxEnt*

MaxEnt<sup>10</sup> es un software gratuito de código cerrado diseñado para generar modelos de distribución utilizando un algoritmo que combina el aprendizaje automático con el principio de máxima entropía.

*MaxEnt es el único programa sin licencia libre de este proyecto, pero sí es de libre uso para investigación y enseñanza.*

---

7 <http://www.gdal.org/>

8 <http://openmodeller.sourceforge.net/>

9 Centro de Referência em Informação Ambiental: <http://www.cria.org.br/>

10 <http://www.cs.princeton.edu/~schapire/maxent/>

Se utiliza para este proyecto, adicionalmente a OpenModeler, por sus buenos resultados, y el fácil acceso a las funcionalidades del programa a través de la línea de comandos, que permite su ejecución mediante scripts, o directamente desde un actor modificado de Kepler.

### *Octave*

Octave<sup>11</sup> es un lenguaje de cálculo numérico matricial libre y gratuito (licencia GPL). Es muy potente, y permite un rápido desarrollo de algoritmos de cálculo gracias a un lenguaje de programación sencillo y bien documentado. Se accede a sus funciones a través de la línea de comandos o de scripts.

### *R*

El lenguaje de cálculo estadístico R<sup>12</sup> es con seguridad el programa libre (licencia GPL) más utilizado en el ámbito académico. Tiene como ventaja un lenguaje de programación muy potente, que facilita la programación de scripts.

Destaca también por la calidad de su motor gráfico y la enorme cantidad de funciones de cálculo de las que dispone. Kepler está especialmente preparado para conectar con R, y muchos actores utilizan expresiones en este lenguaje.

*También tiene algunos interfaces gráficos interesantes, como R-Commander, JGR o RKward.*

### *Bash*

Bash es un intérprete libre (licencia GPL) de órdenes en sistemas UNIX/Linux, útil para programar tareas de distinta complejidad, como creación de directorios, renombrado y edición de ficheros, bucles para la ejecución de tareas iterativas, etc.

*Bash es una auténtica navaja suiza, capaz de casi cualquier cosa en un sistema Linux.*

### *Otras herramientas*

Algunas herramientas adicionales han sido **rpl**, para sustituir cadenas de texto en gran número de ficheros; **pdftk**, para editar ficheros pdf con gráficos de R; ImageMagick, para automatizar tareas de edición de imágenes.

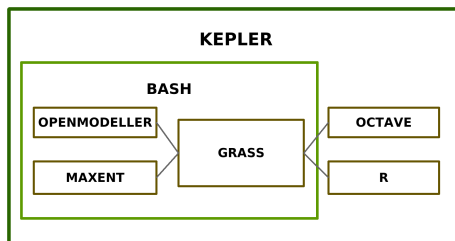
---

<sup>11</sup> <http://www.gnu.org/software/octave/>

<sup>12</sup> <http://cran.r-project.org/>

### *Integración de las herramientas*

Con la idea de automatizar el proceso de modelado, las distintas herramientas deben integrarse en lo que llamaré *arquitectura del sistema*, una jerarquía en la que unas herramientas gobiernan la ejecución de otras, y el trasvase de información entre aplicaciones es fluido y automático. La arquitectura del sistema puede observarse en la Figura 3. Según este esquema, Kepler es el que decide cuándo y qué se hace en cada momento (y es el que contiene y ejecuta el flujo de trabajo), llamando a cada aplicación cuando es necesario. Bash ejecuta todas las tareas que no pueden ser ejecutadas directamente por actores de Kepler, y controla la ejecución de OpenModeller y MaxEnt. Bash también lanza los scripts de GRASS, y ejecuta distintas tareas dentro de esos scripts. Octave y R dependen directamente de Kepler para ejecutar cálculos numéricos sobre los datos proporcionados por GRASS.



**Figura 3:** Arquitectura del sistema. Las líneas entre bloques representan flujo de datos entre aplicaciones. Kepler controla directamente a través de sus actores a Bash, Octave y R. Son órdenes de Bash las que ejecutan modelos con OpenModeller y MaxEnt, y los analizan con GRASS.

### *Datos de entrada*

Este proyecto de modelado necesita cierto tipo y gran cantidad de datos iniciales, procedentes de diversas fuentes. Parte importante de la información geográfica primaria utilizada en este trabajo la he obtenido de la Red de Información

Ambiental de la Consejería de Medio Ambiente de la Junta de Andalucía, a través de un convenio de colaboración suscrito entre esta entidad y el grupo de investigación al que pertenezco. Los datos meteorológicos y las simulaciones de Cambio Climático proceden de la Agencia Estatal de Meteorología (AEMET<sup>13</sup>) y la Fundación para la Investigación del Clima (FIC<sup>14</sup>):

### *Presencia de especies*

La información de presencia de especies vegetales procede de tres fuentes:

1. FAME<sup>15</sup> (Flora Amenazada): Contiene la localización geográfica de poblaciones de plantas catalogadas por la Ley Andaluza de Flora y Fauna. Se trata de una cartografía de poblaciones relativamente precisa (aunque no exenta de errores), generada por los técnicos de la Red Andaluza de Jardines Botánicos durante sus trabajos de campo. El formato original es vectorial, ESRI shapefile de geometría poligonal.
2. Mapa de Usos y Coberturas<sup>16</sup>: Este mapa de usos de escala 1:50000 corresponde a la situación del año 1999, y contiene gran cantidad de información sobre presencia de especies comunes, especialmente del estrato arbóreo y arbustivo.
3. Información de recolección propia: Durante varios años el grupo de investigación ha estado recogiendo geolocalizaciones de poblaciones de especies amenazadas mediante dispositivos GPS, especialmente de los ecosistemas áridos y altas montañas de Andalucía Oriental.

Estos datos los procesé manual y automáticamente (según casos) para unir en ficheros individuales la información

---

13 [http://www.aemet.es/es/elclima/cambio\\_climat/escenarios](http://www.aemet.es/es/elclima/cambio_climat/escenarios)

14 <http://www.ficlima.org/>

15 Consejería de Medio Ambiente de la Junta de Andalucía

16 Enlaces a los datos y Guía Técnica en el Anexo *Fuentes de Datos en la Red*

sobre cada especie de las distintas fuentes de datos.<sup>17</sup> Como resultados se dispone de ficheros vectoriales con polígonos de presencia de más de quinientas especies, tanto comunes como amenazadas. Sin embargo para muchas de ellas el área de presencia es escasa, por lo que no serán apropiadas para generar MDE.

### *Variables ambientales*

Para generar las variables ambientales fueron necesarios los siguientes conjuntos de datos:

- Modelo digital de elevaciones<sup>18</sup> de 10 metros de resolución espacial.
- Registros climáticos de la red de estaciones meteorológicas de la AEMET, obtenidos a través del Subsistema Clima de la Consejería de Medio Ambiente, en formato DBF, cubriendo el periodo 1971-2000.
- Simulaciones climáticas (temperatura y precipitación) del futuro según dos escenarios de cambio climático del IPCC<sup>19</sup> (A2 y B2) generadas por la Fundación para la Investigación del Clima utilizando el *método de Análogos FIC*. Están accesibles en la web de la AEMET<sup>20</sup>, y se trata de un extenso banco de datos en formato texto, con simulaciones diarias de temperatura máxima, mínima y precipitación para cada estación meteorológica del territorio español.

A partir de estos datos de partida preparé un amplio conjunto de variables topográficas y climáticas aptas para calibrar las simulaciones. En la selección de las variables tuve en cuenta la correlación espacial entre ellas, y su importancia en la distribución de plantas.

---

17 Los datos de muchas de las especies catalogadas por ley de Flora y Fauna fueron preparados por Raúl Casado Barbero, al que agradezco su esfuerzo y desinterés al ofrecérmelos.

18 Publicado por la Junta de Andalucía en 2005

19 [http://www.grida.no/publications/other/ipcc\\_sr/?src=/climate/ipcc/emission/094.htm](http://www.grida.no/publications/other/ipcc_sr/?src=/climate/ipcc/emission/094.htm)

20 [http://www.aemet.es/es/elclima/cambio\\_climat/escenarios](http://www.aemet.es/es/elclima/cambio_climat/escenarios)

#### 6.4.2 *Proceso de modelado*

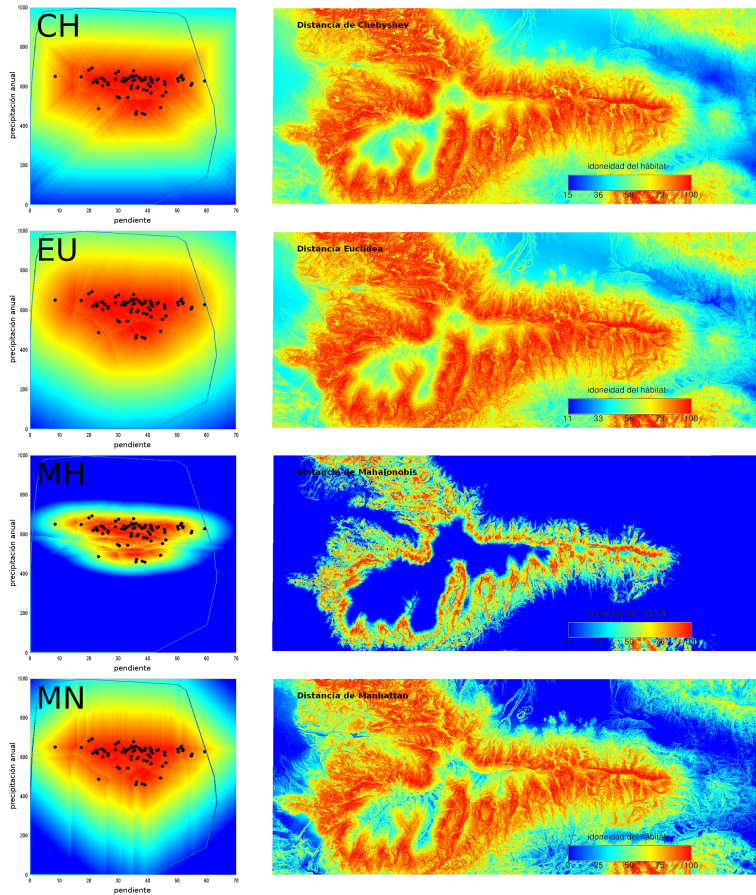
El proceso de modelado que implementado en el flujo de trabajo sigue las pautas ofrecidas en el Capítulo 2, adhiriéndose por tanto a las últimas tendencias en el mundo de los MDE. La clave del método para ofrecer simulaciones robustas se basa en modelado múltiple y ensamblado de los resultados, teniendo en cuenta la incertidumbre. A continuación se describen detalles fundamentales del proceso de modelado que implementado en el flujo de trabajo.

##### *Algoritmos*

Como motores de modelado el flujo de trabajo incluye algunos de los algoritmos disponibles en OpenModeller, y MaxEnt, hasta un total de ocho:

**Environmental Distance:** Implementado en OpenModeller, es un método general de análisis de similaridad, que calcula la distancia ecológica entre cualquier celda del territorio y el punto de presencia de la especie más cercano (en el espacio ecológico). OpenModeller proporciona cuatro métricas diferentes, que se considerarán cuatro algoritmos distintos (ver Figura 4):

- **Chebyshev (CH):** ofrece como resultado la máxima distancia posible entre dos puntos, en cualquier dimensión de los ejes de coordenadas. Se conoce también como distancia del tablero de ajedrez.
- **Euclídea (EU):** mide al recorrido más corto entre el punto que representa a la celda y el punto de presencia de la especie más cercano, según la geometría más convencional e intuitiva.
- **Mahalonobis (MH):** es similar a la euclídea, pero tiene en cuenta la correlación entre las variables a la hora de ponderar la medida.
- **Manhattan (MN):** es la suma de las diferencias absolutas de las coordenadas de los dos puntos entre los que se quiere medir la distancia.



**Figura 4:** Modelos de idoneidad del hábitat según los algoritmos CH, EU, MH y MN de la planta *Moehringia fontqueri* según las variables pendiente y precipitación acumulada anual, mostrados en el espacio ecológico (eje x: pendiente; eje y: precipitación anual) y geográfico.

Los cuatro algoritmos pueden configurarse según dos parámetros:

- Puntos más cercanos: número de puntos de presencia tomados como referencia para medir la distancia. Pueden ser todos o varios, y entonces la distancia se calcula utilizando el centroide de los puntos, o puede ser uno, en cuyo caso se mide la distancia al más cercano.



- **Máxima distancia:** es un parámetro de referencia que indica una distancia (ecológica) umbral a partir de la cual ya no se asigna a una celda un valor de distancia, por considerarse suficientemente alejadas de las condiciones apropiadas para la especie.

**GARP (GA):** es un algoritmo genético implementado en OpenModeller. Utiliza un mecanismo similar a la evolución genética biológica para desarrollar un conjunto de reglas válidas para predecir la distribución de la especie. Una ejecución de **GA** comienza con una población de reglas generadas al azar (una ejemplo de regla: la especie está presente entre los 500 y los 800 milímetros de precipitación anual). Estas reglas se generan mediante distintos métodos: regresión logística, envuelta climática (la especie está entre los valores  $a$  y  $b$  de la variable), envuelta climática negativa (la especie **no** está entre los valores  $a$  y  $b$  de la variable). **GA** genera un modelo de distribución con cada regla, los evalúa todos, y a la siguiente generación solo pasan aquellas reglas que han superado cierto umbral en el proceso de evaluación. El proceso itera  $n$  ocasiones hasta que se alcanza un criterio de parada (todas las reglas por encima de determinado valor de evaluación). **GA** añade un proceso de mutación y recombinación de reglas estocástico en el paso de una generación a otra, por lo que el resultado no es determinista. Una sola simulación de **GA** es una superficie binaria de presencia-ausencia. Para generar una superficie continua, el algoritmo genera  $n$  modelos **GA**, selecciona los mejores y los suma. En la Figura 5 puede observarse la curiosa forma que dibuja **GA** en el espacio ecológico.

**Redes Neuronales Artificiales (NN):** son conjuntos de neuronas artificiales organizadas en capas, que tienen como objetivo proporcionar respuestas a un problema concreto del mismo modo que lo haría un cerebro. La red neuronal del algoritmo implementado en OpenModeller tiene tres capas. En la capa de entrada la red tiene tantas neuronas como variables ambientales se usan para calibrar el modelo. En la segunda un número de neuronas determinado empíricamente, y una sola neurona en la capa de salida. El algoritmo de aprendizaje de la red se basa en un mecanismo de retro-

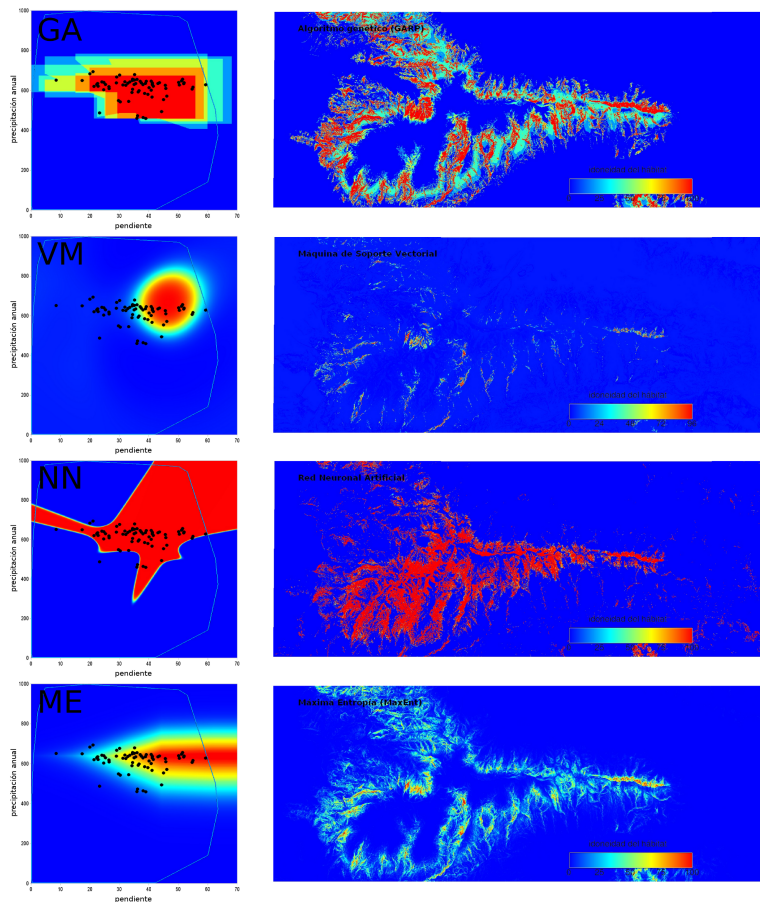
propagación. La red lee los valores de las variables sobre los puntos de presencia y  $n$  puntos aleatorios para construir un mecanismo de clasificación. En el siguiente paso la red clasifica celdas, y compara el resultado con la realidad. La diferencia entre la predicción de la red y la observación real da lugar a un reajuste en la ponderación de las conexiones entre neuronas *aguas arriba* (retropropagación). Esta operación se repite hasta que el error posible queda por debajo de cierto grado de tolerancia.

**Support Vector Machines (VM):** es un método de aprendizaje artificial diseñado para clasificar elementos según su pertenencia a distintos grupos. El algoritmo toma como entrada dos vectores de datos (presencias y puntos aleatorios) en un espacio de  $n$  dimensiones determinado por las variables ambientales. Utiliza un algoritmo iterativo para generar un hiperplano (el equivalente de un plano en un espacio de  $n$  dimensiones) que maximiza las diferencias entre ambos vectores. Cuanto mayor es la distancia entre un objeto de un vector y el hiperplano, mayor será la probabilidad de pertenencia al grupo a ese *lado* del hiperplano.

**MaxEnt (ME):** es un método de propósito general diseñado para generar predicciones a partir de información incompleta. MaxEnt utiliza un mecanismo de inteligencia artificial para buscar iterativamente la distribución de probabilidad de presencia de la especie más uniforme posible (con la máxima entropía). Esta ecuación se proyecta al espacio geográfico con un rango de valores de 0 a 100 que expresa en valores relativos la idoneidad del hábitat para la especie.

### *Ensamblado de modelos*

Para conseguir pronósticos robustos el flujo de trabajo aplica un método de ensamblado no selectivo (Marmion *et al.*, 2009), promediando los resultados de los ocho algoritmos en un mapa de idoneidad media (EC, de Ensamblado Continuo). La función promedio ofrece mejores resultados que la mediana (y que otros métodos de ensamblado) según Marmion *et al.* (2009). Según estos autores, el buen desempeño del método de promedio se debe a la habilidad como filtro de paso bajo que tiene esta función, que produce un



**Figura 5:** Modelos de idoneidad del hábitat según los algoritmos GA, NN, VM y ME.

efecto de limpieza similar al que difumina errores en una fotografía al aplicarle un filtro de desenfoque.

Como novedad respecto a otros estudios en los que se aplica el ensamblado de modelos introduzco el estudio de la dispersión de los valores de cada celda, calculando la desviación estándar de los valores de idoneidad del ensamblado. La desviación estándar mide el grado de coincidencia de los distintos componentes del ensamblado en cuanto a valores de idoneidad. Cuando la coincidencia es baja, la desviación es alta, y la incertidumbre sobre el valor de idoneidad media también. Por este motivo se toma la desviación estándar co-

mo una medida directa de la incertidumbre de la idoneidad media.

Para mejorar la visualización de los resultados, el flujo de trabajo implementa un algoritmo basado en las ideas sobre visualización de la incertidumbre de Hengl *et al.* (2004)<sup>21</sup>, que combina en un solo mapa la idoneidad media y la desviación estándar utilizando una técnica denominada *whitening*, que se basa en cálculos con los colores de los mapas (según el modelo RGB). En el mapa resultante, que en adelante llamaré **FID** (Fusión de Idoneidad y Desviación), los valores de idoneidad del hábitat se codifican mediante distintos colores, mientras que la incertidumbre se codifica mediante la intensidad de los colores. Cuanto más hacia el blanco se desplaza un color (de ahí la denominación *whitening*), mayor es la desviación estándar asociada al valor de idoneidad.

Para interpretar el mapa FID es necesaria una leyenda bidimensional, con la idoneidad del hábitat en el eje *y* la incertidumbre en el eje *x*, tal y como puede verse en la Figura 6c.

### *Evaluación de modelos*

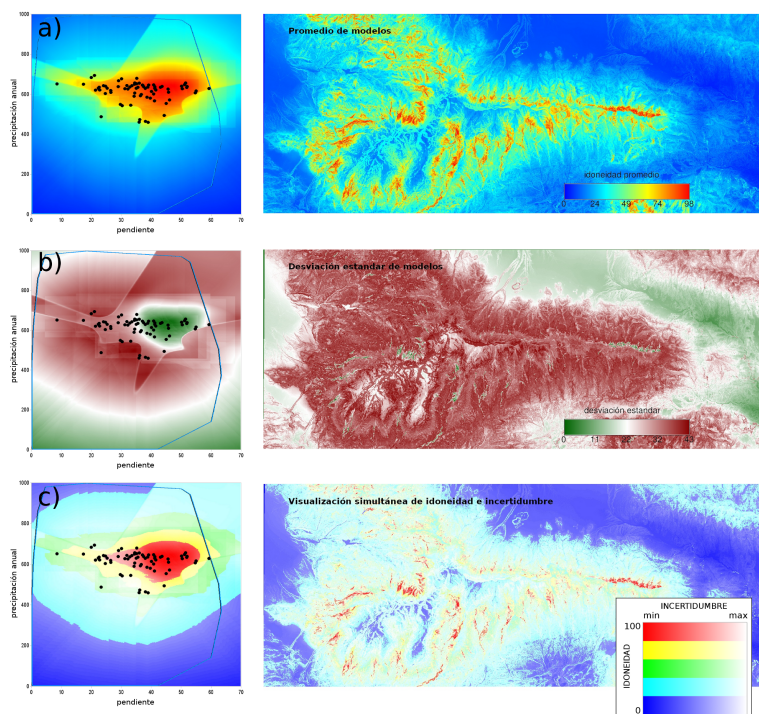
A lo largo de los capítulos anteriores se ha podido comprobar que la evaluación estadística de los MDE es un paso fundamental para obtener información objetiva sobre su fiabilidad. Esta cuestión se vuelve especialmente importante cuando se trabaja con un gran número de especies, y es difícil afrontar una evaluación experta de todos los modelos.

El flujo de trabajo, además de una herramienta de modelado, es un auténtico laboratorio virtual que permite hacer distintos tipos de pruebas con gran número de especies. Aprovechando esta cualidad, en el flujo de trabajo implementé otros cuatro índices de evaluación.

- **AUC global:** Se trata de una evaluación convencional mediante el método AUC descrito en el Capítulo 3 (sustituyendo el análisis ANOVA por uno gráfico). El

---

<sup>21</sup> En su sitio web profundiza en la metodología: [http://spatial-analyst.net/wiki/index.php?title=Uncertainty\\_visualization](http://spatial-analyst.net/wiki/index.php?title=Uncertainty_visualization)



**Figura 6:** Visualización de la incertidumbre. **a** representa el promedio de idoneidad, mayor cuanto más cálido es el color. **b** representa la desviación estándar (o incertidumbre), menor cuando el color es verde y mayor cuando es rojo. **c** es el mapa FID, la fusión de los anteriores mediante un algoritmo de *whitening*. Según los distintos colores tienden al blanco, mayor es la incertidumbre sobre el valor de idoneidad que representan.

apelativo de global se refiere a que los puntos aleatorios se generan sobre todo el área de trabajo.

- **AUC local:** Calculada como en el caso anterior, pero los puntos aleatorios se generan dentro del ámbito de presencia de la especie, que es el mínimo rectángulo que envuelve todas las poblaciones conocidas.
- **Índice Compuesto (IC):** Combina los dos índices anteriores, ponderando **AUC local** al doble que **AUC global**, y de forma que el máximo posible en la suma

de ambos sea igual a 1, para mantener la escala de referencia original entre 0 y 1.

- **Correlación entre valores de idoneidad y desviación de las presencias de evaluación (CIDp):** La idea que se explora con este índice es la siguiente: cuanto mayor es el promedio de idoneidad y menor es la desviación estándar en las presencias de evaluación, más robusto es el ensamblado, porque está determinando el valor de idoneidad con menor dispersión. La medida de evaluación propuesta es el coeficiente de correlación de Pearson obtenido del análisis de correlación.

Las tres primeras medidas se aplican a los resultados de todos los algoritmos y al promedio de idoneidad, para comparar el desempeño del ensamblado respecto a los resultados individuales.

#### *Umbral de corte*

El mapa de promedio de idoneidad tiene valores continuos de 0 a 100, pero para obtener valores como área de ocupación potencial, o migración altitudinal es necesario transformarlo en un mapa binario de presencia y ausencia, seleccionando un umbral de corte.

Según Liu *et al.* (2005) el promedio de los puntos de evaluación puede ser un buen criterio, pero unas pruebas realizadas durante los trabajos del Capítulo 5 ofrecieron unos resultados demasiado restrictivos según este criterio de promedio.

En su lugar, por ser más conservador, he seleccionado un criterio basado en un porcentaje de omisión, y será más fácil explicarlo con un ejemplo: seleccionando un porcentaje de omisión igual a 10, el umbral de corte se establece en el valor de idoneidad que deja fuera del modelo el 10% de los puntos de evaluación, que pasan a ser considerados *outliers*.

#### *Fases de modelado*

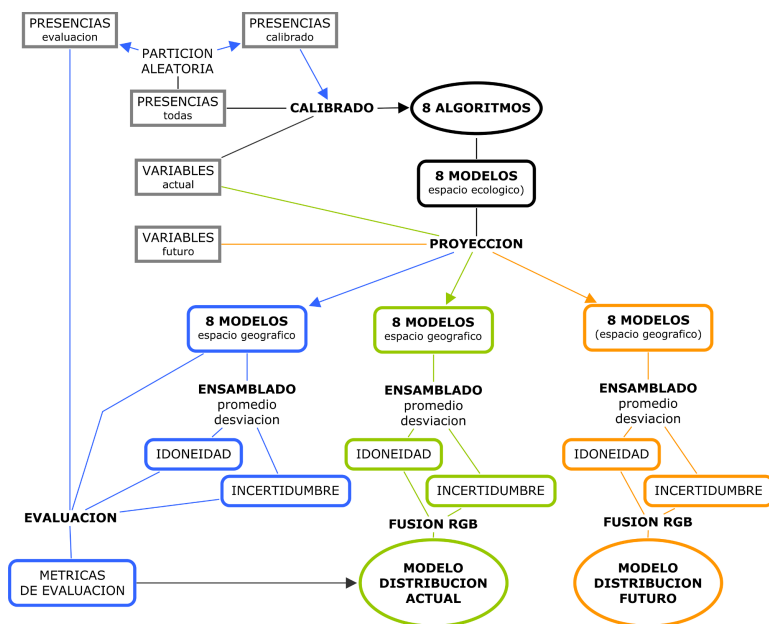
El proceso de modelado supone tres fases, tal y como se aprecia en la Figura 7.

1. **Calibrado inicial y evaluación:** una muestra de calibrado extraída al azar de los registros de presencia se utiliza para calibrar (usando variables actuales) los ocho algoritmos de modelado. Los ocho modelos numéricos resultantes se proyectan al espacio geográfico (sobre las mismas variables usadas para calibrarlos), obteniéndose ocho mapas digitales de idoneidad del hábitat. Estos mapas de idoneidad se ensamblan según el método explicado en el apartado anterior. Todos los modelos (el ensamblado y sus componentes) se evalúan según distintos métodos para obtener métricas de evaluación.
2. **Calibrado:** usando todos los registros de presencia disponibles, se calibran de nuevo los ocho algoritmos, y se repite el proceso de proyección sobre las variables actuales y ensamblado para obtener el modelo de distribución actual, punto de partida de las simulaciones. A este *nuevo* modelo se le asignan como medidas de precisión los valores de evaluación obtenidos en la fase anterior.
3. **Proyección sobre distintos escenarios:** A partir de los modelos numéricos obtenidos en la fase anterior, se realizan tantas proyecciones sobre variables futuras como escenarios e intervalos de tiempo se quieren simular, repitiendo el proceso de ensamblado.

## 6.5 EL FLUJO DE TRABAJO

A continuación resumiré los detalles fundamentales del funcionamiento del flujo de trabajo.

El flujo de trabajo diseñado consta realmente de dos flujos diferenciados. El primero, denominado **PRP** (Preparación de Registros de Presencia) se encarga de transformar los ficheros vectoriales de presencia de las especies en pares de coordenadas, imponiendo ciertas restricciones. El segundo, denominado **MODPLAN** (MOdelado de Distribución de PLANtas) es el que ejecuta las simulaciones. La separación



**Figura 7:** Proceso de modelado: Los rectángulos grises representan datos de entrada. La sección azul representa la Fase 1 (Evaluación); la sección verde representa la Fase 2 (Calibrado); la sección naranja representa la Fase 3 (Proyección).

del flujo de trabajo en dos secciones de ejecución independiente tiene un motivo sencillo. PRP prepara los puntos de presencia según condiciones establecidas por el investigador, como se verá en la próxima sección, y ejecutar el proceso de preparación de presencias independientemente al de simulación permite tomar decisiones sobre las especies a incluir (o excluir) del análisis, o cambios en los parámetros de selección. De este modo, se ejecuta PRP, se observan los resultados para tomar decisiones, y posteriormente se ejecuta MODPLAN para generar las simulaciones.

### 6.5.1 PRP

La información de presencia de las especies está almacenada como *shapefiles* de geometría poligonal, no aprovechables directamente por los programas de modelado MaxEnt y OpenModeller, por lo que los polígonos deben transformarse



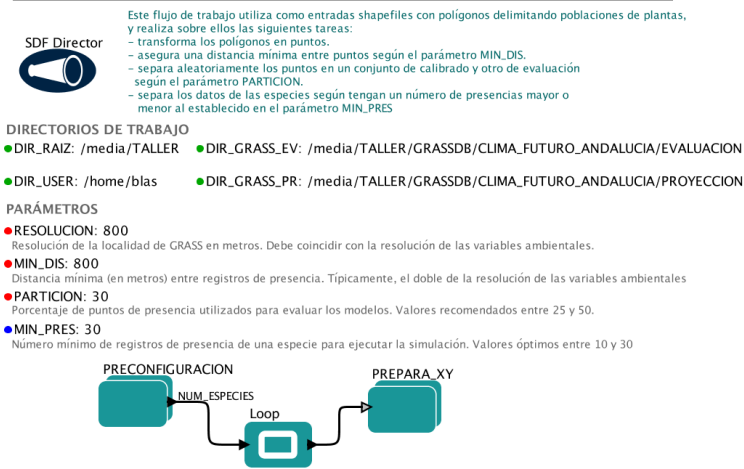
en pares de coordenadas, que deben tener una serie de características:

- Baja autocorrelación espacial: puntos demasiado cercanos entre sí aportan información redundante al modelo. Un modo de disminuir esta redundancia por proximidad es garantizar una distancia mínima entre pares de coordenadas.
- Deben ser separados al azar en un conjunto de calibrado (proporción mayoritaria) y otro de evaluación, aplicando un proceso de partición aleatoria de datos. También es importante disponer de ambos conjuntos fusionados, para poder calibrar los modelos con toda la información disponible.
- Los pares de coordenadas deben estar formateados como ficheros de texto plano, según una estructura de columnas específica de cada programa de modelado.
- Debe existir un número mínimo de puntos que garantice cierta calidad en los modelos resultantes. Una única muestra con un número escaso de puntos no proporciona a ningún algoritmo la información suficiente para caracterizar el hábitat apropiado para una especie.

En el flujo de trabajo **PRP** (ver Figura 8) el usuario determina varios parámetros de importancia para la preparación de los puntos de presencia:

- **RESOLUCION**: Resolución de trabajo deseada para ejecutar las simulaciones. Lo ideal es seleccionar la resolución nativa de las capas de variables ambientales.
- **MIN\_DIS**: Distancia mínima entre pares de coordenadas de presencia de la especie. Idealmente, debería ser al menos el doble de la resolución de trabajo, para garantizar la presencia una celda vacía entre punto y punto de presencia.

## PRP (Preparacion de Registros de Presencia)



**Figura 8:** Flujo de trabajo PRP. PRECONFIGURACIÓN y PREPARA\_XY son actores compuestos que contienen flujos de trabajo anidados.

- **PARTICIÓN:** Porcentaje de los registros de presencia que se utilizará como conjunto de evaluación. Lo habitual es utilizar entre un 25 y un 50 por ciento de los puntos con este propósito.
- **MIN\_PRES:** Número mínimo de registros de presencia necesarios para que una especie pueda pasar a la fase de modelado.

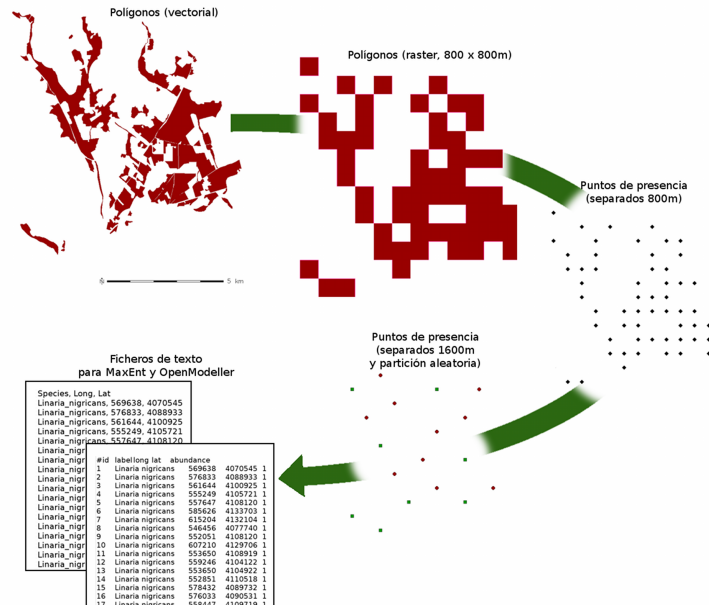
Otros parámetros de menor importancia técnica pero imprescindibles para el funcionamiento del flujo son los directorios de ejecución y las rutas de acceso a los directorios de mapas de GRASS.

PRP está formado por dos actores compuestos que contienen flujos de trabajo anidados, PRECONFIGURACIÓN y PREPARA\_XY, y un actor Loop. El actor PRECONFIGURACIÓN solo se ejecuta una vez y prepara el entorno de trabajo creando los directorios en los que se almacenarán los resultados. También cuenta el número total de especies para iterar el proceso de preparación sobre una de ellas. El actor Loop recibe el número de especies de PRECONFIGURACIÓN

*Un número mínimo conservador de presencias es 30, aunque es posible que el método de ensamblado pueda funcionar con números algo menores.*

RACIÓN, y lo envía a PREPARA\_XY para que se ejecute tantas veces como especies se quieren procesar.

La secuencia de acciones que realiza el flujo queda representada en la Figura 9. Un *script* de GRASS importa el fichero vectorial, y lo transforma en raster según la resolución indicada en los parámetros de PRP. El raster se vuelve a transformar en vectorial, pero usando puntos (un punto por celda) en lugar de polígonos.



**Figura 9:** Resumen visual de las tareas ejecutadas por PRP, usando como ejemplo la distribución en Tabernas de *Linaria nigricans*, a una resolución de 800 metros por celda.

Los puntos, exportados como pares de coordenadas, son tratados mediante un algoritmo en Octave, que actúa del siguiente modo: Tomando como referencia las coordenadas del primer punto de la lista (punto **a**), mide la distancia al siguiente (punto **b**) aplicando el teorema de Pitágoras. Si la distancia es menor que la indicada por el parámetro MIN\_DIS, elimina el punto **b** y pasa a repetir la tarea con el siguiente punto de la lista (punto **c**). Si la distancia es mayor, pasa medir la distancia con el siguiente punto (punto **d**).

Cuando ha terminado de comprobar todo el listado, pasa al punto **b** y repite la secuencia. Una vez terminado este filtrado según distancias, el programa hace una partición aleatoria de los datos para preparar un conjunto de calibrado y otro de evaluación.

Posteriormente Kepler termina de editar los ficheros de coordenadas para MaxEnt y OpenModeller, y los sitúa en los directorios apropiados.

Por último Kepler mide el número de presencias total de la especie, y si es menor que el valor MIN\_PRES, envía ficheros de la especie a un directorio de especies descartadas.

### 6.5.2 MODPLAN

El flujo de trabajo **MODPLAN** (ver estructura en Figura 10) utiliza los resultados de **PRP** y las variables ambientales del presente y simulaciones de clima futuro para generar las simulaciones realizando el proceso de modelado descrito en la Figura 7. Su estructura y funcionamiento son más complejas, pero trataré de resumirlas (Para más detalles acudir al ANEXO\_DETALLES\_FLUJO).

En la Figura 10 puede observarse que **MODPLAN**, además de varios actores interconectados, presenta un parámetro llamado PORCENTAJE\_OMISIÓN, con el que el usuario determina el valor de omisión a partir del cuál el sistema calculará el umbral de corte para transformar los ensamblados de valores continuos en mapas binarios de presencia-ausencia. Los actores que forman el flujo son NUMERO\_ESPECIES y EJECUCIÓN\_MODELOS, un actor Loop y un actor Display. NUMERO\_ESPECIES lee los resultados de **PRP** e indica a Loop el número de especies que cumplen los requisitos para ser modeladas. También lee el directorio de variables ambientales para listar cuáles son esas variables, y los escenarios de cambio climático que representan. NUMERO\_SELECCIONADAS le muestra al usuario el número de especies que se van a modelar. Loop obliga a EJECUCIÓN\_MODELOS a iterar sobre el listado completo de especies seleccionadas para el modelado. EJE-

## MODPLAN (Modelado de Distribucion de Plantas)

Este flujo de trabajo ejecuta las siguientes tareas:

- Calibra, ensambla y evalua modelos de distribucion de "n" especies
- Proyecta los modelos sobre escenarios de cambio climatico
- Genera un completo informe de resultados

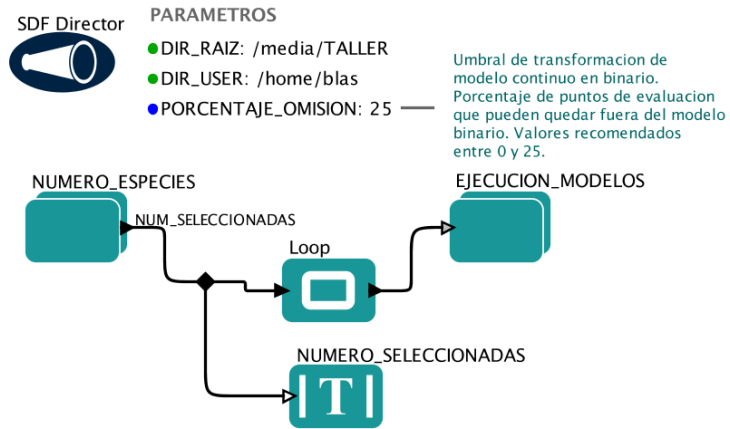


Figura 10: Flujo de trabajo MODPLAN.

CUCIÓN\_MODELOS Figura 11 efectúa el trabajo duro de MODPLAN, llevando a cabo las siguientes tareas:

- Análisis gráfico de los requerimientos ecológicos de la especie.
- Calibrado inicial de los algoritmos con los puntos de calibrado y ensamblado de los resultados.
- Evaluación de los modelos usando distintas métricas.
- Re-calibrado con todos los puntos de presencia.
- Proyección de los modelos sobre los distintos escenarios e intervalos de tiempo.
- Extracción de datos de los ensamblados en GRASS y análisis gráfico en R.
- Almacenamiento de los ensamblados en una base de datos geográfica de GRASS.
- Preparación de los resultados en un informe.

Para analizar los requerimientos ecológicos de la especie **MODPLAN** compara la distribución de frecuencia de los registros de presencia (sobre las variables de calibrado) con la distribución de 10000 puntos distribuidos aleatoriamente a lo largo del territorio de trabajo. Es GRASS el encargado generar los puntos aleatorios y extraer los valores de estos y los puntos de presencia sobre las variables ambientales (del periodo de referencia o *actuales*) usadas para calibrar los modelos. Con esos valores R dibuja gráficas de densidad de fácil interpretación.

Para calibrar el modelo de evaluación, Kepler genera los ficheros de configuración que OpenModeller necesita para generar los modelos, y Bash lanza la ejecución de OpenModeller y MaxEnt.

Los modelos resultantes se importan y ensamblan (idoneidad media, desviación estándar y mapa FID) en GRASS, donde se extraen los valores de los puntos de evaluación y 10000 puntos aleatorios, que pasan a Octave para el cálculo de la curva ROC y el índice compuesto. Los resultados de evaluación de Octave pasan a R para generar las gráficas.

El sistema modifica los ficheros de configuración de OpenModeller cambiando la dirección al fichero de puntos de calibrado por la del fichero de puntos completos para calibrar los modelos con toda la información disponible. El calibrado de los algoritmos produce como resultado 8 modelos numéricos. En OpenModeller se denominan *modelos serializados*, y son ficheros *xml*<sup>22</sup> con las especificaciones del modelo. En MaxEnt se trata de un fichero de texto plano con los coeficientes de la ecuación (*lambdas*). Estos modelos numéricos se proyectan sobre el espacio geográfico y los resultados se importan y ensamblan (idoneidad media, desviación estándar y mapa FID) en GRASS.

Bash utiliza los modelos numéricos resultantes del paso anterior para ejecutar iterativamente las proyecciones sobre los distintos escenarios e intervalos de tiempo. En cada iteración, GRASS extrae los valores de idoneidad y desviación de las presencias y los puntos aleatorios sobre los ensamblados

---

<sup>22</sup> [http://es.wikipedia.org/wiki/Extensible\\_Markup\\_Language](http://es.wikipedia.org/wiki/Extensible_Markup_Language)

(idoneidad y desviación), almacenándolos en una tabla, y dibuja mapas de los modelos.

La tabla resultante del paso anterior se procesa en R para obtener un análisis gráfico muy completo de las consecuencias potenciales de los distintos escenarios sobre las especies.

Todos los resultados quedan organizados y almacenados en una base de datos geográfica de GRASS, para permitir análisis que no hayan sido incluidos en el flujo de trabajo.

Finalmente, los gráficos y mapas que describen los resultados se disponen sobre una plantilla escrita en *html* para visualizar los resultados del modelado en un navegador web, por lo que están listos para su publicación en web.

*El lenguaje html se utiliza para dar estructura y contenido a las páginas web.*

## 6.6 RESULTADOS DEL FLUJO DE TRABAJO

La intención del flujo de trabajo es facilitar la labor de interpretación de los resultados al investigador, para lo que prepara un completo informe. El informe separa la información en tres secciones que cubren los aspectos más importantes para interpretar los resultados:

1. Distribución actual y requerimientos ecológicos de la especie.
2. Modelo de idoneidad del hábitat en el periodo 1971-2000.
3. Efectos potenciales del Cambio Climático en el periodo 2000-2050.

A continuación describo el contenido de los resultados más importantes que aparecen en los distintos apartados, tomando como ejemplo la especie *Moehringia fontqueri*, rupícola endémica de Sierra Nevada almeriense.

### 6.6.1 *Distribución actual y requerimientos ecológicos de la especie*

En esta sección se trata de responder a dos preguntas: ¿Dónde está? y ¿cuáles son sus requerimientos ecológicos?.

### *Distribución conocida*

Para responder a la primera cuestión el informe proporciona dos mapas de distribución conocida de la especie, uno a escala global (todo el área de trabajo) y otro en el ámbito de presencia de la especie (mínimo rectángulo que acoge todas las poblaciones conocidas, más un área de influencia de 2000 metros).

### *Requerimientos ecológicos*

Los requerimientos ecológicos de la especie se expresan mediante gráficas de densidad (obtenidas en R), teniendo en cuenta las variables de calibrado de las simulaciones, los puntos de presencia de la especie y 10000 puntos aleatorios dispersos por el área de trabajo. En las gráficas se comparan las distribuciones de los puntos de presencia y los puntos aleatorios (ver Figura 11) indicando los promedios de cada conjunto. Como norma general, cuanto mayor es la diferencia entre promedios para una variable, más importante será la variable en la distribución de la especie.

#### 6.6.2 *Modelo de idoneidad del hábitat en el periodo 1971-2000*

Esta sección del informe es la más compleja, porque trata los aspectos más técnicos de la simulación.

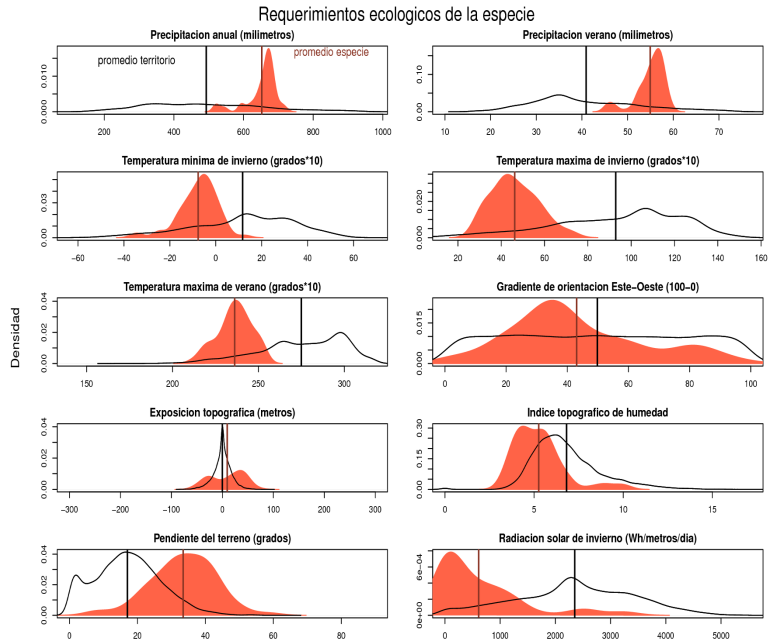
### *Calibrado inicial de algoritmos*

Este apartado del informe muestra los resultados de los ocho algoritmos, que se han entrenado con los puntos de calibrado, y evaluado con los puntos de evaluación. La intención es proporcionar una idea previa de cómo funcionan los distintos algoritmos con la especie objetivo.

### *Evaluación*

La Figura 12 muestra el análisis gráfico de la evaluación de los algoritmos y el ensamblado según los cuatro métodos propuestos anteriormente en este capítulo. Los gráficos **a** y **b** permiten saber qué algoritmos se ajustan mejor a los



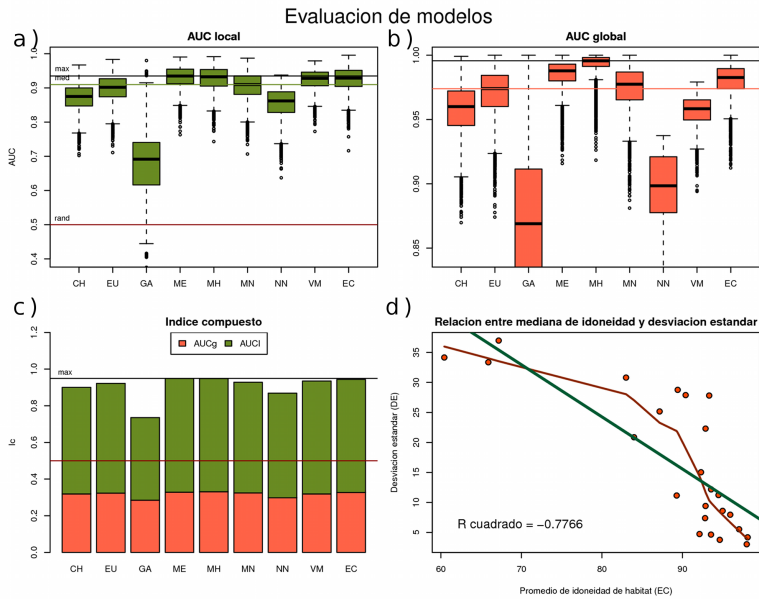


**Figura 11:** Gráfico de requerimientos ecológicos de la especie. La curva negra representa los puntos aleatorios dispersos por todo el territorio. La curva naranja representa los puntos de presencia de la especie. Las líneas verticales indican la posición del promedio de cada grupo. Cuanto mayor es la distancia entre los promedios, más probable es que la variable tenga una gran influencia en la distribución de la especie.

datos de evaluación en dos escalas geográficas distintas, y entender hasta que punto el ensamblado (EC) puede llegar a representar la distribución de la especie. También permiten identificar algoritmos que han fallado en la predicción.

### *Recalibrado*

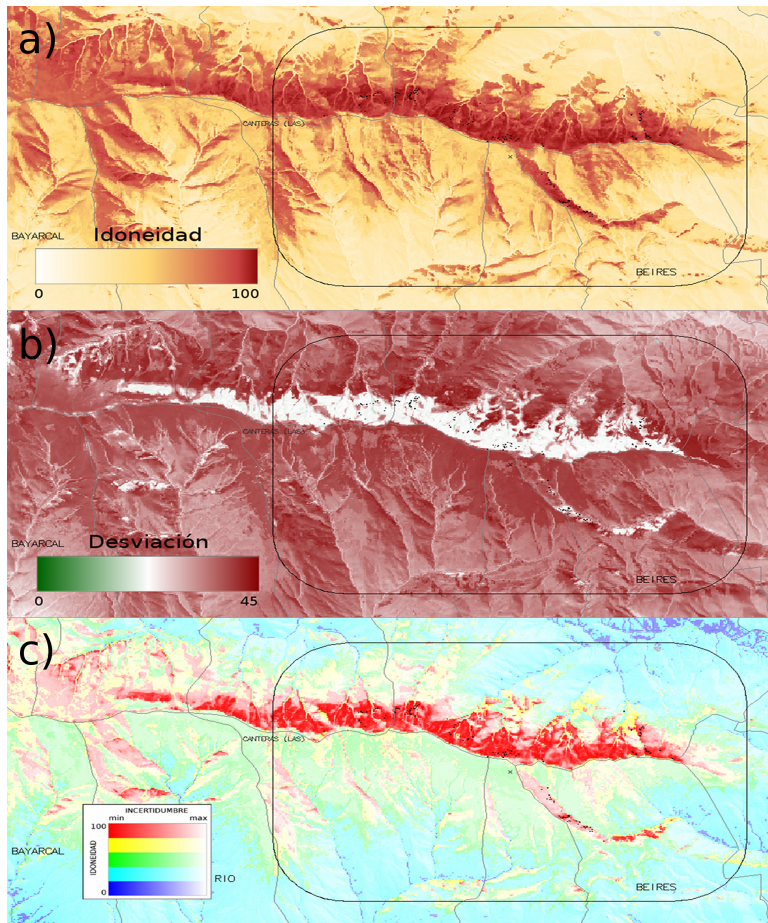
En esta sección el informe muestra los resultados del ensamblado, recalibrado con todos los puntos de presencia de la especie, en sus tres versiones: promedio de idoneidad, desviación estándar y mapa FID (Fusión de Idoneidad y Desviación; ver Figura 13).



**Figura 12:** Evaluación del modelo: Los gráficos **a** y **b** comparan entre algoritmos (CH, EU, GA, ME, MH, MN, NN, VM)) y ensamblado (EC) los valores de AUCg y AUCI. El gráfico **c** muestra el índice compuesto (IC). En **d** se muestra el análisis CIDp. En verde, la recta de regresión de los puntos de evaluación para la idoneidad y la desviación en el ensamblado, con su coeficiente de regresión (valor de CIDp).

### *Diferencias entre el modelo de evaluación y el modelo de calibrado*

El modelo de evaluación se calibra con un determinado porcentaje de puntos de presencia (puntos de calibrado), mientras que el modelo definitivo se calibra con todos los puntos de presencia. Esta diferencia entre los datos de entrada provoca que ambos modelos no sean exactamente idénticos. En esta sección se exploran las diferencias entre ambos modelos mediante gráficos y un mapa de diferencias (ver Figura 14). El análisis gráfico que se presenta es útil para determinar hasta que punto las métricas de evaluación obtenidas del primer modelo pueden asignarse al segundo, y para conocer el efecto sobre el modelo de la adición de nuevos puntos de presencia.

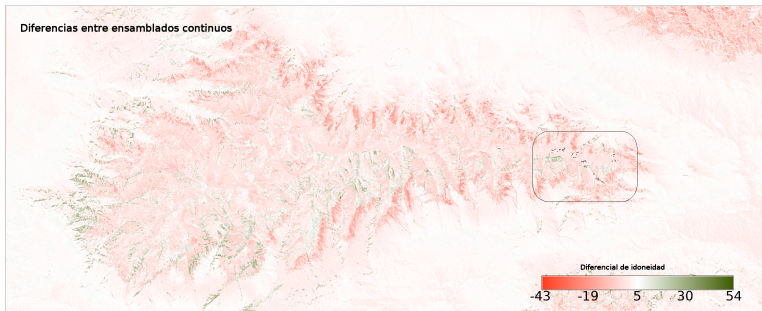


**Figura 13:** La figura **a** muestra un detalle del mapa de promedio de idoneidad. **b** muestra la misma zona, pero con los valores de desviación estándar. **c** muestra el mapa FID obtenido mediante la técnica de *whitening*.

### *Curvas de respuesta del modelo*

En este apartado se muestra un análisis de la relación entre los valores de idoneidad del ensamblado y cada una de las variables de calibrado (ejemplo en la Figura 15).

El ensamblado está formado por algoritmos con distintas lógicas internas, agregadas mediante una función de promedio. La propia dispersión de los componentes del ensamblado hace que sea muy difícil que el promedio de

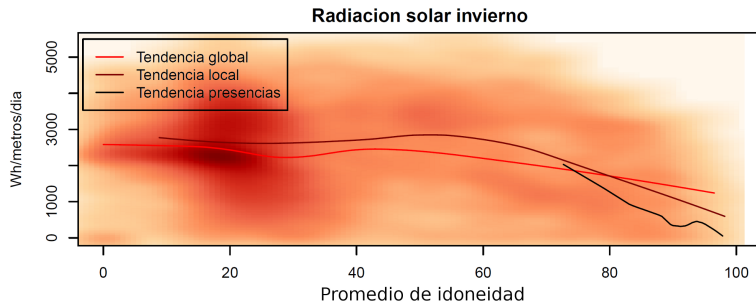


**Figura 14:** Mapa de diferencias entre el modelo de evaluación y el modelo de calibrado. El color rojo indica áreas en las que el modelo recalibrado presenta valores de idoneidad menores que el modelo de evaluación. La aparición de estas áreas con diferencial negativo puede interpretarse como una mejora en el modelo producida por la adición de nuevos puntos de calibrado.

idoneidad tenga una relación curvilínea fuerte con alguna de las variables de calibrado. Para tratar de encontrar patrones de respuesta entre la idoneidad y los valores de las variables el análisis se hace a tres escalas espaciales. La línea (**Tendencia global**) relaciona los valores de idoneidad y los de la variable para 10000 puntos aleatorios dispersos por todo el área de trabajo. La línea **Tendencia local** relaciona 10000 puntos aleatorios dispersos en el ámbito de presencia de la especie con la misma variable. La línea **Tendencia presencias** representa la misma relación tomando como referencia los puntos de presencia de la especie.

#### *Efectos potenciales del Cambio Climático en el periodo 2000-2050*

Esta sección pretende mostrar información suficiente para evaluar el posible efecto del cambio del clima en la distribución de la especie, ofreciendo animaciones de las simulaciones, analizando la variación temporal de la idoneidad del hábitat en las poblaciones conocidas de la especie, las variaciones en área potencial, la fragmentación del hábitat, y la migración, tanto altitudinal como latitudinal.

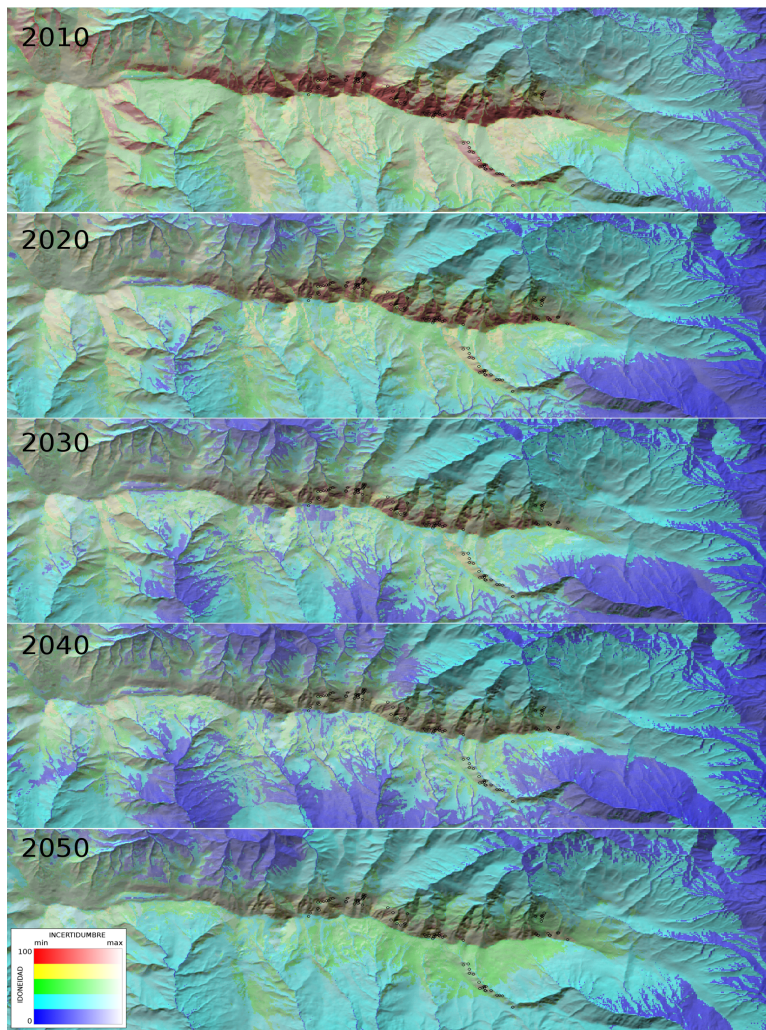


**Figura 15:** Curvas de respuesta del ensamblado: La nube de fondo (10000 puntos aleatorios dispersos por el área de trabajo) representa en una escala de color relativa la proporción de celdas del territorio que coinciden en los valores de idoneidad y radiación solar. Como ejemplo de interpretación, hay una gran proporción de celdas con idoneidad en torno a 20 y radiación en torno a 2300, pero apenas hay celdas con idoneidad 100 y radiación 5000. Respecto a la interpretación de las líneas, **Tendencia presencias** muestra que la idoneidad en las localidades de presencia aumenta según disminuye la radiación solar. Esta relación también puede detectarse en **Tendencia local**, pero podría pasar casi desapercibida según **Tendencia global**.

### *Animación de series temporales*

La animación es un lenguaje visual potente, y en el informe de resultados se aplica para mejorar la experiencia de interpretación de resultados. Pueden visualizarse las series temporales de los cuatro escenarios para la idoneidad del hábitat, la desviación estándar y el mapa FID, que es con seguridad la más interesante (ver Figura 16). En las animaciones el mapa FID se fusiona con un mapa de relieve (que representa de modo intuitivo mediante sombras el relieve de la región) mediante el módulo HIS de GRASS. Este módulo permite fusionar hasta tres mapas digitales, de forma que el primero determina el tono (**Hue**), el segundo la intensidad del tono (**Intensity**) y el tercero la saturación del color (**Saturation**). En este caso la capa H es el mapa FID, y la capa I el mapa de relieve. Este tipo de visualización, aunque altera los colores originales del mapa FID, mejora la interpretación de los resultados permitiendo al investiga-

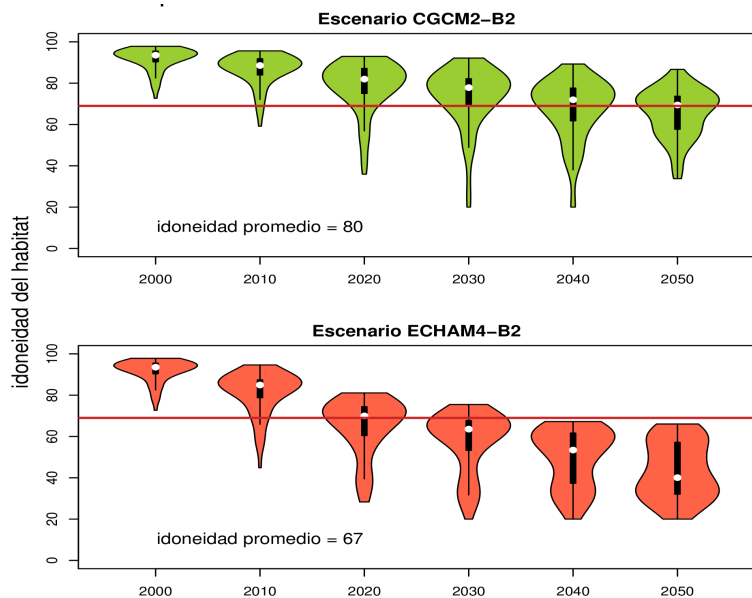
dor observar simultáneamente la idoneidad del hábitat, su desviación del promedio y el tipo de topografía asociada al hábitat.



**Figura 16:** Serie temporal del mapa FID *Moehringia fontqueri* según el escenario CGCM2-B2.

### Cambios en la idoneidad del hábitat sobre las poblaciones actuales de la especie

En este apartado se analizan gráficamente las series temporales de idoneidad del hábitat en los puntos de presencia de la especie (ver Figura 17). El análisis se basa en unos gráficos similares a los gráficos de cajas (*boxplot*) llamados gráficos de violín, generados mediante el paquete de R *violinplot*. El gráfico muestra para cada intervalo de tiempo la distribución de los valores de idoneidad de los puntos de presencia usados para calibrar el modelo. Es una información de fácil interpretación en el ámbito de la conservación, porque muestra hasta qué punto van a cambiar las condiciones del hábitat en las poblaciones conocidas.



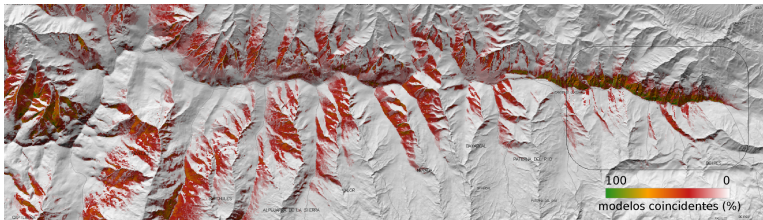
**Figura 17:** Idoneidad del hábitat en las poblaciones conocidas de la especie (puntos de calibrado) según los escenarios CGCM2-B2 y ECHAM4-B2. Las curvas con forma de violín representan la proporción de puntos de presencia que corresponde a cada nivel de idoneidad del hábitat. La **zona roja** determina el nivel mínimo de idoneidad por debajo del cual es poco probable encontrar poblaciones viables de la especie.

### *Tendencias potenciales del hábitat idóneo*

Este apartado ofrece un análisis de las tendencias del hábitat idóneo de la especie desde cuatro perspectivas: área de hábitat idóneo (como medida de área potencialmente disponible para la especie), número de parches de hábitat idóneo (medida simple de la fragmentación del hábitat), migración altitudinal y migración latitudinal. Como todas las medidas se refieren al hábitat potencial, no puede esperarse que la especie realmente se desplace vertical u horizontalmente lo que indican las respectivas gráficas, sino que los valores deberían tomarse como el desplazamiento potencial que la especie debería realizar para mantenerse en su hábitat idóneo.

### *Mapa de áreas de persistencia potencial*

Esta sección muestra un mapa que puede resultar de interés para planes de conservación de las especies. El mapa en cuestión condensa la información de todas las simulaciones para localizar aquellos lugares que tienen mayores posibilidades de conservar las condiciones adecuadas para la especie durante mayor tiempo. En la Figura 18 puede verse un ejemplo de áreas de persistencia.



**Figura 18:** Detalle del mapa de persistencia de *Moehringia fontqueri*. Las áreas verdes son las que tienen valores mayores de persistencia potencial. En este caso coinciden con las áreas de presencia conocida, pero hay otras que proporcionarían las mismas condiciones a la especie en el oeste de Sierra Nevada.





## EFFECTOS POTENCIALES DEL CAMBIO DEL CLIMA EN LA DISTRIBUCIÓN GEOGRÁFICA DE LA FLORA DE ZONAS ÁRIDAS Y ALTAS MONTAÑAS

---

### 7.1 INTRODUCCIÓN

#### 7.1.1 *La cuestión biológica*

En el Capítulo 5 traté los potenciales efectos del Cambio del Clima en la distribución geográfica de las formaciones vegetales más importantes de Sierra Nevada. Pero la utilización de formaciones vegetales como sujetos de modelado presenta algunas desventajas en las proyecciones de cambio climático, porque no pueden mantener su composición en especies constante en el tiempo debido a las distintas velocidades de migración de sus elementos constituyentes. Para superar esta limitación, y avanzar un paso en complejidad de los MDE, este nuevo capítulo tratará las especies desde una perspectiva individualista.

Andalucía Oriental es una región rica en gradientes ecológicos extremos, reuniendo hábitats muy diferentes en un espacio geográfico limitado. Esta riqueza de hábitats da lugar a una alta diversidad vegetal, entre las más ricas de la cuenca Mediterránea (Médail y Quèzel, 1999). Nuestro grupo de investigación está especialmente interesado en la flora rara y amenazada de las zonas áridas y las altas montañas, y es una prioridad establecer líneas de actuación para su conservación. Por este motivo, la simulación de cambios potenciales en la distribución de estas especies es una herramienta fundamental en nuestro trabajo.

Una de las cuestiones que se ha planteado durante distintas fases de la investigación, ha sido la posibilidad de que el cambio del clima afecte de modo diferente a hábitats áridos y hábitats de alta montaña. Para afrontar esta cuestión, y disponiendo de una nueva y potente herramienta de modelado

descrita en el capítulo anterior, seleccioné Sierra Nevada (en adelante, SN) como ejemplo paradigmático de alta montaña mediterránea y la depresión de Guadix-Baza (en adelante, GB) como caso de zona árida con alta diversidad vegetal.

Analizar la influencia del cambio del clima en las posibilidades de persistencia de las especies no es sencillo, porque entran en juego muchos aspectos biológicos y metodológicos que es difícil tener en cuenta, como la migración potencial de los organismos, o el efecto de las propiedades de los datos sobre los resultados de los MDE.

*Este tipo de análisis es probablemente el menos especulativo que se puede afrontar con los MDE.*

Para limitar el rango de incertidumbres, el análisis que voy a realizar se limita a los valores de idoneidad del hábitat de las localidades de presencia de las especies. De este modo me centraré en analizar las posibilidades de persistencia de las especies relacionándola con las variaciones de idoneidad del hábitat de sus poblaciones entre los periodos de comienzo y final de las simulaciones. Analizaré también la influencia de las propiedades de la muestra de presencia (extensión de presencia y número de registros) en el diferencial de idoneidad de cada especie, para comprobar si los resultados obtenidos se deben a un artefacto metodológico, o concuerdan con la realidad biológica.

### 7.1.2 *La cuestión metodológica*

Aprovechando que la resolución de la cuestión requiere un trabajo de simulación de cierta envergadura, y será por tanto la primera prueba importante para el flujo de trabajo MOD-PLAN, en este capítulo trato de resolver algunas cuestiones sobre el funcionamiento del flujo de trabajo, y revelar algunas carencias metodológicas, con la intención de proponer mejoras para futuros estudios. Particularmente, me centraré en el funcionamiento de la técnica de ensamblado, tratando de comprobar si realmente el ensamblado de modelos tiene ventajas respecto al uso de algoritmos individuales, aunque los trabajos de [Araújo y New \(2007\)](#) y [Marmion et al. \(2009\)](#) ya adelantan datos al respecto.

## 7.2 OBJETIVOS

El objetivo principal de este trabajo es analizar las variaciones en idoneidad del hábitat en las poblaciones conocidas de varias especies de plantas raras y amenazadas de SN y GB, para obtener una vista previa de los efectos potenciales del cambio del clima sobre la flora comparando los resultados de ambos hábitats.

En torno a este objetivo principal, hay varias cuestiones metodológicas a resolver, referentes al propio proceso de simulación:

- ¿Funciona mejor el ensamblado que los resultados individuales de los distintos algoritmos?.
- ¿Cuáles son las principales carencias de la metodología de simulación implementada en MODPLAN?.

## 7.3 MATERIAL Y MÉTODOS

### 7.3.1 Áreas de trabajo y especies

#### *La depresión de Guadix-Baza*

El territorio de la *depresión de Guadix-Baza* se corresponde mayoritariamente con los límites establecidos por Valle (2004) para el distrito biogeográfico Guadiciano-Bastetano (ver Figura 1), y está compuesto geográficamente por tres subcuencas sedimentarias (Mioceno superior-Pleistoceno): la de Guadix, la de Gorafe-Huélago y la de Baza (García-Aguilar, 2000). La depresión presenta unas irregulares y exiguas precipitaciones (330mm anuales entre 1971 y 2000, cálculos propios a partir de datos de la AEMET), debidas a las barreras geográficas que forma el relieve circundante, que la aísla de la influencia del mar, ocasionando un efecto de sombra de lluvias. La temperatura media anual es de 15.4°C, con extremos que pueden llegar desde los -15°C hasta los 40°C.

La conjunción de condiciones litológicas, topográficas y climáticas explica la naturaleza esteparia de los hábitats, y la abundancia de especies terofíticas en la región. El catálogo

de flora amenazada de la región consta de 43 taxones (6 En Peligro Crítico, 9 En Peligro, 18 Vulnerables y las restantes con datos insuficientes o Casi Amenazadas), de los que solo 6 están legalmente protegidas.

Las especies con suficientes datos de presencia para calibrar las simulaciones fueron: *Astragalus oxyglottis*, *Boreava aptera*, *Centaurea saxifraga*, *Cynomorium coccineum*, *Dorycnium gracile*, *Gypsophila tomentosa*, *Haplophyllum bastetanum*, *Krascheninnikovia ceratoides*, *Limonium majus*, *Limonium que-sadense*, *Loeflingia baetica*, *Plantago maritima serpentina*, *Senecio auricula* subsp. *auricula*, *Vella pseudocytisus* subsp. *pseudocytisus*.

Los mapas digitales de GB presentaron las siguientes características: 40 m de resolución; 2447 filas por 2373 columnas; 5806731 celdas; Límites (UTM ED-50): N=4200249, S=4102369, E=555969, O=461049.

### *Sierra Nevada*

Sierra Nevada es un sistema montañoso situado en el sur de la Península Ibérica (ver Figura 1), que se originó durante el Mioceno (hace unos 20 millones de años). La cordillera se desarrolla en dirección este-oeste a lo largo de 90 kilómetros, con un área total de 2100 kilómetros cuadrados. En estas montañas destacan los gradientes de temperatura, con mínimas absolutas en torno a -20°C y máximas absolutas (condiciones puntuales en zonas más expuestas a la radiación solar) en torno a 30°C en las altas cumbres. La precipitación media anual presenta un patrón descendente en el sentido oeste-este, pero existen grandes variaciones, con áreas en las que la precipitación es menor a 350 mm (áreas basales de la cordillera) y zonas puntuales con precipitaciones en torno a 1200 mm *Blanca et al.* (2000).

Como consecuencia de su estratégica situación, como punto de paso de distintas corrientes migratorias y estación de refugio durante las glaciaciones SN se ha convertido en el centro de diversidad más importante de la Región Mediterránea occidental (*Blanca et al.*, 1998), con 2100 especies de plantas vasculares catalogadas, de las cuales unas 80 son endémicas. Con el objetivo de preservar esta diversidad vege-

tal, SN fue declarada Reserva de la Biosfera por la UNESCO en 1986, Parque Natural en 1987 y Parque Nacional en 1999.

Las especies de SN con suficientes datos para calibrar las simulaciones fueron: *Acer opalus* subsp. *granatense*, *Aconitum burnatii*, *Anthyllis vulneraria* subsp. *arundana*, *Arenaria nevadensis*, *Armeria filicaulis* subsp. *trevenqueana*, *Armeria splendens*, *Artemisia granatensis*, *Betula pendula* subsp. *fontqueri*, *Carex camposii*, *Centaurea pulvinata*, *Chamaespartium undulatum*, *Convolvulus boissieri* subsp. *boissieri*, *Cytisus galianoi*, *Dryopteris submontana*, *Erodium boissieri*, *Euphorbia nevadensis*, *Euphrasia willkommii*, *Festuca indigesta* subsp. *indigesta*, *Gentiana pneumonanthe* subsp. *depressa*, *Helianthemum appenninum* subsp. *estevei*, *Helianthemum pannosum*, *Hippocrepis prostrata*, *Leontodon boryi*, *Leontodon microcephalus*, *Moehringia fontqueri*, *Narcissus nevadensis*, *Papaver lapeyrosianum*, *Pimpinella procumbens*, *Primula elatior* subsp. *lofthousei*, *Reseda complicata*, *Salix caprea*, *Santolina elegans*, *Sorbus hybrida*, *Taxus baccata*, *Teucrium bicolorum*.

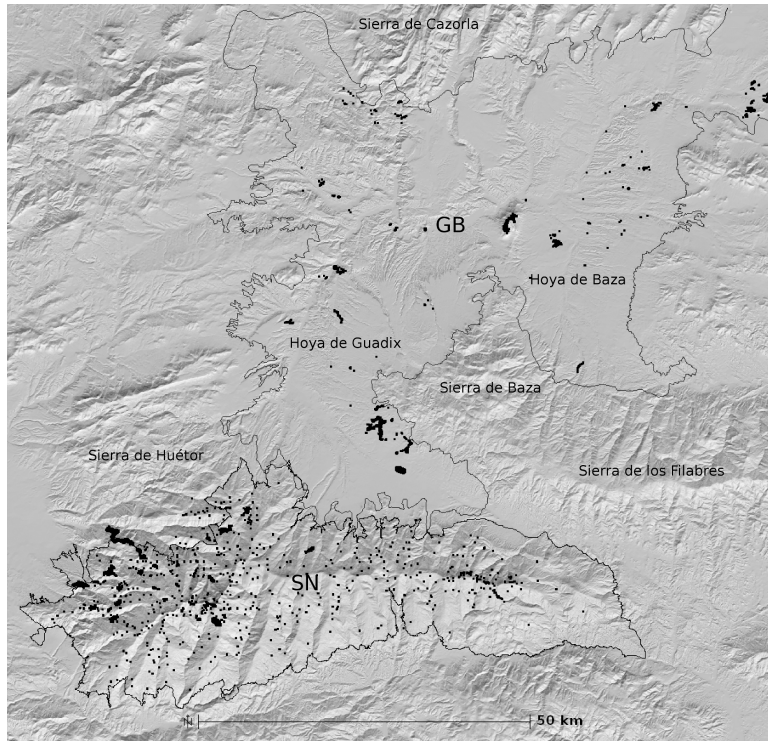
Los mapas digitales de SN presentaron las siguientes características: 25 m de resolución; 1478 filas por 3715 columnas; 5490770 celdas; Límites: N=4123251, S=4086300, E=535800, O=442920.

Durante el resto del capítulo omitiré la abreviatura **subsp.** de las denominaciones de las especies.

### 7.3.2 Análisis de cambio potencial del clima en SN y GB

Las simulaciones de clima proporcionadas por la AEMET (ya descritas en el capítulo anterior) están calculadas según los escenarios del IPCC A2 y B2. El escenario A2 describe un mundo heterogéneo, en el que la autosuficiencia y la conservación de las identidades sociales son características. El crecimiento económico y tecnológico es lento y fragmentado, aunque se fomenta una agricultura más sostenible. La población mundial crece constantemente hasta alcanzar 15000 millones en 2100. El escenario B2 es el resultado de un mundo heterogéneo, con iniciativas locales de sostenibilidad económica, ambiental y social. Las distintas regiones adaptan sus consumos energéticos a los recursos naturales

La **extensión** en MODPLAN es el área del mínimo rectángulo que acoge todas las poblaciones conocidas de la especie.



**Figura 1:** Delimitación de áreas de trabajo y localidades de presencia de especies. Las manchas negras representan localidades de presencia de las especies seleccionadas para este trabajo. Las áreas de modelado están formadas por los respectivos mínimos rectángulos que envuelven ambos polígonos, pero no se muestran para mantener claridad en el mapa.

de los que disponen. La población mundial crece pausadamente hasta alcanzar los 10000 millones a finales de siglo. Las emisiones antropogénicas estimadas de gases de efecto invernadero en cada uno de estos escenarios es uno de los condicionantes del clima simulado a lo largo del siglo XXI.

Para conocer las variaciones potenciales en el clima de las regiones de trabajo, realicé un análisis sencillo de la variación de seis variables climáticas, según los escenarios A2 y B2, utilizando los valores promedio de las áreas delimitadas por los polígonos SN y GB (ver Figura 1) en el año 2000 y en el 2050, para generar una tabla con los valores y un gráfico comparando los cambios en las distintas va-

riables climáticas según hábitats y escenarios. Las variables consideradas fueron precipitación media anual (PA), precipitación media de verano (PV), temperatura media máxima de invierno (TMXI), temperatura media máxima de verano (TMXV), temperatura media mínima de invierno (TMNI) y temperatura media mínima de verano (TMNV).

### 7.3.3 Calibrado de simulaciones

Las simulaciones de cambio en la distribución de las distintas especies en ambos hábitats las generó el flujo de trabajo MODPLAN, según los parámetros de modelado expuestos en la Tabla 7.3.3.

Los parámetros de calibrado del flujo de trabajo fueron diferentes, en algunos aspectos, entre regiones, tal y como muestra la Tabla 7.3.3.

Parámetro	Sierra Nevada	Guadix-Baza
RESOLUCIÓN	25	40
MIN_DIS	50	80
MIN_PRES	15	15
PARTICIÓN	33	33
P_OMISIÓN	5	5

**Tabla 1:** Parámetros de calibrado del flujo de trabajo. Ver Capítulo ??.

Las variables seleccionadas para calibrar los modelos fueron: precipitación media anual, precipitación media de verano, temperatura media mínima de invierno, temperatura media máxima de invierno, temperatura media máxima de verano, orientación este-oeste, exposición topográfica, índice topográfico de humedad, pendiente y radiación solar de invierno. Todas mostraron una correlación entre sí menor a 0.75 (coeficiente de correlación de Pearson).

MODPLAN realizó las simulaciones sobre los escenarios de cambio climático A2 y B2 (según el modelo de circulación global CGCM2), en intervalos de 10 años, desde el año 2000 al año 2050 (ambos incluidos).

*Descarté ECHAM4 porque pronostica valores muy extremos, según los resultados del Capítulo 5.*



Los parámetros de los algoritmos aplicados para calibrar los modelos fueron:

*Seleccioné los parámetros mediante prueba y error usando varias especies bien conocidas como controles.*

- Chebyshev (CH): NearestPoints=8; MaxDistance=0.6.
- Euclidean (EU): NearestPoints=8; MaxDistance=0.5.
- Mahalanobis (MH): NearestPoints=8; MaxDistance=0.1.
- Manhattan (MN): NearestPoints=8; MaxDistance=0.3.
- GARP (GA): TotalRuns=20; MxGenerations=100; PopulationSize=100; Resamples=1000.
- Red Neuronal (NN): Configuración por defecto de OpenModeller.
- Support Vector Machines (VM): Configuración por defecto de OpenModeller.
- MaxEnt (ME): Configuración por defecto de MaxEnt.

#### 7.3.4 Selección de modelos

La selección de modelos tiene como objetivo descartar aquellas especies cuyos modelos no ofrecieron resultados fiables. Para descartarlas apliqué secuencialmente dos criterios:

*En mi opinión, las métricas de evaluación no proporcionan criterio suficiente para rechazar o aceptar un MDE como bueno.*

- **CRITERIO A:** Análisis de la consistencia del modelo, mediante examen visual de los mapas de idoneidad, desviación y FID (fusión de color de idoneidad y desviación, ver Capítulo 6). Conociendo el funcionamiento de los modelos y la lógica que hay detrás de los mapas FID, y tomando como referencia modelos de especies de distribución bien conocida, es posible reconocer modelos que no han capturado una señal de presencia consistente. Se trata de un criterio subjetivo, basado en la experiencia del investigador con los MDE.
- **CRITERIO B:** Opinión de expertos en la distribución de las especies de los distintos hábitats (por los doctores en Botánica Julio Peñas, Juan Lorite y Eva Cañadas).

*El CRITERIO A es un criterio experto que no se basa en el conocimiento de la flora, sino en la experiencia sobre MDE.*

### 7.3.5 *Análisis de efectos del Cambio del Clima en las especies y hábitats*

Una de las posibilidades más realistas para comparar los efectos potenciales del cambio del clima entre hábitats o especies según los resultados de un MDE es analizar la variación de idoneidad del hábitat de los puntos de presencia de las especies en un periodo de tiempo determinado (en este caso, todo el intervalo de simulación). Aquellas especies que presenten mayores diferenciales de idoneidad (un mayor descenso) serán las que probablemente padezcan peores consecuencias (estrés climático, reducción de eficacia biológica, cambios fenológicos, etc.) a causa del cambio del clima.

Para analizar la variación de idoneidad, extraje los valores de idoneidad de las presencias de todas las especies en el año 2000, y en el año 2050 para ambos escenarios. Con los datos resultantes realicé cuatro fases de análisis.

1. Test  $t$  para comparar los valores de idoneidad de los puntos de presencia de cada especie entre los años 2000 y 2050 para ambos escenarios. Las diferencias medias ofrecidas por el test las tabulé para el siguiente análisis.
2. Test de Wilcoxon para muestras pareadas, comparando los valores de diferencial de idoneidad de todas las especies en ambos escenarios, para conocer el efecto global de cada uno de ellos en el conjunto de la flora de SN y GB, como representantes de hábitats mediterráneos.
3. Test de Wilcoxon-Mann-Whitney sobre los valores de diferencial de cada especie en ambos escenarios, tomando el hábitat (SN y GB) como factor de agrupación, para conocer en qué hábitats y escenarios las localidades de presencia de las especies pierden mayor idoneidad.
4. Análisis de regresión múltiple para comprobar la relación entre las propiedades de la muestra (extensión

*En las simulaciones no he tenido en cuenta la tolerancia de las especies al cambio. En el siguiente capítulo trato de incluir este factor en las simulaciones. Los análisis los realicé en R, mediante scripts o R-Commander, según el caso.*

y número de presencias) y los valores de diferencial obtenidos, comparando entre escenarios.

Los resultados permitirán conocer en qué hábitat y escenario las especies estarán sometidas a un mayor riesgo por descenso de idoneidad.

### 7.3.6 Análisis de MODPLAN

#### Funcionamiento del ensamblado

Para valorar el funcionamiento del ensamblado respecto a los algoritmos individuales, analicé los resultados de AUCg de todos los algoritmos y el ensamblado (EC) para todas las especies mediante un test ANOVA y un test de diferencias múltiples de Tukey.

*Seleccioné AUCg para comparar EC con los algoritmos porque su cálculo es similar al de AUC en otros trabajos de modelado.*

*El tiempo de computación total supuso unas 1200 horas (unas 24h por especie) repartidas entre 5 PCs de gama media, con una producción total de 121.5 GB de información digital.*

## 7.4 RESULTADOS

### 7.4.1 Tendencias del clima hasta el 2050

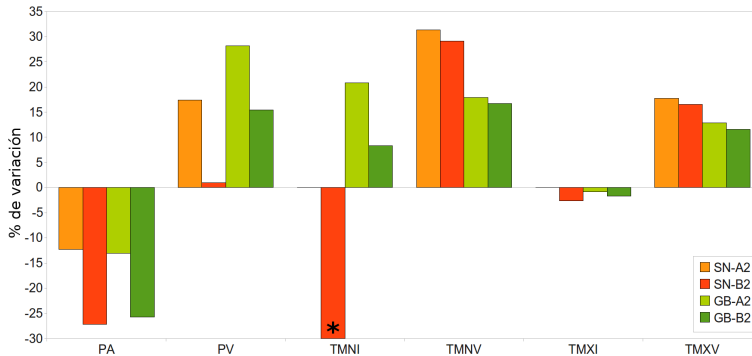
Las tendencias climáticas en ambos hábitats para la primera mitad del siglo XXI quedan reflejadas en la siguiente tabla.

	Sierra Nevada			Guadix-Baza		
	2000	A2-2050	B2-2050	2000	A2-2050	B2-2050
PA	577	506	420	330	287	245
PV	46	54	46	39	50	45
TMNI	1	1	-1	2.4	2.9	2.6
TMNV	13.4	17.6	17.3	16.2	19.1	18.9
TMXI	7.7	77	7.2	11.8	11.7	11.6
TMXV	26	30.6	30.3	31.2	35.2	34.8

**Tabla 2:** Variaciones potenciales del clima en GB y SN según los escenarios A2 y B2 hasta el año 2050. PA = precipitación media anual; PV = precipitación media de verano; TMNI = temperatura media mínima de invierno; TMXI = temperatura media máxima de invierno; TMNV = temperatura media mínima de verano; TMXV = temperatura media máxima de verano.

La Figura 2 muestra los patrones de variación de las distintas variables climáticas para ambos escenarios y hábitats.

Es interesante el patrón común de disminución de la precipitación anual junto al incremento de la precipitación en verano (sobre todo en el escenario A2), lo que indica una redistribución temporal de las precipitaciones en la depresión de Guadix-Baza. El patrón de incremento de las temperaturas en verano (TMNV y TMXV) es sólido entre escenarios y hábitats, y comparativamente mayor en Sierra Nevada.



**Figura 2:** Variaciones de las variables entre hábitats y escenarios para el año 2050. La escala está en porcentaje de variación, tomando como referencia los valores climáticos del intervalo 1971-2000. La barra marcada llega hasta -200, debida a un cambio desde 1 a -1.

#### 7.4.2 Selección de modelos

Según los criterios de selección, algunos modelos no representaron correctamente la distribución de las especies.

Las especies rechazadas según el CRITERIO A fueron *Aconitum burnatii*, *Anthyllis vulneraria arundana*, *Chamaespartium undulatum*, *Cytisus galianoi*, *Dorycnium gracile*, *Erodium boissieri*, *Festuca indigesta indigesta*, *Leontodon microcephalus*, *Limonium majus* y *Primula elatior lofthousei*. El signo de mal ajuste más común observado fue un exceso de áreas de alta idoneidad con gran desviación, y pocas áreas de alta idoneidad y baja desviación.

Según el CRITERIO B, además de algunas de las anteriores, las especies rechazadas fueron *Carex camposii*, *Euphrasia willkommii* y *Teucrium bicoloreum*. El motivo de rechazo fue una mala representación de la distribución conocida de las especies, por sobrepredicción.

#### 7.4.3 Efectos del Cambio del Clima en las especies y hábitats

El test t mostró diferencias significativas en las variaciones de promedio de idoneidad del hábitat de los puntos de presencia entre el año 2000 y el año 2050 para todas las especies y escenarios. Las diferencias de idoneidad oscilaron entre los 13.62 de *Astragalus oxyglottis* en B2 hasta los 63.64 de *Krascheninnikovia ceratioides* en A2 (ver Figura 3).

Realmente los valores de diferencial habría que interpretarlos como pérdida de idoneidad, con un signo negativo delante: -13.62 y -63.64 respectivamente.

Según los resultados del test de Wilcoxon para muestras pareadas, ambos escenarios no mostraron diferencias significativas entre hábitats en cuanto a los valores de diferencial de idoneidad ( $V = 392$ ;  $p = 0.3624$ ).

Según el test de Wilcoxon-Mann-Whitney (A2:  $W = 203$ ,  $p = 0.0486$ ; B2:  $W = 119$ ,  $p = 0.4164$ ) en el escenario A2 hay diferencias significativas entre ambos hábitats, con una situación desfavorable para las especies de GB. En el escenario B2 la tendencia podría invertirse, aunque los resultados no son estadísticamente significativos (ver Figura 4).

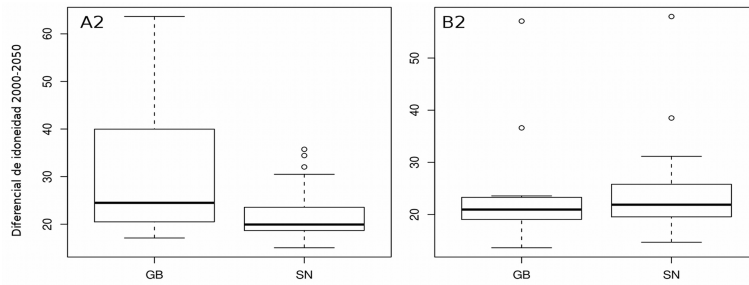
El análisis de regresión múltiple reveló que las propiedades de la muestra (realicé una transformación logarítmica para normalizar los datos) pueden explicar entre un 23 % (A2,  $R^2 = 0.2316$ ,  $F = 4.97$   $p=0.0129$ ) y un 27 % (B2,  $R^2 = 0.2748$ ,  $F = 6.25$   $p=0.0049$ ) de la variabilidad de los valores de diferencial de idoneidad. En ambos escenarios la influencia del número de presencias fue significativa ( $p < 0.0010$ ), y solo en B2 la influencia de la extensión tuvo importancia marginalmente significativa ( $p = 0.0558$ ). La Figura 5 muestra el efecto de cada propiedad de la muestra sobre el diferencial de idoneidad en el escenario B2 (A2 se omite por no mostrar diferencias importantes).



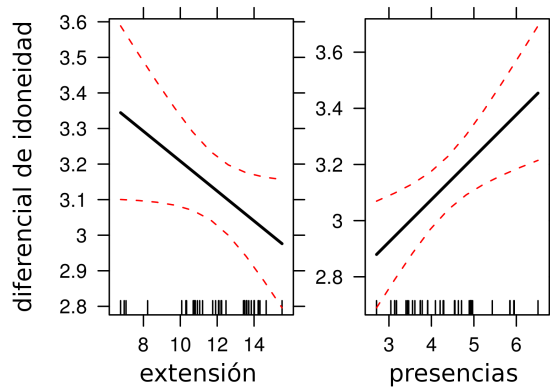
**Figura 3:** Diferencial de idoneidad de las poblaciones conocidas de todas las especies según ambos escenarios. Las especies marcadas (\*) fueron rechazadas en el proceso de evaluación.

#### 7.4.4 Funcionamiento del ensamblado

El análisis ANOVA de los valores AUCg entre algoritmos para todas las especies reveló diferencias muy significati-



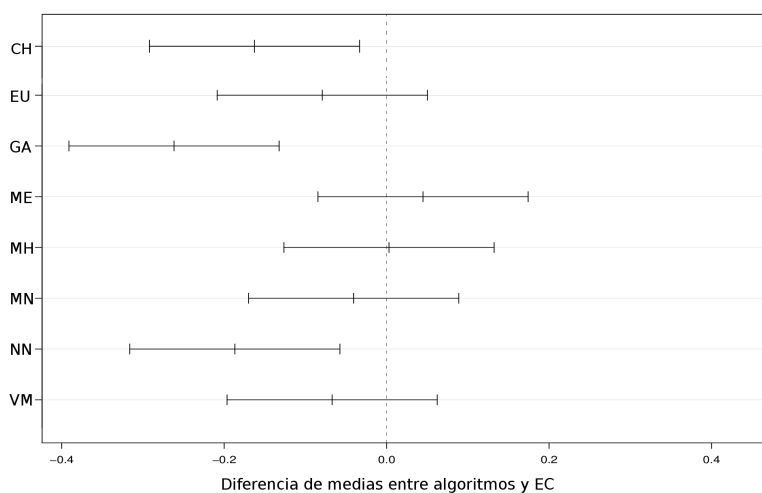
**Figura 4:** Comparación de diferencial de idoneidad entre hábitats para los distintos escenarios.



**Figura 5:** Efectos de la extensión de presencia y el número de presencias sobre el diferencial de idoneidad para el escenario B2. Los valores son transformaciones logarítmicas de las variables originales.

vas ( $F=11.88$ ;  $p=2.34 \cdot 10^{-15}$ ) entre los distintos métodos de modelado. El test de Tukey (ver Figura 6) confirmó que solo dos algoritmos proporcionan resultados significativamente menos precisos que los del ensamblado (GA,  $p=0$ ; NN,  $p=0.0003$ ). De los que proporcionan resultados similares, solo ME obtuvo un promedio de AUCg ligeramente superior.

En la Tabla 3 puede observarse que EC solo ha obtenido mejor valor AUCg que los algoritmos individuales en 3 especies (*Acer opalus granatense*, *Hippocrepis prostrata* y *Sorbus hybrida*), quedando muy por detrás de ME y MH.



**Figura 6:** Resultados gráficos del test de Tukey. 0.0 representa el valor promedio del ensamblado (EC). La amplitud de las líneas que representan a cada algoritmo se corresponden con un intervalo de confianza para la media del 95%.

algoritmos	CH	EU	GA	ME	MH	MN	NN	VM	EC
<b>primero</b>	0	1	0	21	20	1	0	3	3
<b>último</b>	10	0	23	0	2	0	11	3	0

**Tabla 3:** Número de especies en las que el algoritmo quedó primero o último según sus valores de **AUCg**

## 7.5 DISCUSIÓN

### 7.5.1 Efectos del Cambio del Clima en las especies y hábitats

Los pronósticos de cambio de clima de escenarios A2 y B2, según los resultados obtenidos para GB y SN, no pueden calificarse como mejores o peores, porque se diferencian en cuanto al comportamiento de la temperatura y la precipitación. A2 propone un futuro más cálido (+2.01°C, media anual para GB y SN) con precipitaciones algo menores (-57 mm/año, media de SN y GB) pero más distribuidas a lo lar-



go del año, y B2 propone un futuro menos cálido (+1.54°C) pero significativamente más árido (-121 mm/año).

Esta diferencia de comportamiento entre escenarios según la variable considerada provoca que el efecto global (tomando todas las especies excepto las descartadas) en la flora sea el mismo en cuanto a pérdida de idoneidad. Entre hábitats la tendencia mostrada por las simulaciones según el escenario A2 se aleja de lo que cabría esperar, con un mayor descenso de la idoneidad en las especies de GB. Esta situación se debe a varios factores actuando en conjunto, pero analicemos dos que considero más importantes:

1. **El incremento de la precipitación en verano.** Los MDE capturan los requerimientos ecológicos de la especie según los datos disponibles, y cuando las condiciones cambian, no pueden valorar si estos cambios suponen o no un beneficio para la especie, y solo miden cuanto se alejan las nuevas condiciones de las anteriormente conocidas. Por este motivo, un incremento de la precipitación en verano, que tal vez favorecería a algunas especies, es percibido por el MDE como un alejamiento de las condiciones apropiadas. Esta circunstancia afecta en menor medida a algoritmos de modelado que calculan curvas de respuesta, como MaxEnt, GLM o GAM, que tienen cierta capacidad de extrapolar valores de idoneidad del hábitat sobre nuevas condiciones.
2. **La concentración de los puntos de presencia de la especie.** Algunas especies de GB y SN presentan la mayoría de sus puntos de presencia concentrados en un solo polígono, o en unos pocos muy cercanos entre sí, y tal vez alguna observación más alejada. Es el caso de *Centaurea saxifraga*, *Haplophyllum bastetanum*, *Krascheninnikovia ceratioides*, *Papaver lapeyrosianum*, *Plantago maritima serpentina* y *Senecio auricula*, entre otras. Si en distintos puntos dentro del polígono (en pequeños polígonos adyacentes) las condiciones ecológicas son muy similares (algo esperable en cuanto a condiciones climáticas, aunque no necesariamente en las topográficas), las condiciones apropiadas que el MDE

captura para la especie ocupan un pequeño volumen  $n$ -dimensional. Si este volumen queda cerca, en el espacio ecológico, de alguno o varios extremos de algunas de sus dimensiones, y en el siguiente periodo de simulación algunos de estos extremos se recortan, desaparecen de golpe los valores apropiados de varias dimensiones para la especie, originando una bajada importante de la idoneidad. En el caso de un mal muestreo de la presencia de la especie, este efecto es un artefacto debido a la mala calidad de los datos. Si el muestreo es bueno, y la especie presenta una distribución restringida por motivos naturales, el efecto explicado podría estar pronosticando una dramática realidad.

De estas consideraciones puede extraerse una norma general de funcionamiento de las proyecciones de cambio de distribución geográfica futura basadas en los MDE: cuanto más concentrados están los registros de presencia de una especie en los espacios geográfico y ecológico, mayor será el descenso promedio de idoneidad del hábitat pronosticado para el futuro por la simulación, independientemente de lo que ocurra en el resto del territorio.

Otra cuestión, que está sin resolver, es la verdadera influencia del descenso de idoneidad del hábitat sobre las probabilidades de supervivencia de las poblaciones. No hay, por ahora, ningún estudio que relacione un gradiente de idoneidad del hábitat con alguna medida de eficacia biológica o de decaimiento de poblaciones, por lo que no pueden relacionarse los valores de diferencial de idoneidad con probabilidades de supervivencia.

Estos valores deben tomarse como una medida relativa de cuanto cambian las condiciones del hábitat de la especie. Incluso desde este punto de vista, para algunas especies puede resultar conflictiva esta medida, como en aquellas que tienen rangos geográficos mayores que las áreas de estudio. Por ejemplo, *Astragalus oxyglottis* está presente en otros lugares del CE y NE de España y O de Asia. Es posible que su rango ecológico sea más amplio que el muestreado en GB, y probablemente esta especie puede tener potencial adaptativo para

desarrollarse en condiciones que no aparecen actualmente en el área de trabajo. En la misma circunstancia están en GB *Boreava aptera*, *Cynomorium coccineum*, *Dorycnium gracile*, *Gypsophila tomentosa*, *Krascheninnikovia ceratoides*, *Loeflingia baetica*, *Plantago maritima serpentina*, *Senecio auricula*. En SN, aunque algunas de las especies de trabajo son endémicas, otras están presentes en otros lugares, como *Acer opalus granatense*, *Anthyllis vulneraria arundana*, *Betula pendula fontqueri*, *Dryopteris submontana*, *Primula elatior lofthousei*, *Salix caprea*, *Sorbus hybrida* y *Taxus baccata*.

No he incluido en esta lista las especies endémicas del sureste peninsular.

Una aproximación sencilla para estudiar la posible incidencia de este efecto en los MDE consiste en comparar las características climáticas de todas las localidades conocidas (a escala global) de la especie con las de las localidades de la zona de trabajo. La posición y la amplitud relativa de las características climáticas de las presencias locales respecto a las globales determinará si el MDE de la especie puede ser útil o no para proyectar su distribución sobre escenarios de cambio climático.

Otra posibilidad consiste en generar las simulaciones con todos los registros de presencia conocidos (obtenidos de una base de datos distribuida como GBIF <http://data.gbif.org>) y variables climáticas de todo el planeta a alta resolución (Microclima <http://www.worldclim.org/>, por ejemplo), pero se pierde el detalle local, tan importante para análisis a escala de paisaje y acciones de conservación y gestión de la Biodiversidad.

### 7.5.2 Funcionamiento del ensamblado

Los resultados indican que el ensamblado de modelos (EC) ocupa la tercera posición en cuanto a ajuste con los datos de evaluación si se compara con el resto de algoritmos. Aunque solo ha sido el mejor en tres ocasiones, en ningún caso ha proporcionado el peor resultado, que normalmente ha estado ocupado por GA o NN. El ensamblado no siempre proporciona mejores resultados que los modelos individuales, pero este comportamiento concuerda con el descrito por otros autores (Marmion *et al.*, 2009).

En la comparación global de algoritmos, GA es el que peores registros ha obtenido. Según [Peterson et al. \(2008\)](#) los valores de idoneidad de un modelo generado con GARP se concentran en la zona superior de la curva ROC, mientras que los resultados de otros algoritmos cubren todas las regiones de la curva, por lo que GARP queda en desventaja al usar la curva ROC al completo como criterio de evaluación. La cuestión planteada por [Peterson et al. \(2008\)](#) no deja de ser interesante, pero probablemente la parametrización del algoritmo tenga importancia en este caso. GARP consume más o menos tiempo según algunos de sus parámetros, que en la configuración por defecto de OpenModeller están preparados para minimizar el tiempo de cálculo. Una parametrización específica podría mejorar los resultados de GARP, a costa de ralentizar todo el proceso de simulación significativamente.

Probablemente el caso de NN es similar al de GARP, porque no he modificado sus parámetros de ejecución. Encontrar los mejores parámetros para un número amplio de algoritmos es un proceso largo y tedioso, que requiere multitud de pruebas con distintas especies, y es una cuestión que poco o nada se ha tocado en la literatura, excepto para algún algoritmo [Phillips y Dudík \(2008\)](#). Para este trabajo realicé varios experimentos previos con dos especies de distribución bien conocida por nuestro grupo de trabajo (*Linaria nigricans* y *Moehringia fontqueri*) con el objetivo de mejorar en lo posible la parametrización de los algoritmos, pero modificar y probar parámetros de NN, GA y VM hubiera llevado mucho tiempo, por lo que solo se cambiaron las configuraciones de los algoritmos de distancias.

La verdadera fortaleza de usar el ensamblado de algoritmos individuales, además de en los resultados robustos del promedio de idoneidad, está en el análisis de la desviación estándar, que proporciona una visión clara de la consistencia de las predicciones, sobre todo cuando se combina con la idoneidad en un mapa FID.

Es cierto que los algoritmos que peor están funcionando (GA y NN) podrían estar afectando negativamente al funcionamiento global del ensamblado, incrementando el valor de la desviación estándar. Tal vez una parametrización más

precisa de ambos algoritmos, o su sustitución por otros con mejor desempeño mejore los resultados globales, pero comprobar este punto no está, por motivos de tiempo, dentro de los objetivos de esta tesis.

### 7.5.3 *Limitaciones técnicas del flujo de trabajo y posibles mejoras*

Aunque en este capítulo he utilizado solo una parte de la información que genera MODPLAN, el trabajo con los resultados me permite hacer una crítica global.

La limitación más importante de MODPLAN es la no consideración del potencial dispersivo de las especies en las simulaciones. Esta cuestión no tiene gran incidencia en los resultados de este trabajo, porque solo he tratado el descenso de idoneidad en las poblaciones conocidas, pero sí la tiene si se pretende interpretar los resultados de las simulaciones fuera de estas localidades.

Los procesos dispersivos tienen una gran importancia para aportar consistencia a las proyecciones de cambio de distribución, pero solo muy recientemente han sido incluidos en los procesos de modelado mediante autómatas celulares (Engler *et al.*, 2009). Para implementar este factor son necesarios datos muy precisos sobre las mecánicas dispersivas de las especies, y generalmente estos datos son escasos o inexistentes. En este sentido resulta útil la revisión de Vittoz y Engler (2007), en la que se asignan distancias de dispersión medias y máximas a distintos mecanismos de dispersión. En ausencia de esta información, las simulaciones de distribución futura asumen automáticamente una dispersión ilimitada. Este escenario solo es factible mediante una intervención humana intensa para la conservación de las especies, transportando ejemplares o semillas a los lugares en los que aparecerán condiciones apropiadas para las especies de interés. Esta perspectiva tiene un trasfondo ético que no pretendo abordar, aunque es el momento de empezar a tratarlo con seriedad.

Otro factor que las simulaciones no tienen en cuenta es la tolerancia potencial de las especies a un cambio en las con-

diciones climáticas. Al trabajar con las simulaciones surgen preguntas como: ¿Puede *Papaver lapeyrosianum* soportar un incremento de 2°C de temperatura media anual?; ¿puede *Astragalus oxyglottis* tolerar un descenso de 85 mm en la precipitación media anual?. Afrontar estas preguntas exige dos perspectivas (combinadas o no). Por una parte, la opinión experta y, por otra, la perspectiva experimental, ya sea en ambientes controlados (ensayos de germinación, por ejemplo), o translocaciones de individuos son algunas de las posibilidades. Conocer la tolerancia de las especies al cambio del clima es vital para simular correctamente cambios en la distribución de plantas debidos al cambio del clima, pero de nuevo estos datos son escasos o inexistentes.

Finalmente, sobre las métricas de evaluación implementadas en MODPLAN (ver Capítulo 6). AUCg fue útil para comparar el ajuste obtenido por los distintos algoritmos y el ensamblado, ya que dicha comparación implicaba modelos de la misma especie y misma extensión, único ámbito de aplicabilidad real de los resultados de AUC (Lobo *et al.*, 2008). Pero durante la evaluación experta de los modelos, los valores de las métricas de evaluación no aportaron ningún criterio satisfactorio para apoyar las decisiones sobre inclusión o no de especies en el análisis. Y es que AUCg y AUCl se mostraron sensibles (especialmente AUCg) a las propiedades de la muestra (número de presencias y extensión). El criterio CIDp tampoco ofreció una respuesta satisfactoria, pero abrió las puertas a una nueva posibilidad, la evaluación del modelo por sus propiedades: valores de idoneidad y desviación sobre los puntos de presencia de la especie y sobre  $n$  puntos aleatorios. Esta posibilidad se desarrolla brevemente en el ANEXO\_EVALUACIÓN, en el que analizo a fondo la relación de las métricas con las propiedades de la muestra, y de los puntos de presencia sobre los modelos, y a partir de los resultados desarrollo un criterio más simple para evaluar modelos sin partición de datos y calibrado previo.

## 7.6 CONCLUSIONES

1. Hasta el año 2050 en SN y GB las simulaciones climáticas indican que pueden incrementarse las temperaturas máximas y mínimas de verano significativamente (+4°C) y las disminuir las precipitaciones anuales (-121 mm). En GB se aprecia un patrón de redistribución temporal de las precipitaciones hacia el verano.
2. Todas las especies de trabajo mostraron descensos de idoneidad durante el periodo simulado (2000-2050).
3. Las localidades de presencia de la flora de GB presentan un mayor descenso de idoneidad que las de la flora de SN. En el escenario B2 la tendencia se invierte, pero no es estadísticamente significativa.
4. El diferencial de idoneidad de una especie está relacionado con su extensión de presencia y número de registros de presencia.
5. El ensamblado de modelos mediante promedio funcionó mejor que 6 de los 8 algoritmos individuales.
6. Para mejorar MODPLAN es necesario implementar las siguientes características:
  - Incorporación de la migración potencial de las especies en los modelos.
  - Incorporación de la tolerancia potencial de las especies al cambio.
  - Eliminación de la fase de partición de datos, calibrado y evaluación de los modelos, y sustitución por una medida basada en las propiedades del modelo.

*En el próximo capítulo se tratan de implementar estas mejoras en MODPLAN.*

## IMPACTO POTENCIAL DEL CAMBIO CLIMÁTICO EN LA DIVERSIDAD VEGETAL

---

### 8.1 INTRODUCCIÓN

En el Capítulo 7 he utilizado MODPLAN para tratar de conocer el posible efecto del Cambio Climático en las poblaciones conocidas de las especies seleccionadas, estudiando sus variaciones de idoneidad. Pero los MDE no solo ofrecen información valiosa sobre esta variable en las poblaciones conocidas, sino que también (y ahí está la mayor de sus posibilidades) proporcionan valores de idoneidad para todo el territorio.

El transcurrir del tiempo y el cambio en el espacio geográfico de los valores de las variables climáticas ocasiona una dinámica de expansión y/o contracción de las áreas de hábitat idóneo, que obliga a las especies a migrar para mantenerse dentro de su rango climático. En adelante, a las áreas de expansión (donde domina el proceso de migración) del hábitat las denominaré *frentes de avance* y a las de contracción (donde desaparecen individuos gradualmente) las llamaré *frentes de retroceso*.

*Esta nomenclatura ya apareció en el Capítulo 5.*

#### 8.1.1 *En el frente de avance*

Las evidencias paleoecológicas y los estudios observacionales sugieren que la migración es la respuesta más habitual de las especies ante un cambio climático (Davis y Shaw, 2001; Parmesan, 2006). En el pasado las velocidades de migración pudieron estar en torno a 1000 m/año (Clark *et al.*, 1998), pero a causa del rápido cambio climático esperado durante el presente siglo, las necesidades en velocidad de migración para seguir las isothermas apropiadas deberían estar entre 3000 y 5000 m/año (Davis y Shaw, 2001).



Según *Thuiller et al. (2008)*, la migración es el resultado de cuatro procesos: fecundidad, dispersión, reclutamiento y crecimiento poblacional. Algunos estudios confirman que las limitaciones dispersivas son más importantes que las limitaciones de reclutamiento, al menos en las especies forestales (*Brunet et al., 2000; Verheyen y Hermy, 2001; Honnay et al., 2002*). Pero la migración también depende de las características del paisaje, como las barreras geográficas, o la fragmentación geógena del hábitat. En la actualidad, al contrario que en el pasado, la actividad humana ha fragmentado severamente los hábitats (fragmentación antropógena), imponiendo nuevas barreras que las especies no pueden superar (*Walther, 2004*).

*La actividad humana también permite a algunas especies colonizar territorios que han sido inaccesibles según sus propios medios de dispersión.*

Teniendo en cuenta estos factores, una predicción precisa de la distribución futura de las especies requiere un conocimiento detallado de la habilidad migratoria de las especies (algo que los MDE no tienen en cuenta) y de la fragmentación del hábitat. En los trabajos de modelado y pronóstico de efectos del cambio climático en la distribución de las especies, generalmente se asume que aquellas con gran capacidad dispersiva ocuparán todo el espacio idóneo futuro (dispersión ilimitada), mientras que las de escasa capacidad dispersiva solo ocuparán aquellas áreas de su distribución actual que permanezcan idóneas en el futuro (*Peterson et al., 2001*). Pero, si debido a las características del paisaje las especies no tienen posibilidad de migrar una velocidad suficiente, los resultados de las simulaciones de distribución futura son erróneos.

A pesar de la importancia de la migración como respuesta al cambio del clima (*Higgins y Cain, 2003*), son pocos los estudios de simulación que incluyen los procesos migratorios (algunos son *Lischke et al. (2006); Engler et al. (2009)*). Y esta circunstancia tiene una doble explicación: la falta de datos de capacidad migratoria de una gran mayoría de las especies, y la escasez de programas informáticos que permitan simular la dispersión sobre modelos de distribución.

*La escasez de información sobre la mecánica dispersiva de las especies ha sido un problema importante durante la elaboración de este trabajo.*

En una reciente revisión, *Vittoz y Engler (2007)* han llevado a cabo una completa revisión bibliográfica de las distancias dispersivas (medias, máximas y mínimas) asociadas a cada tipo de dispersión. A partir de su trabajo, es posible

asignar distancias de dispersión a un amplio conjunto de especies solo conociendo su tipo de dispersión.

En cuanto a programas capaces de simular el proceso de dispersión sobre los MDE, recientemente Engler *et al.* (2009) han aplicado TREEMIG (Lischke *et al.*, 2006), que pasa por ser el simulador de dispersión que incluye la representación más completa de la dispersión de semillas y procesos de regeneración, incluyendo antagonismo densidad-dependiente y competición intra e interespecífica (Thuiller *et al.*, 2008).

### 8.1.2 En el frente de retroceso

El frente de retroceso es el gran olvidado (o tal vez solo sobreentendido) en los trabajos de modelado. En teoría, los frentes de retroceso se sitúan en aquellos márgenes de grandes poblaciones en los que, por una situación geográfica desventajosa, las condiciones apropiadas para la especie comienzan a empeorar antes que en el resto de la población. Según Davis y Shaw (2001), el inicio del proceso de extinción de las especies podría comenzar en estos lugares, especialmente en los periodos de rápido cambio climático.

Hampe y Petit (2005) afirman que el comportamiento de una población en el frente de retroceso puede oscilar entre dos extremos: la desaparición de los individuos del frente y la persistencia. Estos autores reservan el término *frente de retroceso* (*trailing edge*) para la primera de las posibilidades y *frente estable* para la segunda. En este trabajo considero ambos tipos de frentes una misma cosa, cuya percepción solo depende de la escala temporal considerada.

*Porque, a largo plazo,  
no hay nada estable...*

En el frente de retroceso pueden darse distintas circunstancias según la especie. Imaginemos una planta anual, cuyo umbral de temperatura máxima de germinación está en 20°C. Según se va alcanzando esta temperatura año a año en el frente de retroceso, el porcentaje de semillas que germinan es progresivamente menor, por lo que el número de individuos reproductores también decrece. Cuando finalmente la temperatura supera el umbral, independientemente de la potencia del banco de semillas, la población total en la zona esa temporada es cero. Como respuesta compensato-

En plantas anuales no hay estudios que atestigüen esta posibilidad, y los pocos que podrían parecerse, solo se refieren a plantas perennes (Hampe y Petit, 2005).

ria la planta podría adelantar su fenología, pero dependerá del rango diario de temperatura que puedan soportar las plántulas. Siendo una planta de corto tiempo de generación, también podría adaptarse, pero en las condiciones actuales de rápido cambio la probabilidad debería ser baja.

En el otro extremo están las especies con un largo ciclo de vida, que tienen mayores posibilidades de persistencia, al menos a escalas de tiempo pequeñas (Thuiller *et al.*, 2008). Una especie arbórea puede ver disminuida su capacidad reproductiva, o inhibida la germinación de sus semillas ante cambios bruscos en disponibilidad de agua y temperatura, pero sobrevivirá un tiempo indeterminado (hasta varios años). Un ejemplo de esta situación podría ser el de *Pinus sylvestris nevadensis* (Castro, 2000) en Sierra Nevada y la Sierra de Baza.

Los MDE no tienen en cuenta la posibilidad de persistencia, y pronostican la extinción inmediata de las poblaciones según desaparece su hábitat idóneo. Y esta cuestión es técnicamente sencilla de implementar en los programas de modelado, por lo que el problema solo viene de los datos disponibles sobre tolerancia de las especies a cambios en su hábitat, que, cuando están presentes, suelen ser más cualitativos que cuantitativos, y en todo caso, muy difíciles de relacionar con los valores de idoneidad del hábitat que proporcionan los MDE.

### 8.1.3 Efectos del cambio del clima en la biodiversidad

La dinámica entre los frentes de avance y retroceso, las características migratorias de las especies y la configuración del territorio (tanto geógena como antropógena) determinan el desplazamiento de las poblaciones de las especies dentro de una región. Tomando el escenario descrito desde una perspectiva temporal, las especies de una región forman un mosaico fluido que se mueve al compás marcado por los cambios ambientales, en el que podríamos observar la aparición, expansión, contracción y desaparición de poblaciones de distintas especies. Ésta es la expresión espacio-temporal de la biodiversidad, con dos niveles de apreciación: la escala

regional, en la que priman las variaciones en el patrón geográfico de la riqueza, y la escala local, donde la propiedad más interesante (desde la perspectiva de este trabajo) es la variación en el número total y la composición de especies a lo largo del tiempo.

## 8.2 OBJETIVOS

El objetivo principal de este capítulo es simular el comportamiento de la diversidad vegetal de Andalucía en dos escenarios de cambio climático (A2 y B2) para el periodo 2000-2100 a intervalos de 10 años, teniendo en cuenta las capacidades de migración y tolerancia de las especies y la fragmentación antropógena del hábitat.

Con los resultados de las simulaciones trataré de dar respuesta a varias cuestiones:

- ¿Cuántas especies podrían extinguirse?.
- ¿Cuáles son las tendencias previstas en área de ocupación potencial para las distintas especies modeladas?.
- ¿Cómo van a cambiar los patrones geográficos de la riqueza de especies, y cómo afectarán estos cambios a la flora andaluza en general y a la de los Espacios Naturales Protegidos en particular?.
- ¿Cuál es la relación que existe entre la variación climática y la tasa de extinción?, ¿puede predecirse?.

## 8.3 MATERIAL Y MÉTODOS

### 8.3.1 Área de trabajo

Andalucía ocupa una posición geográfica estratégica como puente entre dos continentes, que explica su gran diversidad vegetal, resultado de las distintas corrientes migratorias que la han atravesado y los refugios glaciares que se han constituido en su territorio. Su configuración geográfica y su amplia extensión (87597 km<sup>2</sup>), junto a una elevada geodiversidad (Rivas-Martínez *et al.*, 1997), propician la aparición de

*En este capítulo no trataré la composición de especies, ni propiedades individuales de los taxa, como la rareza o el estado de amenaza, aunque son cuestiones de gran interés.*

distintos tipos de hábitats que cobijan más de 4000 taxones de flora (Hernández-Bermejo y Clemente, 1994), de los que unos 516 son endémicos (Melendo *et al.*, 2003).

He seleccionado como límites del área de trabajo los de la Comunidad Autónoma Andaluza porque constituye una unidad de gestión ambiental sobre la que se aplica la misma política de protección de la biodiversidad. La representación digital de Andalucía utilizada para las simulaciones tiene las siguientes características: 100 m de resolución; 3022 filas por 5324 columnas; 16089128 celdas totales; 8760194 celdas de trabajo (el resto pertenece al mar y otras comunidades); Límites (UTM 30N ED-50): N=4289200, S=3987000, E=620400, O=88000.

*Como en el capítulo anterior, descarté los escenarios calculados según ECHAM4 por considerarlos muy poco probables.*

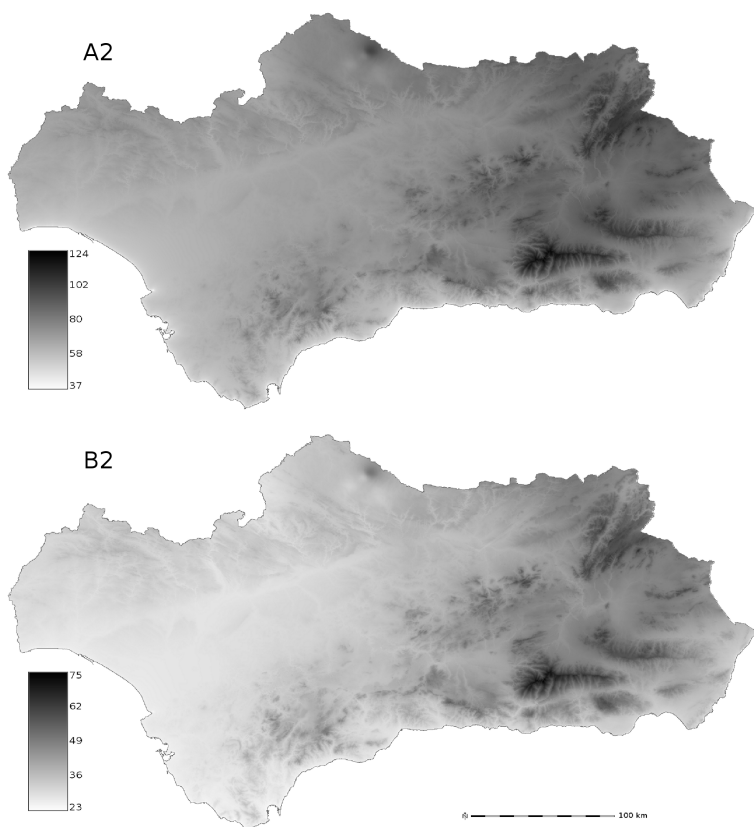
Según las simulaciones climáticas disponibles (escenarios A2 y B2 según el modelo de circulación global CGCM2), en Andalucía a final del siglo XXI la temperatura media máxima de verano podría incrementarse (desde los 31.7°C entre 1971 y 2000) en +3.7 y +7.4°C, según los escenarios B2 y A2 respectivamente. La variación esperada para las temperaturas medias mínimas de invierno (partiendo de 4.3°C) supondría un incremento entre +1.1 y +1.9°C. La Figura 1 muestra como los mayores incrementos de temperatura se esperan en las zonas montañosas de Andalucía Oriental.

La precipitación media anual (537mm) podría disminuir entre -97 y -183mm (según B2 y A2), con una cierta redistribución por incremento en las precipitaciones de verano (37mm) entre +14 y +16mm. La distribución geográfica de las precipitaciones en el futuro (ver Figura 2) podría homogeneizar las diferencias entre territorios, descendiendo en las zonas montañosas de la costa, y manteniéndose o incrementándose puntualmente en las zonas áridas del este de la comunidad.

### 8.3.2 *Especies de trabajo*

Quiero agradecer a Mario Ruiz y Rosa Fernández la aportación de sus propios registros.

La información de presencia de las especies procede del Mapa de Usos y Coberturas del año 1999 (escala 1:50000), del



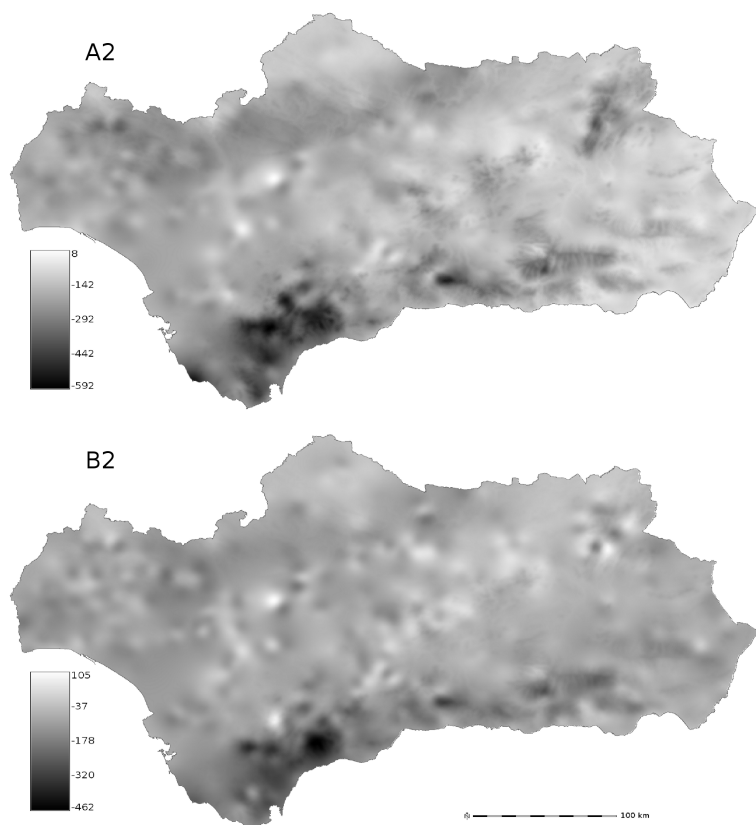
**Figura 1:** Variación de temperatura media máxima de verano (en décimas de grado) entre el año 2000 y el 2100 para ambos escenarios.

proyecto FAME (Flora Amenazada de Andalucía) y de y de trabajos de campo propios y de otros investigadores. Las dos primeras fuentes ofrecen información poligonal, mientras que las recolecciones propias constan de registros puntuales obtenidos mediante GPS. A partir de las tres fuentes de datos obtuve información de presencia para 514 especies.

### 8.3.3 Modificaciones a MODPLAN

En el Capítulo ?? llegué a algunas conclusiones sobre las mejoras que podrían introducirse en MODPLAN para mejorar

*Más detalles sobre las fuentes de información de presencia de las especies en el Capítulo ??.*



**Figura 2:** Diferencias de precipitación media anual entre el año 2100 y el 2000 para ambos escenarios.

las simulaciones. A continuación expongo las modificaciones más importantes efectuadas.

#### *Eliminación de la fase de evaluación*

Durante el desarrollo de los trabajos que describo en el Capítulo ?? llegué a la conclusión de que el proceso previo de partición de datos, calibrado del modelo con los puntos de calibrado, y evaluación mediante los puntos de evaluación aplicando las métricas AUGg, AUCl y CIDp no proporcionaba ninguna ventaja, porque algunas métricas son sensibles a las propiedades de la muestra, y no proporcionan ninguna ventaja como criterio para seleccionar modelos (esta idea se

desarrolla en el Anexo ??). Por tanto, eliminé de MODPLAN la fase de evaluación, sustituyéndola por un análisis gráfico de los valores de idoneidad y desviación de los puntos de presencia y 10000 puntos aleatorios, que ofrece una rápida visión del poder de discriminación del modelo.

### *Migración y la tolerancia*

Para las especies que obtuvieron más de 30 puntos de presencia en el flujo de trabajo PRP, realizamos una búsqueda bibliográfica tratando de constatar, o en su defecto, inferir, su tipo de dispersión. Para asignar las distancias de dispersión a cada especie apliqué los criterios descritos en **Vittoz y Engler (2007)**, que identifican 7 tipos dispersivos. Tratando de aplicar un criterio conservador, tomé como distancia de dispersión máxima para cada especie aquella dentro de la que caen el 50% de las semillas (ver Tabla 1 en **Vittoz y Engler (2007)**). Las distancias las transformé de metros a celdas por década (siendo las celdas de 100 x 100 metros), para adaptarlas a la resolución espacial y temporal de las simulaciones. Para compensar el tiempo de generación de las especies arbóreas y poder obtener una estimación de la migración dividí sus distancias de dispersión por dos. De este modo, las distancias de migración potencial obtenidas, en celdas por decenio, para cada tipo dispersivo de los considerados por **Vittoz y Engler (2007)** fueron:

- **Tipos 1, 2 y 3:** 100 metros/10 años, para especies sin adaptaciones aparentes para la dispersión o con pequeñas distancias de dispersión (mirmecoria, epizoocoria y anemocoria en herbáceas).
- **Tipo 4:** 400 metros/10 años para especies anemocoria en especies arbóreas y diszoocoria (dispersión por pequeños animales).
- **Tipo 5:** 1000 metros/10 años para anemocoria con plumas eficientes.
- **Tipo 6:** 2000 metros/10 años para especies arbóreas y 4000 metros/10 años para arbustivas dispersadas mediante endozoocoria o epizoocoria (aves y mamíferos).

*Este método de evaluación no es perfecto, pero facilita la interpretación del modelo.*

*En la revisión sobre métodos de dispersión recibí la ayuda inestimable de los doctores Juan Lorite y José Ángel Sánchez.*



Se trata de una simplificación grosera del proceso de migración, que sólo tiene en cuenta el factor dispersión, pasando por alto fecundidad, reclutamiento y crecimiento poblacional.

Para calcular la tolerancia de las especies al cambio en las condiciones de su hábitat, ante la falta de datos, desarrollé una escala arbitraria de valores de tolerancia según el biotipo de la especie y el tamaño medio de los individuos (altura). La escala asigna a cada biotipo y a cada rango de tamaño un valor de tolerancia entre 1 y 5, cuya suma ofrece el valor de tolerancia final de cada especie.

Los biotipos y sus valores de tolerancia asignados fueron: leñosas perennifolias (5), leñosas caducifolias (4), herbáceas perennifolias (3) y herbáceas anuales (2).

Los rangos de tamaño y valores asignados fueron: más de 10 metros (5), entre 5 y 10 metros (4), entre 1 y 5 metros (3) y menos de 1 metro (2).

Según esta escala, las especies más resistentes (arbóreas de gran tamaño) podrán soportar una idoneidad del hábitat 10 puntos inferior al valor mínimo de idoneidad del conjunto de puntos de presencia. Un caso como ejemplo: si el menor valor de idoneidad para un punto de presencia de una especie arbórea perennifolia es 70, en cualquier celda del territorio con presencia de la especie en la que aparezcan valores de idoneidad iguales o menores a 60 desaparecerá la población que la ocupa. Estas unidades de tolerancia son valores de idoneidad del hábitat, por lo que para la especie del ejemplo el hábitat puede *empeorar* un 10 % por debajo del mínimo conocido antes de que desaparezca.

Para implementar la migración y la tolerancia en MODPLAN diseñé un algoritmo sencillo basado en el cálculo de áreas de influencia y álgebra de mapas con GRASS. Los parámetros que el algoritmo tiene en consideración son:

- **UMBRAL OMISIÓN:** valor de idoneidad que deja por debajo la proporción de puntos de presencia calculada a partir del parámetro PORCENTAJE OMISIÓN establecido en MODPLAN. Es el valor de idoneidad dónde se separan el hábitat idóneo y el no idóneo.

*Se trata de una propuesta metodológica, que para una aplicación realista requiere un conocimiento previo del nivel de tolerancia real de las especies al cambio.*

- **IDONEIDAD MÍNIMA:** Es el menor valor de idoneidad obtenido por un registro de presencia de la especie.
- **MIGRACIÓN:** Distancia de migración potencial en número de celdas por intervalo de tiempo. El valor original calculado para la especie se toma como referencia de máximo, pero realmente la distancia de migración no se toma como constante, y se recalcula en cada intervalo de tiempo tomando un número aleatorio entre cero y MIGRACIÓN. Por tanto podemos distinguir  $MIGRACIÓN_{2000}$  (distancia de migración dentro del intervalo 2000-2010),  $MIGRACIÓN_{2010}$ , etc.
- **TOLERANCIA:** Es el valor de tolerancia de la especie en unidades de idoneidad del hábitat, según he descrito en los párrafos anteriores. Es constante en el tiempo.
- **IDONEIDAD:** el promedio de idoneidad del hábitat de los ocho algoritmos para cada intervalo de tiempo. Los nombraré sucesivamente:  $IDONEIDAD_{2010}$ ,  $IDONEIDAD_{2020}$ , etc.
- **PRESENCIA:** Área de presencia conocida de la especie.  $PRESENCIA_{2000}$  representan la *distribución conocida* inicial.  $PRESENCIA_{2010}$  y las sucesivas ya incorporan la migración y la tolerancia, por lo que realmente ya son *áreas de ocupación potencial*.

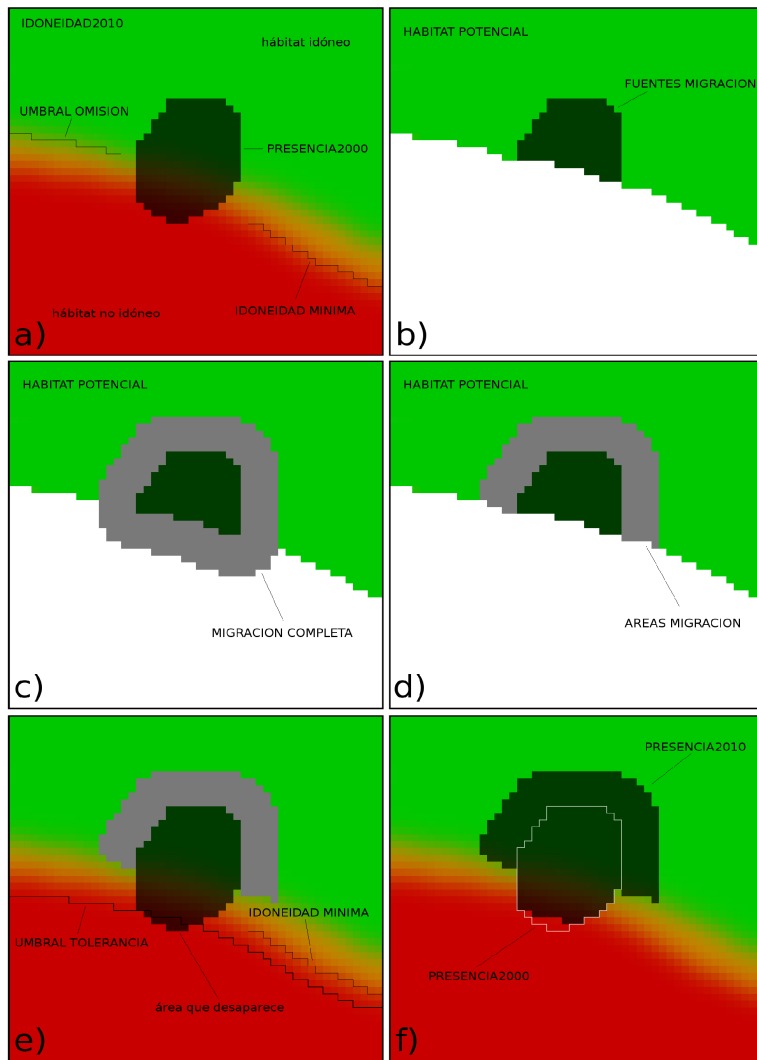
*Si el parámetro PORCENTAJE OMISIÓN es igual a cero, UMBRAL OMISIÓN e IDONEIDAD MÍNIMA son iguales.*

*La tolerancia es constante en el tiempo porque no asumo la posibilidad de adaptación, aunque para plantas anuales 100 generaciones son muchas generaciones.*

El algoritmo de simulación funciona como sigue (tomando como ejemplo la transición del año 2000 al 2010; ver Figura 3):

1. **Cálculo de fuentes de migración** (cuadros a y b en la Figura 3): Las fuentes de migración son celdas de presencia ocupadas por la especie en las que se da un hábitat suficientemente idóneo para permitir la producción de diásporas. De  $PRESENCIA_{2000}$  solo son fuentes de migración las que mantienen idoneidad del hábitat por encima de UMBRAL OMISIÓN en  $IDONEIDAD_{2010}$ .

*Las celdas que no están en hábitat idóneo no son fuentes de migración, asumiendo que en estas áreas aparecerán problemas*



**Figura 3:** Descripción del algoritmo de migración y tolerancia.

- El mapa  $IDONEIDAD_{2010}$  se recodifica para transformarlo en binario según el valor de  $UMBRAL\ OMISION$ , resultando  $HÁBITAT\ POTENCIAL_{2010}$ , que delimita las zonas de hábitat óptimo (Figura 3 cuadro b).
- $HÁBITAT\ POTENCIAL_{2010}$  se cruza con  $PRESENCIA_{2000}$ , para obtener  $FUENTES\ MIGRA-$

CIÓN, que contiene las poblaciones capaces de migrar con éxito.

2. **Cálculo de la migración** (cuadros c y d en la Figura 3): requiere tres pasos.

- Cálculo de  $MIGRACION_{2000}$ , que es un número aleatorio entre cero y el valor  $MIGRACION$ .
- Cálculo de un área de influencia sobre FUENTES  $MIGRACION_{2000}$  según la distancia  $MIGRACION_{2000}$  para obtener  $MIGRACION COMPLETA_{2000}$ . En este mapa están las zonas que quedan dentro de la distancia de migración desde FUENTES  $MIGRACION_{2000}$ , independientemente de la idoneidad del hábitat.
- Cálculo de áreas de migración con hábitat idóneo, cruzando  $MIGRACION COMPLETA_{2000}$  con  $OCUPACION POTENCIAL_{2010}$ , para eliminar las zonas de migración que están fuera del hábitat idóneo, obteniendo  $AREAS MIGRACION_{2010}$ .

*MIGRACION COMPLETA ocupa todo el área alrededor de FUENTES MIGRACION dentro de la distancia MIGRACION\_{2000}.*

3. **Desaparición de áreas de presencia** (cuadro e en la Figura 3):

- Cálculo de UMBRAL TOLERANCIA. Es el valor de idoneidad por debajo del cual las poblaciones de la especie desaparecen. Se calcula restando TOLERANCIA a IDONEIDAD MÍNIMA.
- Cálculo del mapa de áreas de tolerancia. Todas las zonas de  $IDONEIDAD_{2010}$  por encima de UMBRAL TOLERANCIA.
- Eliminación de las celdas de presencia fuera del área de tolerancia. Todas las celdas de  $PRESENCIA_{2000}$  que quedan en áreas de  $IDONEIDAD_{2010}$  por debajo del UMBRAL TOLERANCIA se eliminan.

4. **Cálculo de la distribución potencial al final de intervalo** (cuadro f en la Figura 3): Uniendo FUENTES  $MIGRACION_{2000}$  con  $AREAS MIGRACION_{2010}$  y habiendo eliminado las áreas de presencia bajo el umbral

de tolerancia se obtiene PRESENCIA<sub>20010</sub>, que representa el área de ocupación potencial de la especie en 2010.

### *Fragmentación antropógena del hábitat*

*La fragmentación geógena está implícita en los modelos de idoneidad del hábitat.*

Para tener en cuenta las limitaciones migratorias que impone la fragmentación antropógena del paisaje realicé un estudio sobre el Mapa de Usos y Coberturas de 1999, clasificando los usos del suelo como COMPATIBLES y NO COMPATIBLES con el desarrollo de poblaciones naturales de plantas. Para desarrollar el criterio superpuse todos los polígonos de presencia de las especies sobre el mapa de usos, y analicé todos aquellos usos sin presencia de alguna especie para evaluar su potencial real para acoger poblaciones naturales de plantas. Finalmente, con los usos COMPATIBLES generé una máscara en GRASS que evita que todos los cálculos de migración y tolerancia se realicen sobre áreas NO COMPATIBLES.

*En los mapas de salida las áreas NO COMPATIBLES aparecen en color gris, para diferenciarlas del resto de componentes de los mapas.*

### 8.3.4 *Calibrado y análisis de las simulaciones*

Los parámetros introducidos en PRP y MODPLAN fueron: RESOLUCIÓN = 100; MIN\_DIS = 200; MI\_PRES = 30; P\_OMISIÓN = 5.

Las variables seleccionadas para calibrar los modelos fueron: Precipitación anual, precipitación en verano, temperatura mínima de invierno, temperatura máxima de invierno, temperatura máxima de verano, curvatura del terreno, exposición topográfica, índice topográfico de humedad, pendiente y radiación solar de invierno.

*Respecto al Capítulo 7 he cambiado la variable orientación este-oeste (que resultó poco importante como descriptor de la distribución de las especies) por la curvatura del terreno.*

Las simulaciones se generaron sobre los escenarios A2 y B2 (según el modelo de circulación global CGCM2) para el periodo 2000-2100, en intervalos de 10 años.

Los parámetros de los algoritmos aplicados para calibrar los modelos fueron los mismos indicados en el Capítulo 7.

## *Evaluación de modelos*

Para clasificar las simulaciones según su fiabilidad, examiné los nuevos gráficos que ofrece MODPLAN tras su modificación, basados en las propiedades del modelo (ver Anexo ??). Clasifiqué cada modelo en una de las siguientes categorías, en orden descendente de fiabilidad: 1<sup>a</sup> (\*), 2<sup>a</sup>(\*\*), 3<sup>a</sup> (\*\*\*) y 4<sup>a</sup> (\*\*\*\*).

### *Respuesta individual de las especies al cambio del clima*

Analicé la respuesta del área de ocupación potencial de cada especie modelada para describir su comportamiento general en cada escenario. Las distintas posibilidades y su codificación están descritas a continuación:

- **Estabilidad (ES):** el área de ocupación se mantiene relativamente constante en el tiempo.
- **Contracción (C):** el área de ocupación potencial decrece desde el primer intervalo de tiempo. Incluye periodos de estabilidad transitorios.
- **Expansión (E):** el área de ocupación potencial se expande desde el primer intervalo de tiempo. Incluye periodos de estabilidad transitorios.
- **Expansión - contracción (EC):** la curva del área de ocupación tiene forma de campana, pero puede incluir periodos de estabilidad transitorios.
- **Contracción - expansión (CE):** la curva muestra forma de campana invertida, incluyendo periodos de estabilidad.
- **Extinción (EX):** el área de ocupación potencial alcanza el valor cero. Al lado se indica el intervalo de extinción (año de inicio del intervalo en el que la especie ya no aparece).

*La asignación de categorías es subjetiva. Un ejemplo de interpretación de los gráficos de evaluación puede verse en el Anexo ??*

### 8.3.5 Variaciones geográficas y temporales en la riqueza de especies

En este apartado aproveché la alta resolución espacial de las simulaciones para estudiar los patrones de variación geográfica y temporal de la riqueza de especies. En este trabajo el concepto *riqueza de especies* se refiere al número total de especies presentes en una celda cualquiera del territorio de estudio en un intervalo concreto de tiempo. La *variación en riqueza* de especies en el tiempo para una celda dada será por tanto la diferencia entre el número de especies entre  $t_0$  y  $t_1$ . Estudié ambos aspectos para todo el territorio andaluz, analizando la distribución geográfica de la riqueza de especies y la variación en riqueza entre el inicio y el final de las simulaciones.

Andalucía cuenta con más de un millón y medio de hectáreas de Espacios Naturales Protegidos (ENP). Con el objetivo de conocer como puede afectar el cambio del clima a su diversidad vegetal, realicé un análisis gráfico y comparativo de la distribución de valores de variación de riqueza de todas las celdas de cada ENP.

*Para simplificar el análisis solo considero como ENP los Parques Naturales y Parques Nacionales.*

#### *Curvas de extinción y factores causales*

Los resultados de MODPLAN muestran para cada especie el intervalo de tiempo para el que pronostica la extinción por desaparición de las poblaciones (cuando se da el caso de que se extinguen, según el modelo). Resumí esos datos de tiempo de extinción para generar dos series temporales de porcentaje acumulado de especies extinguidas (una para cada escenario), con el objetivo de conocer cual es el escenario con mayor potencial de extinción.

Para tratar de descubrir cuales son las variables climáticas con mayor influencia en las curvas de extinción, recopilé los valores medios (para toda Andalucía, en todos los intervalos y escenarios) de las variables predictoras de mayor interés: precipitación anual, precipitación de verano, temperatura media mínima de invierno y temperatura media máxima de verano.

Tras una exploración previa de los datos, los sometí a un análisis de regresión para localizar las variables con mayor influencia en la tasa de extinción y estimar ecuaciones lineales capaces de pronosticar el porcentaje de extinción según los valores de las variables climáticas con mayor influencia en el proceso.

#### 8.4 RESULTADOS

En total 198 especies presentaron más de 30 registros de presencia

A continuación aparece el listado con el análisis resumido del comportamiento de cada especie en ambos escenarios (incluye los resultados de la evaluación codificados mediante asteriscos; cuanto menor es el número de asteriscos, más fiable es el modelo).

*Las simulaciones de las 198 especies llevaron 28 días en 7 ordenadores de sobremesa (de diferentes características), con un promedio de 23 horas por especie.*

- Abies pinsapo* \*: **A2**, C; **B2**, ES.  
*Acer monspessulanum* \*: **A2**, E; **B2**, E.  
*Acer opalus granatense* \*: **A2**, EC; **B2**, EC.  
*Adenocarpus decorticans* \*: **A2**, C; **B2**, ES.  
*Alnus glutinosa* \*: **A2**, EC; **B2**, E.  
*Anthyllis cytisoides* \*\*: **A2**, E; **B2**, E.  
*Anthyllis plumosa* \*: **A2**, EX 2020; **B2**, EX 2030.  
*Arbutus unedo* \*\*: **A2**, E; **B2**, E.  
*Arenaria delaguardiae* \*: **A2**, EX 2020; **B2**, EX 2030.  
*Arenaria pungens* \*: **A2**, EX 2050; **B2**, EX 2040.  
*Arenaria racemosa* \*: **A2**, EX 2020; **B2**, EX 2050.  
*Arenaria tetraqueta amabilis* \*: **A2**, EX 2100; **B2**, C.  
*Armeria velutina* \*\*\*: **A2**, E; **B2**, E.  
*Artemisia barrelieri* \*\*: **A2**, C; **B2**, C.  
*Artemisia campestris glutinosa* \*: **A2**, EX 2090; **B2**, C.  
*Artemisia granatensis* \*: **A2**, C; **B2**, ES.  
*Arundo donax* \*\*\*\*: **A2**, E; **B2**, E.  
*Asparagus acutifolius* \*: **A2**, C; **B2**, ES.  
*Asplenium billotii* \*\*: **A2**, E; **B2**, E.  
*Astragalus edulis* \*: **A2**, EX 2060; **B2**, EX 2080.  
*Atropa baetica* \*: **A2**, EC; **B2**, E.  
*Avenella iberica* \*: **A2**, EX 2090; **B2**, EC.



*Berberis hispanica* \*: **A2**, EX 2100; **B2**, EC.  
*Berberis vulgaris australis* \*: **A2**, C; **B2**, ES.  
*Boreava aptera* \*: **A2**, EX 2050; **B2**, EX 2050.  
*Brachypodium retusum* \*: **A2**, C; **B2**, C.  
*Buxus balearica* \*: **A2**, EX 2100; **B2**, C.  
*Buxus sempervirens* \*: **A2**, EC; **B2**, E.  
*Calicotome villosa* \*: **A2**, C; **B2**, ES.  
*Calluna vulgaris* \*: **A2**, EC; **B2**, E.  
*Carex camposii* \*: **A2**, EX 2100; **B2**, EC.  
*Castanea sativa* \*\*: **A2**, EC; **B2**, E.  
*Celtis australis* \*\*: **A2**, E; **B2**, E.  
*Ceratonia siliqua* \*: **A2**, C; **B2**, C.  
*Chamaerops humilis* \*\*\*\*: **A2**, E; **B2**, E.  
*Cistus albidus* \*\*\*\*: **A2**, EC; **B2**, E.  
*Cistus clusii* \*\*: **A2**, C; **B2**, ES.  
*Cistus crispus* \*: **A2**, EC; **B2**, ES.  
*Cistus ladanifer* \*\*\*\*: **A2**, EC; **B2**, E.  
*Cistus laurifolius* \*: **A2**, E; **B2**, E.  
*Cistus monspeliensis* \*\*\*\*: **A2**, EC; **B2**, E.  
*Cistus populifolius* \*\*: **A2**, E; **B2**, E.  
*Cistus populifolius major* \*: **A2**, C; **B2**, ES.  
*Cistus salvifolius* \*\*: **A2**, EC; **B2**, E.  
*Corema album* \*: **A2**, C; **B2**, ES.  
*Corylus avellana* \*: **A2**, E; **B2**, E.  
*Crataegus laciniata* \*: **A2**, EC; **B2**, E.  
*Crataegus monogyna* \*: **A2**, EC; **B2**, E.  
*Crepis oporinoides* \*: **A2**, C; **B2**, C.  
*Cynomorium coccineum* \*: **A2**, EX 2100; **B2**, E.  
*Cytisus galianoii* \*: **A2**, EX 2080; **B2**, C.  
*Cytisus malacitanus* \*: **A2**, EC; **B2**, E.  
*Cytisus scoparius* \*: **A2**, EX 2100; **B2**, C.  
*Cytisus striatus* \*: **A2**, C; **B2**, E.  
*Dactylis juncinella* \*: **A2**, EX 2090; **B2**, C.  
*Daphne gnidium* \*: **A2**, C; **B2**, ES.  
*Draba hispanica laderoi* \*: **A2**, EX 2050; **B2**, EX 2040.  
*Drosophyllum lusitanicum* \*: **A2**, C; **B2**, EC.  
*Echinospartium boissieri* \*: **A2**, EC; **B2**, EC.  
*Erica andevalensis* \*: **A2**, C; **B2**, E.  
*Erica arborea* \*\*\*\*: **A2**, E; **B2**, E.  
*Erica australis* \*\*: **A2**, EC; **B2**, E.

*Erigeron frigidus* \*: **A2**, EX 2100; **B2**, E.  
*Erinacea anthyllis* \*: **A2**, E; **B2**, E.  
*Erodium rupicola* \*: **A2**, C; **B2**, E.  
*Eucalyptus camaldulensis* \*\*\*\*: **A2**, E; **B2**, E.  
*Eucalyptus globulus* \*\*\*\*: **A2**, EC; **B2**, E.  
*Euphorbia gaditana* \*\*\*\*: **A2**, EX 2080; **B2**, EC.  
*Euzomodendron bourgaeum* \*: **A2**, EX 2090; **B2**, C.  
*Festuca clementei* \*: **A2**, C; **B2**, ES.  
*Festuca indigesta* \*: **A2**, EX 2100; **B2**, C.  
*Festuca pseudoeskia* \*: **A2**, EX 2020; **B2**, EX 2040.  
*Festuca scariosa* \*: **A2**, EX 2100; **B2**, C.  
*Flueggea tinctoria* \*\*: **A2**, E; **B2**, E.  
*Frangula alnus baetica* \*: **A2**, E; **B2**, E.  
*Galium rosellum* \*: **A2**, EX 2010; **B2**, EX 2010.  
*Galium viridiflorum* : **A2**, EC; **B2**, EC.  
*Genista cinerea* \*: **A2**, EC; **B2**, E.  
*Genista florida* \*: **A2**, EX 2090; **B2**, C.  
*Genista hirsuta* \*\*\*\*: **A2**, E; **B2**, E.  
*Genista pumila pumila* \*: **A2**, EX 2090; **B2**, C.  
*Genista scorpius* \*\*: **A2**, EC; **B2**, E.  
*Genista spartioides retamoides* \*: **A2**, EX 2100; **B2**, C.  
*Genista umbellata* \*\*: **A2**, C; **B2**, ES.  
*Genista versicolor* \*: **A2**, EX 2100; **B2**, C.  
*Halimium atriplicifolium* \*: **A2**, EX 2070; **B2**, C.  
*Halimium halimifolium* \*\*: **A2**, C; **B2**, C.  
*Halimium viscosum* \*\*: **A2**, EX 2080; **B2**, EX 2080.  
*Hammada articulata* \*: **A2**, EX 2070; **B2**, EX 2050.  
*Helictotrichon filifolium cazorlense* \*: **A2**, EX 2100; **B2**, C.  
*Holcus caespitosus* \*: **A2**, EX 2060; **B2**, EX 2100.  
*Hormatophylla spinosa* \*: **A2**, C; **B2**, ES.  
*Hyparrhenia hirta* \*\*: **A2**, E; **B2**, E.  
*Ilex aquifolium* \*: **A2**, E; **B2**, E.  
*Juniperus communis* \*\*: **A2**, C; **B2**, ES.  
*Juniperus oxycedrus* \*\*: **A2**, E; **B2**, E.  
*Juniperus oxycedrus macrocarpa* \*: **A2**, C; **B2**, C.  
*Juniperus phoenicea* \*: **A2**, C; **B2**, ES.  
*Juniperus phoenicea turbinata* \*: **A2**, EC; **B2**, EC.  
*Juniperus sabina* \*: **A2**, EX 2090; **B2**, C.  
*Laurus nobilis* \*: **A2**, E; **B2**, E.  
*Lavandula lanata* \*: **A2**, C; **B2**, C.

*Lavandula latifolia* \*: **A2, C; B2, C.**  
*Lavandula stoechas* \*\*\*\*: **A2, EC; B2, E.**  
*Leontodon boryi* \*: **A2, EX 2090; B2, EC.**  
*Leontodon microcephalus* \*: **A2, EX 2090; B2, EC.**  
*Linaria benitoi* \*: **A2, EX 2020; B2, EX 2040.**  
*Linaria nigricans* \*: **A2, C; B2, C.**  
*Linaria tursica* \*: **A2, EX 2100; B2, C.**  
*Lithodora nitida* \*: **A2, EX 2100; B2, E.**  
*Loeflingia baetica* \*\*\*: **A2, E; B2, E.**  
*Lygeum spartum* \*: **A2, ES; B2, E.**  
*Moehringia fontqueri* \*: **A2, EX 2080; B2, C.**  
*Myrtus communis* \*\*: **A2, E; B2, E.**  
*Narcissus bugei* \*: **A2, E; B2, E.**  
*Narcissus fernandesii* \*: **A2, EC; B2, EC.**  
*Narcissus longispathus* \*: **A2, EC; B2, EC.**  
*Narcissus nevadensis* \*: **A2, EX 2060; B2, EX 2090.**  
*Nerium oleander* \*\*\*\*: **A2, E; B2, E.**  
*Olea europaea sylvestris* \*\*\*\*: **A2, E; B2, E.**  
*Ononis tridentata* \*: **A2, EX 2020; B2, EX 2030.**  
*Phillyrea angustifolia* \*\*: **A2, E; B2, E.**  
*Phillyrea latifolia* \*: **A2, EC; B2, E.**  
*Phlomis purpurea* \*\*: **A2, EC; B2, ES.**  
*Phragmites australis* \*\*\*\*: **A2, E; B2, E.**  
*Picris willkommii* \*: **A2, EX 2020; B2, EX 2100.**  
*Pinguicula vallisneriifolia* \*: **A2, ES; B2, E.**  
*Pinus halepensis* \*\*\*\*: **A2, E; B2, E.**  
*Pinus nigra* \*: **A2, EX 2090; B2, C.**  
*Pinus nigra salzmannii* \*: **A2, C; B2, ES.**  
*Pinus pinaster* \*\*\*\*: **A2, E; B2, E.**  
*Pinus pinea* \*\*\*\*: **A2, E; B2, E.**  
*Pinus sylvestris* \*: **A2, EX 2100; B2, C.**  
*Pinus sylvestris nevadensis* \*: **A2, C; B2, ES.**  
*Pistacia lentiscus* \*\*\*\*: **A2, E; B2, E.**  
*Pistacia terebinthus* \*: **A2, E; B2, E.**  
*Populus alba* \*\*\*: **A2, E; B2, E.**  
*Populus nigra* \*\*\*: **A2, E; B2, E.**  
*Primula elatior lofthousei* \*: **A2, EC; B2, E.**  
*Prunus avium* \*: **A2, EC; B2, E.**  
*Prunus dulcis* \*\*\*: **A2, EC; B2, E.**  
*Prunus insititia* \*\*: **A2, EC; B2, E.**

*Prunus mahaleb* \*: **A2**, E; **B2**, E.  
*Quercus canariensis* \*: **A2**, E; **B2**, E.  
*Quercus coccifera* \*: **A2**, E; **B2**, E.  
*Quercus faginea* \*\*\*\*: **A2**, EC; **B2**, E.  
*Quercus fruticosa* \*: **A2**, C; **B2**, ES.  
*Quercus ilex ballota* \*\*\*\*: **A2**, E; **B2**, E.  
*Quercus pyrenaica* \*: **A2**, C; **B2**, ES.  
*Quercus suber* \*\*\*\*: **A2**, E; **B2**, E.  
*Retama sphaerocarpa* \*\*\*\*: **A2**, E; **B2**, E.  
*Rhododendron ponticum baeticum* \*: **A2**, EX 2100; **B2**, C.  
*Rosa sempervirens* \*: **A2**, EC; **B2**, E.  
*Rosmarinus officinalis* \*\*\*\*: **A2**, E; **B2**, E.  
*Rosmarinus tomentosus* \*: **A2**, C; **B2**, EC.  
*Rothmaleria granatensis* \*: **A2**, EX 2070; **B2**, EX 2090.  
*Rubus ulmifolius* \*\*\*\*: **A2**, E; **B2**, E.  
*Rupicapnos africana decipiens* \*: **A2**, EC; **B2**, EC.  
*Ruscus aculeatus* \*: **A2**, EX 2060; **B2**, EX 2090.  
*Salix atrocinerea* \*: **A2**, C; **B2**, E.  
*Salix caprea* \*: **A2**, EC; **B2**, ES.  
*Salix eleagnos angustifolia* \*: **A2**, EC; **B2**, E.  
*Salsola genistoides* \*: **A2**, C; **B2**, C.  
*Salsola papillosa* \*: **A2**, EX 2100; **B2**, EC.  
*Salsola vermiculata* \*\*\*\*: **A2**, EC; **B2**, E.  
*Santolina canescens* \*\*: **A2**, EC; **B2**, E.  
*Satureja obovata* \*: **A2**, C; **B2**, ES.  
*Scirpus holoschoenus* \*\*\*\*: **A2**, E; **B2**, E.  
*Senecio auricula* \*: **A2**, EX 2080; **B2**, C.  
*Sideritis incana* \*: **A2**, C; **B2**, C.  
*Silene fernandezii* \*: **A2**, EX 2070; **B2**, EX 2060.  
*Silene mariana* \*\*: **A2**, EX 2100; **B2**, CE.  
*Sorbus aria* \*: **A2**, E; **B2**, E.  
*Spartium junceum* \*: **A2**, EC; **B2**, EC.  
*Stipa tenacissima* \*\*\*\*: **A2**, EC; **B2**, E.  
*Tamarix canariensis* \*: **A2**, EX 2100; **B2**, EC.  
*Tamarix gallica* \*\*\*\*: **A2**, EC; **B2**, E.  
*Taxus baccata* \*: **A2**, EC; **B2**, E.  
*Teucrium charidemi* \*: **A2**, EX 2080; **B2**, EX 2040.  
*Thymbra capitata* \*: **A2**, EX 2100; **B2**, C.  
*Thymelaea tartonraira* \*: **A2**, EX 2020; **B2**, EX 2030.  
*Thymus albicans* \*\*: **A2**, C; **B2**, C.

*Thymus hyemalis* \*: **A2**, C; **B2**, C.  
*Thymus mastichina* \*\*: **A2**, C; **B2**, E.  
*Thymus orospedanus* \*: **A2**, C; **B2**, ES.  
*Thymus serpylloides* \*: **A2**, EX 2060; **B2**, EX 2060.  
*Thymus zygis gracilis* \*\*\*\*: **A2**, C; **B2**, E.  
*Trisetum glaciale* \*: **A2**, EX 2050; **B2**, EX 2040.  
*Ulex baeticus* \*\*: **A2**, C; **B2**, E.  
*Ulex canescens* \*: **A2**, EX 2060; **B2**, EX 2040.  
*Ulex eriocladus* \*\*\*\*: **A2**, EC; **B2**, E.  
*Ulex parviflorus* \*\*\*\*: **A2**, EC; **B2**, E.  
*Ulex parviflorus rivasgodayanus* \*: **A2**, EX 2020; **B2**, EX 2030.  
*Ulmus minor* \*\*\*\*: **A2**, E; **B2**, E.  
*Vella pseudocytisus pseudocytisus* \*: **A2**, EX 2070; **B2**, EC.  
*Viburnum lantana* \*: **A2**, EC; **B2**, E.  
*Viola cazorlensis* \*: **A2**, EC; **B2**, E.  
*Viola crassiuscula* \*: **A2**, EX 2020; **B2**, EX 2020.

30 modelos obtuvieron una clasificación de 4<sup>a</sup> categoría, 10 obtuvieron la 3<sup>a</sup> categoría, 24 la 2<sup>a</sup> y los restantes 134 la 1<sup>a</sup> de máxima fiabilidad. No excluí ningún modelo de los análisis.

La Figura 4 resume gráficamente los datos del listado anterior.

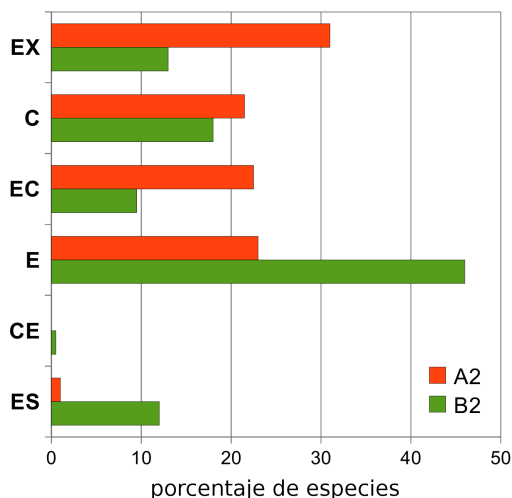
#### 8.4.1 Variaciones geográficas y temporales en la riqueza de especies

La Figura 5 muestra el patrón de diversidad vegetal a partir de los datos de presencia conocida de las 198 especies modeladas.

La Figura 6 muestra la variación de riqueza de especies para ambos escenarios. Los resultados son similares, con un gradiente oeste-este de pérdida de riqueza, que afecta en mayor medida a las montañas y zonas áridas de Andalucía Oriental.

Las Figuras 7 y 8 muestran la distribución de valores de diferencial de riqueza para los ENP. Los resultados indican que el ENP en peor situación en ambos escenarios es Sierra

Las bases de datos  
 geográficas usadas  
 muestran cierto  
 grado de  
 incertidumbre  
 taxonómica. No he  
 modificado los  
 nombres de especies  
 que en ellas aparecen  
 para facilitar la  
 replicabilidad de los  
 resultados.



**Figura 4:** Comparación entre ambos escenarios según el comportamiento del área de ocupación potencial de las especies. EX=extinción, C=contracción, EC=expansión-contracción, E=expansión, CE=contracción-expansión, ES=estabilidad.

Nevada. La segunda posición varía entre escenarios, siendo Sierra de Cazorla en A2 y Doñana en B2.

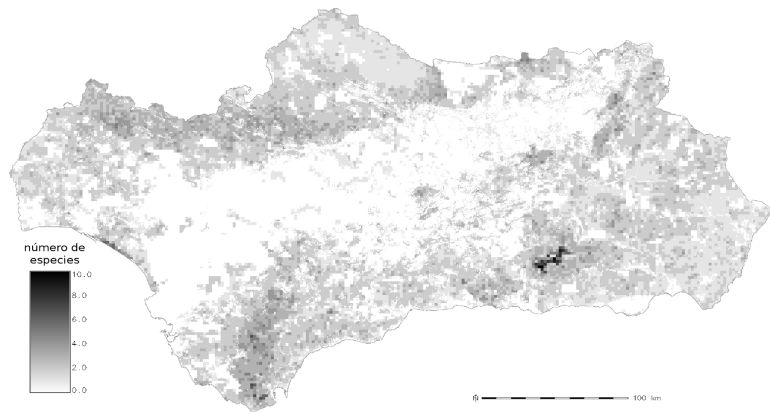
#### *Curvas de extinción y factores causales*

Las curvas de extinción obtenidas para cada escenario pueden verse en la Figura 9.

Las curvas de extinción mostraron un ajuste lineal robusto con algunas de las variables climáticas. La variable climática que por sí sola explicó una mayor variación del porcentaje de extinción para el escenario A2 fue la temperatura media máxima de verano ( $R^2 = 0.9186$ ,  $F = 101.50$   $p=3.4 \cdot 10^{-7}$ ). La mejor combinación de variables para predecir la extinción en este escenario resultó ser temperatura media máxima de verano junto a precipitación anual ( $R^2 = 0.9453$ ,  $F = 69.15$   $p=8.9 \cdot 10^{-6}$ ), con una ecuación de la forma:

$$\%extA2 = 2,244 \cdot TMXV - 0,061 \cdot PA - 39,311 \quad (8.1)$$

*Estas gráficas las preparé mediante spgrass6. Esta librería permite introducir los mapas de GRASS en R para realizar análisis estadísticos.*



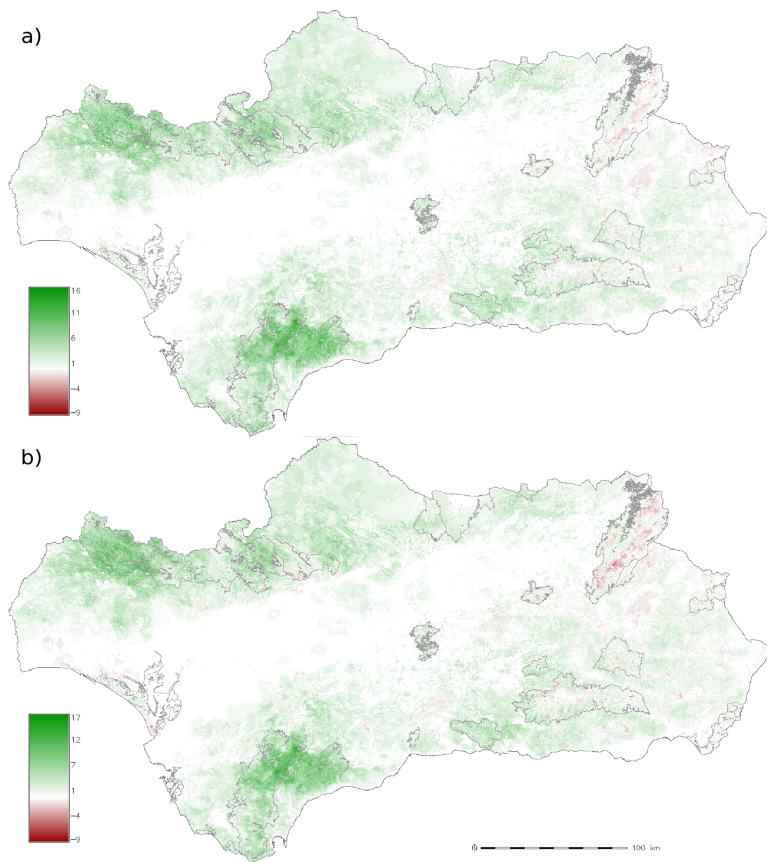
**Figura 5:** Patrón geográfico de diversidad vegetal calculado a partir de los datos de presencia de las 198 especies modeladas. La resolución del mapa se ha reducido a 2000 x 2000 m para facilitar su visualización. El valor de cada celda corresponde al máximo encontrado en el área que representa.

Donde %extA2 es el porcentaje de especies extinguidas, TMXV representa la temperatura media máxima de verano y PA representa la precipitación media anual.

La variable climática más importante para explicar las variaciones en porcentaje de extinción para el escenario B2 también fue la temperatura media máxima de verano ( $R^2 = 0.9490$ ,  $F = 167.40$   $p=4.0 \cdot 10^{-6}$ ). El mejor conjunto de variables para predecir la extinción en este escenario fue el formado por la temperatura media máxima de verano y la temperatura mínima de invierno ( $R^2 = 0.9629$ ,  $F = 103.90$   $p=1.9 \cdot 10^{-6}$ ), según la ecuación:

$$\%extB2 = 2,836 \cdot TMXV + 3,118 \cdot TMNI - 104,233 \quad (8.2)$$

La tabulación de todos los datos en conjunto facilitó el ajuste de un modelo lineal con el que explicar las variaciones en porcentaje de extinción independientemente del escenario, tomando la temperatura media máxima de verano y la

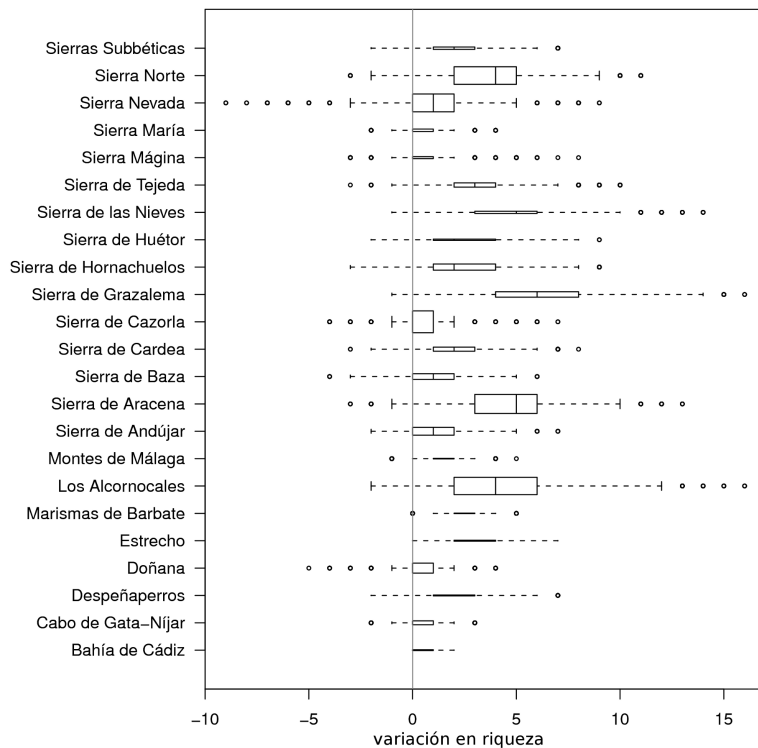


**Figura 6:** Patrón geográfico de variación en la riqueza de especies al final de la simulación (2100). Las figuras **a** y **b** se refieren a los escenarios A2 y B2 respectivamente. En verde las áreas que ganan especies respecto a su riqueza inicial, y en rojo aquellas que las pierden. Aparecen los límites de los ENP como referencia.

precipitación media anual como predictores ( $R^2 = 0.9404$ ,  $F = 150.00$   $p=2.304 \cdot 10^{-12}$ ). La ecuación resultante fue:

$$\%ext = 2,95632 \cdot TMXV - 0,03131 \cdot PA - 78,79813 \quad (8.3)$$



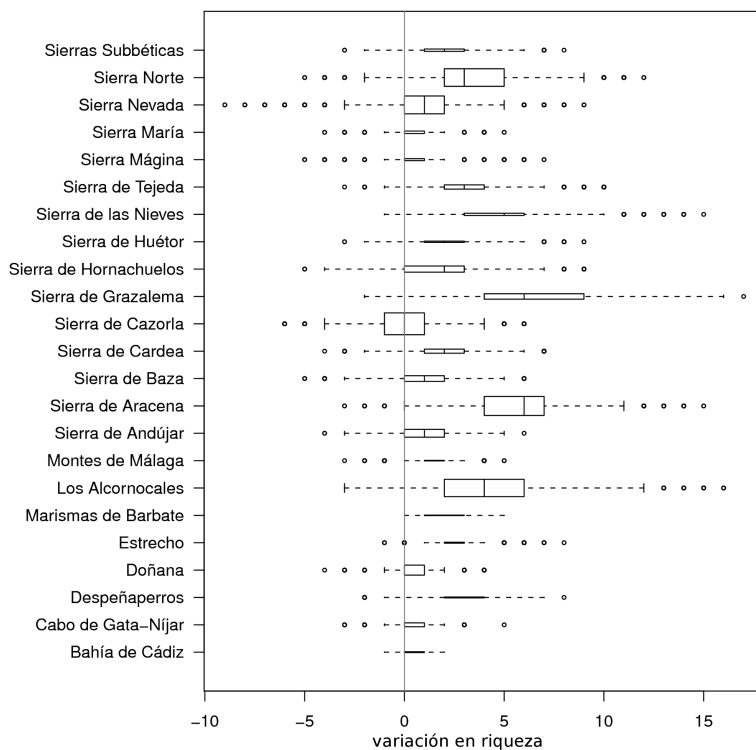


**Figura 7:** Distribución de valores de variación de riqueza dentro de los ENP según el escenario A2 en el año 2100. El alto de las cajas es relativo a la superficie del ENP al que representan.

## 8.5 DISCUSIÓN

### 8.5.1 Limitaciones del algoritmo de migración-tolerancia

Con la incorporación de la migración, cambia la interpretación de los resultados que ofrece MODPLAN. Por una parte genera el **modelo de idoneidad del hábitat**, y sobre él, aplicando la migración, la tolerancia y la fragmentación del hábitat, a partir de los datos de presencia real de la especie, calcula el **área de ocupación potencial**. Según la Figura 1 del Capítulo 2, ahora MODPLAN ofrece como resultados la unión de las áreas **d** (área de presencia real) y **f** (zonas con topografía y clima idóneo a las que la especie puede llegar),

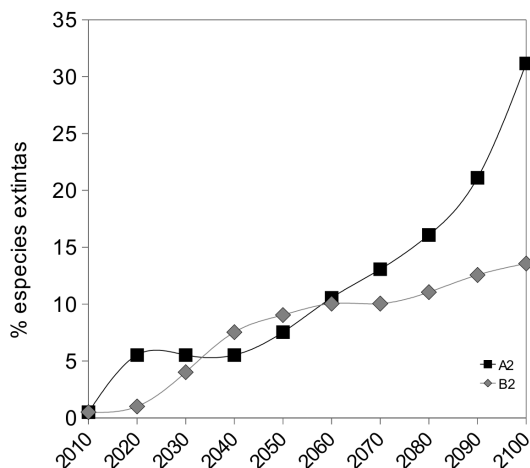


**Figura 8:** Distribución de valores de variación de riqueza dentro de los ENP según el escenario B2 en el año 2100

mucho más cercana a la **distribución efectiva** de la especie que los modelos generados en los capítulos anteriores.

El algoritmo de migración implementado en MODPLAN, no obstante, es simplista, porque solo tiene en cuenta el factor dispersión, obviando otros como la fecundidad o el reclutamiento.

Dos limitaciones a destacar son la isotropía de la simulación de migración, que no tiene en cuenta la tendencia direccional de los vectores de dispersión, y la omisión del efecto de barrera geográfica. Esta última solo afecta a las especies con grandes distancias de dispersión, que por otra parte, suelen ser dispersadas por aves, menos sensibles a este tipo de barreras.



**Figura 9:** Porcentaje de especies extintas para cada intervalo y escenario.

También omite la migración a larga distancia, porque, aun siendo fácil de implementar, no disponemos de datos apropiados para obtener modelos fiables.

Otra cuestión a tener en cuenta que no debe considerarse una limitación, es que los resultados no son deterministas. En primer lugar porque uno de los algoritmos (GARP) del ensamblado no lo es, y en segundo porque no he considerado constante en el tiempo la distancia de dispersión (aunque sí en el espacio, porque es la misma en cada intervalo de tiempo para todo el territorio), cuando realmente es específica de la dinámica de cada población (Thuiller *et al.*, 2008). El no determinismo hace que dos realizaciones de un mismo modelo ofrezcan resultados diferentes. En estas circunstancias la mejor aproximación es simular el proceso  $n$  veces y trabajar con la tendencia del conjunto de simulaciones, pero no ha sido posible asumir el coste computacional de esta aproximación.

La omisión de la adaptación se debe a que las evidencias disponibles en la literatura apuntan a que el ritmo de cambio evolutivo es, en general, mucho más lento que el ritmo previsto para el cambio del clima (Etterson y Shaw, 2001).

Una de las limitaciones más importantes es la omisión de las interacciones entre especies, fundamentalmente la

competencia por el espacio. El algoritmo simula migración a celdas vecinas, sin tener en cuenta qué especies están en esas celdas. Diseñar y desarrollar una máquina de simulación que tenga en cuenta este aspecto requiere un conocimiento previo de las relaciones que se establecen entre todas las especies que se modelan, y es impracticable con los medios técnicos y el conocimiento sobre la biología de las especies de los que disponemos actualmente.

### 8.5.2 Interpretación biológica de los resultados

Los sistemas naturales no son cerrados, y es imposible que los modelos tengan en cuenta todos los factores que afectan a la distribución de las especies (Araújo *et al.*, 2005a), especialmente en un entorno cambiante. Por este motivo, las conclusiones extraídas de este trabajo no deben tomarse en como ciertas, sino como posibilidades que dependen de un cierto número importante de limitaciones metodológicas e incertidumbres biológicas.

En este trabajo he seguido una línea ascendente, desde el caso particular, la especie, hasta una visión más general de la diversidad vegetal a escala regional. A nivel específico la influencia de los parámetros de calibrado (especialmente PORCENTAJE OMISIÓN, MIGRACIÓN y TOLERANCIA) es grande, y uno solo de ellos con un valor poco realista puede llevar al error en los resultados. Las especies para las que probablemente el modelo está cometiendo mayor error son aquellas con un umbral de omisión bajo y una alta migración. En estas especies el hábitat idóneo según el modelo es muy amplio (debido a un umbral de corte bajo), lo que les abre mucho área de migración, que sumada a la gran distancia de migración potencial les otorga un comportamiento invasivo. Esta circunstancia se da en *Anthyllis cytisoides*, *Arbutus unedo*, *Celtis australis*, *Flueggea tinctoria*, *Genista scorpius*, *Juniperus oxycedrus*, *Laurus nobilis*, *Pistacia lentiscus* y *Rubus ulmifolius*.

Otras especies con un comportamiento anómalo son aquellas cuya superficie de ocupación conocida es escasa, y la distancia de migración amplia. Estas especies muestran un

*El comportamiento invasivo de estas especies en las simulaciones, en la realidad está limitado por restricciones en el reclutamiento y la competencia por el espacio.*

salto grande entre el área conocida en el año 2000 y el primer cálculo de área de ocupación potencial en 2010. También se da el caso de especies cuyos factores determinantes de la distribución no se han incluido entre las variables. Por ejemplo, *Arbutus unedo* solo puede migrar a zonas silíceas, o *Flueggea tinctoria*, que aparece principalmente en cursos de agua intermitentes en zonas silíceas.

Es interesante que el comportamiento mayoritario en el escenario A2 es la extinción, con más de un 30 % de especies desaparecidas, mientras que en B2 es la expansión. Las variaciones climáticas previstas para B2 son más suaves, y esto explicaría el patrón general de comportamiento. Sin embargo, el análisis de variación en riqueza de especies para los ENP arroja un resultado contradictorio, porque en ellos el escenario B2 presenta consecuencias más severas, especialmente en Andalucía Oriental.

Para Sierra Nevada es posible explicar el efecto, por su condición de alta montaña, más sensible al incremento de la temperatura, y alto número de especies endémicas propias de las cotas más altas. Por otra parte, estas especies no tienen a dónde migrar, quedando confinadas hasta que superan el límite de tolerancia, cuando desaparecen. En La sierra de Cazorla la situación probablemente se debe, además de a la temperatura, al comportamiento dispar de las precipitaciones entre escenarios. En el escenario A2 se da un descenso generalizado de la precipitación en estas montañas, mientras en el B2, aunque hay un descenso regional del promedio, en algunos lugares podrían incrementarse hasta 100 mm. En cualquiera de los dos casos supone un alejamiento de las condiciones actuales, y el sistema no puede predecir si estos cambios son beneficiosos para las especies implicadas.

En las simulaciones destaca la gran variación positiva de riqueza que alcanzan las sierras de Cádiz y Málaga, especialmente sorprendente si observamos los mapas de pronóstico de precipitaciones en el futuro, que muestran descensos muy importantes en estos sistemas montañosos. Este efecto se debe a la expansión tal vez poco realista que experimentan algunas especies que ya he comentado anteriormente como *Arbutus unedo*, *Celtis australis*, y otras con un grado de expansión más realista, como *Chamaerops humilis*, los distin-

tos *Cistus* presentes en la zona y *Erica arborea*, entre otras. Habría que valorar, además de la riqueza, el *valor* de las especies implicadas en el incremento de riqueza.

La situación de las zonas áridas de Andalucía Oriental es aparentemente preocupante (ya traté el tema en el capítulo anterior), pero es importante entender que en estas áreas los muestreos son poco densos, y la muestra de especies es pequeña. Si de 4 especies 4 se extinguen, obtenemos un área de variación negativa, pero el resultado es poco significativo. Además algunos estudios indican que variaciones en la precipitación de hasta el 25 % tienen escaso impacto en las comunidades de plantas anuales de zonas áridas, debido a las adaptaciones de estas especies a un ambiente irregular (Miranda *et al.*, 2009).

*El análisis pormenorizado de los resultados individuales va más allá de los objetivos de esta tesis, pero la idea es seguir con ellos en el futuro inmediato.*

### 8.5.3 *El patrón regional de extinción*

Ante el escenario A2, que es aproximadamente una proyección lineal de nuestra historia actual, más de un 30 % de la flora estudiada podría extinguirse hasta el final de siglo, con unos 20 o 30 años de estabilidad inicial que pasa a una curva ascendente exponencial. En el escenario B2, que describe un mundo en el que se trabaja por la sostenibilidad ambiental, la tendencia sigue siendo preocupante, aunque realmente solo se diferencia del A2 a partir del 2060.

Teniendo en cuenta que MODPLAN ha modelado 200 especies con un algoritmo de migración simplista, con parámetros estimados según amplios márgenes de incertidumbre, es sorprendente que a escala regional, y tomando solo los valores promedio del clima en todo el territorio, la relación entre variación en el clima y extinción sea prácticamente perfecta, según los modelos lineales desarrollados a partir de los datos.

Este patrón tan perfecto (pero terrible por lo que implica), plasmado en una sola ecuación que pronostica el porcentaje de extinción de especies a partir de solo dos variables climáticas, muestra que independientemente de la complejidad del método de simulación, en el problema de los efectos del

*Fue una verdadera sorpresa encontrar un patrón tan robusto a partir de las simulaciones.*

cambio del clima sobre la flora hay una señal de fondo muy robusta.

## 8.6 CONCLUSIONES

- Los datos sobre dispersión, migración y tolerancia en plantas son escasos. Es importante que la comunidad de botánicos y ecólogos vegetales haga un esfuerzo adicional para tratar de mejorar nuestros conocimientos sobre estos procesos, vitales para entender las consecuencias potenciales del cambio del clima en la diversidad vegetal.
- Andalucía podría enfrentarse a pérdidas de diversidad vegetal entre el 13 y el 32 % según el escenario, y esta pérdida de riqueza vegetal puede afectar en mayor medida a los espacios naturales de la región oriental, especialmente Sierra Nevada, Sierra de Cazorla y las zonas áridas de Granada y Almería.
- Incluso utilizando un sistema de simulación muy complejo, la respuesta del sistema estudiado es simple pero muy significativa: cuanta más temperatura y menos precipitación, mayor porcentaje de extinción.
- La identificación de zonas con ganancias o pérdidas de especies significativas, puede ayudar a realizar una planificación en conservación y adelantarnos a algunos de los cambios que según los modelos pueden producirse

## Parte IV

# REFLEXIONES Y CONCLUSIONES





## REFLEXIONES FINALES

## REFLEXIONES FINALES

## 9.1 LA ECOINFORMÁTICA

El uso de herramientas informáticas es cada vez más usual entre los investigadores en distintas áreas de la Biología, y no me refiero a los procesadores de texto o las hojas de cálculo. Me refiero a las herramientas de análisis geográfico (SIG), las de análisis estadístico, lenguajes de modelado y cálculo, algoritmos predictivos, etc. Estas herramientas, junto a los avances en equipamiento informático (disminución de costes e incremento de potencia) están permitiendo a los investigadores manejar volúmenes de información desconocidos hace apenas 10 años.

En este escenario de asimilación de Tecnologías de la Información, el concepto de *código abierto* y la comunidad que lo desarrolla están teniendo un impacto importante en la investigación.

El software de código abierto no es mejor ni peor que el software de código cerrado, aunque algunas aplicaciones de código cerrado no tienen parangón entre las de código abierto, por su calidad y funcionalidades (algunos ejemplos podrían ser MATLAB, JMP, ArcINFO y ERDAS, entre otras).

La elección de uno u otro debe obedecer en primer lugar a un criterio de funcionalidad y flexibilidad, y en segundo a la disponibilidad de recursos económicos.

El *coste cero* del software libre posibilita la utilización de distintos programas en muchos ordenadores, permitiendo destinar los recursos financieros de un proyecto de investigación a material fungible o investigación de campo.

Por otra parte el catálogo de aplicaciones libres crece día a día, cubriendo campos que atañen a todas las disciplinas de la biología. Las características del código libre facilitan que la calidad de los programas crezca exponencialmente, porque

*En los capítulos 3, 4 y 5 utilicé Idrisi Andes, un programa de relativamente caro pero excepcional. En los capítulos restantes, lo sustituí por GRASS, libre y gratuito, y gané en productividad y posibilidades gracias a su capacidad de integración con otros programas.*

los propios usuarios (con conocimientos en lenguajes de programación, eso sí) puedan añadir nuevas funcionalidades (o sugerir modificaciones a los desarrolladores).

Un esfuerzo en conocer el panorama del software libre aplicado a la ciencia y un intento por mejorar nuestras aptitudes para utilizarlo siempre repercutirá positivamente en la calidad de nuestra investigación.

Los avances en software, hardware, conocimiento ecológico y recolección de datos están permitiendo a la Ecoinformática consolidarse a nivel internacional como una disciplina con entidad propia dentro de la Biología. La Ecoinformática tiene el potencial de proporcionar respuestas a grandes problemas relativos a la conservación de las especies, el pronóstico de efectos de los cambios ambientales e incluso la salud pública. Sin embargo en nuestro país apenas es conocida, y es importante que la comunidad científica en general y la Universidad en particular hagan un esfuerzo para hacerla crecer. Son necesarios planes de formación específicos en la materia, y en este sentido es imprescindible la creación de grupos de investigación interdisciplinarios con especialistas en física, matemáticas, biología, estadística, y ciencias de la computación. Tenemos especialistas, y tenemos los medios, pero falta la visión de futuro y la voluntad de romper barreras entre departamentos para mirar más allá de la propia especialidad.

*Los MDE se usan para pronosticar vías de acceso para vectores de enfermedades infecciosas.*

## 9.2 LOS MDE

Los MDE son una herramienta con un amplio rango de aplicaciones en distintos campos de la biología. Pero debido a limitaciones técnicas, metodológicas, conceptuales y de falta de conocimiento ecológico sobre las especies y sus relaciones, aún tienen un amplio margen de mejora. Es importante hacer un esfuerzo para avanzar en su desarrollo, para convertirlos en herramientas plenamente fiables en dos campos. Primero, la investigación teórica, que atañe a la biogeografía y evolución (entre otras), donde los MDE ayudarán a explicar las variaciones de la biota en el tiempo geológico. Y en segundo lugar, es necesario llevar los MDE más allá

del ámbito científico, preparándolos para su aplicación real en la toma de decisiones. Es el único modo de que nuestro esfuerzo en investigación tenga un impacto real y positivo en el mundo que nos rodea.

Algunos de los factores con gran importancia en la distribución geográfica de las especies, como la migración o las interacciones biológicas, son raramente incluidos en los modelos de distribución. En parte se debe a un inmaduro desarrollo metodológico, pero la principal limitación está en la falta de datos. Sabemos poco sobre la distribución real de muchas especies, sus mecanismos de dispersión, capacidad migratoria e interacciones con otras especies. Es crucial dar un paso adelante en la investigación experimental y observacional en estos campos si pretendemos mejorar nuestras posibilidades de pronosticar el efecto potencial de los cambios ambientales en la distribución de las especies, y conocer sus posibilidades de conservación.

### 9.3 EL CAMBIO DEL USOS DEL SUELO

El desarrollo socio-económico ha estado ligado históricamente al cambio de uso del suelo, ignorando hasta nuestros tiempos su impacto en la biodiversidad, por desconocimiento de la misma y/o falta de interés. Hoy día la falta de conocimiento no es una excusa, pero el interés no está dirigido del todo hacia la sostenibilidad y conservación de la biodiversidad.

La expansión abusiva de los invernaderos en Andalucía Oriental ha destruido hábitats naturales de gran valor, y continúan su expansión poniendo en peligro plantas amenazadas como *Linaria nigricans* (y sus valiosos acompañantes), una de cuyas poblaciones prácticamente ha desaparecido por la construcción de estas estructuras de cultivo intensivo. Salvo cambios en la dinámica económica o en la concienciación ambiental de la administración, este problema seguirá extendiéndose, afectando a más poblaciones de esta planta y otras especies de los valiosos hábitats áridos del sureste de la Península Ibérica. Disponemos de herramientas que pueden ayudar a reducir el impacto del cambio de uso del

suelo, detectando posibles conflictos geográficos entre intereses económicos y especies amenazadas. Es importante avanzar en el desarrollo de estas metodologías para hacerlas confiables, y ofrecer a los gestores herramientas de apoyo a la toma de decisiones.

#### 9.4 EL CAMBIO DEL CLIMA

Las simulaciones climáticas basadas en los escenarios propuestos por el IPCC indican cambios potenciales importantes en las precipitaciones y temperaturas de Andalucía en los próximos años. Es especialmente significativa la reducción de las precipitaciones en las, hasta ahora, lluviosas sierras de Cádiz y Málaga, y el incremento de las temperaturas en las zonas de alta montaña. Sin duda estos cambios van a afectar a la flora de la región, provocando migraciones, expansiones y/o contracciones de áreas de ocupación, extinciones, o desequilibrios en las interacciones biológicas. Conocer con cierta aproximación la magnitud de estos efectos puede proporcionarnos valiosa información para diseñar planes de conservación proactiva.

En los hábitats de alta montaña, representados por Sierra Nevada en este caso, las condiciones climáticas podrían cambiar considerablemente en un periodo corto de tiempo. En esta situación, el desplazamiento vertical de las isoterms obligará a las especies a seguir las montañas arriba, con la consiguiente reducción (o desaparición total) de la superficie de hábitat disponible. Las especies de altas cumbres carecen de posibilidades de migración vertical, y pueden ver seriamente amenazada su persistencia, ya sea por falta de tolerancia a las nuevas condiciones, o por una presión competitiva de las especies que, viniendo desde cotas más bajas, van a colonizar su hábitat. Estrategias de detección temprana de frentes de avance y retroceso como las descritas en el Capítulo 5 de esta Tesis pueden aplicarse a la planificación de las actuaciones de conservación.

No sólo en los hábitats de alta montaña se plantean problemas de conservación debidos al Cambio del Clima. En las zonas áridas de Andalucía Oriental la disminución y

redistribución espacial y temporal de las precipitaciones también puede afectar seriamente a especies de gran valor por su rareza biogeográfica, cambiando las condiciones de su hábitat en una medida superior incluso a la esperada para los hábitats de alta montaña. Parte de estos hábitats áridos no cuentan con ninguna figura de protección (aunque algunas de las especies que los habitan están bajo protección según las leyes de conservación de la naturaleza), por lo que tienen posibilidades de quedar fuera de los planes de conservación.

El cambio del clima en Andalucía podría llevar a la extinción a un porcentaje importante de la flora, aunque esta tendencia es muy variable entre escenarios. La dinámica climática y las características migratorias de las especies puede dar lugar una redistribución de especies que incremente la riqueza local, aunque al precio de perder especies raras o incluso emblemáticas. Los grandes perdedores en estos escenarios son los hábitats naturales de Andalucía Oriental, especialmente los ENP de Cazorla y Sierra Nevada (que podría dejar de ser el máximo exponente de biodiversidad en la Cuenca Mediterránea). Las simulaciones tienen un resultado común robusto: cuanto mayor es la temperatura y menor es la precipitación, mayor es la tasa de extinción. Este resultado debe interpretarse como una llamada de atención. Si la situación actual se dirige hacia donde pronostican los modelos climáticos, la erosión de diversidad biológica en Andalucía puede alcanzar cotas importantes en un corto periodo de tiempo. Estos extremos deberían confirmarse mejorando la metodología de simulación e investigando en profundidad los efectos de la competencia por el espacio (y otras interacciones) en la dinámica de dispersión geográfica de las especies.

La metodología de proyección de cambios en la distribución de especies está sujeta a una cascada de incertidumbres que comienza en el cálculo de las emisiones de los gases de efecto invernadero a partir de los escenarios del IPCC, sigue en los modelos físicos de simulación climática, y termina en los errores algorítmicos y biológicos de los MDE. Sin embargo, distintos investigadores de todo el mundo, usando distintas simulaciones del clima, distintos algoritmos de

modelado y distintas especies están llegando a la misma conclusión: hay un riesgo importante de desaparición de especies si el clima sigue cambiando a lo largo del presente siglo. Esta tendencia común en los resultados de los estudios también ha aparecido en las simulaciones de esta Tesis, y cuanto menos, debe tomarse como una llamada de atención para que nos atengamos a un principio de precaución.

#### 9.5 NUEVAS PERSPECTIVAS EN LOS MDE

Los MDE son modelos estáticos que no describen relaciones causa-efecto entre la presencia de la especie y las variables ambientales. Pero es posible mejorarlos añadiendo más datos.

La **migración**, con los distintos procesos que implica. Técnicamente se está resolviendo su incorporación a los MDE, pero faltan datos experimentales para calibrar los modelos. Apenas se conoce la dinámica dispersiva y la velocidad de migración (o cómo estas dependen de otros factores) para una gran mayoría de las especies. Es muy importante mejorar el conocimiento específico de estos aspectos al menos para especies clave de los distintos ecosistemas.

Las limitaciones de la **germinación**. La utilización de cámaras de germinación puede ofrecer datos de gran interés sobre las temperaturas y niveles de disponibilidad de agua que inhiben el proceso, para incorporarlos como restricciones en los MDE.

Otra de gran importancia, la **tolerancia** de los individuos a condiciones adversas. Es importante realizar experimentos para cuantificar la tolerancia de individuos adultos para conocer sus posibilidades de persistencia. Para especies herbáceas es posible conseguir estos datos de cámaras de cultivo, pero para especies de gran porte es mucho más difícil realizar este tipo de investigación. En estos casos sería posible trabajar con poblaciones translocadas o introducidas, para comparar su eficacia biológica con la de poblaciones naturales, y relacionar las diferencias con distancias entre poblaciones dentro del espacio ecológico.

La incorporación de las **interacciones biológicas** tanto positivas como negativas es probablemente una de las cuestiones más complicadas de incorporar a los MDE, porque requiere estudios minuciosos de estas relaciones a distintas escalas geográficas, y una reformulación del propio proceso de modelado, para trabajar simultáneamente con las especies que requieren la existencia de otras para sobrevivir.

Las **perturbaciones naturales**. Determinados tipos de perturbaciones, como los incendios forestales o las inundaciones, son susceptibles de ser modelados, y sus resultados pueden acoplarse a los MDE para mejorar los pronósticos de cambio en la distribución.

La **dinámica de la fragmentación antropógena** también es un factor que puede incorporarse, generando modelos de cambio del uso del suelo (como el presentado en el Capítulo 4, aunque hay algoritmos mucho más complejos y completos) para largos periodos de tiempo.

La **simulación de planes de conservación** sobre los MDE puede permitir una evaluación *a priori* de la eficacia potencial de las actuaciones de conservación (translocaciones, introducciones, manejo forestal, etc.), permitiendo diseñar estrategias más efectivas para mejorar las posibilidades de conservación de las especies.

## 9.6 PERO...

Hay mucho trabajo por hacer, pero apenas tenemos tiempo. El Cambio Global y los intereses económicos están acabando rápidamente con la biodiversidad en distintas partes de nuestro planeta. Y es necesario trabajar en dos frentes: la investigación teórica sobre los efectos potenciales de la situación actual en la biota, y la aplicación **real y efectiva** de las medidas de conservación. Este segundo frente pertenece por desgracia a la voluntad política, influenciada por un falso conflicto entre conservación de la biodiversidad y el desarrollo económico.

La comunidad de investigadores que trabajamos en conservación de la biodiversidad debemos dar un paso al frente, y utilizar los resultados de nuestros trabajos como argu-



mentos para tratar de mejorar las perspectivas reales de conservación de las especies. Para facilitar el acceso a este objetivo, es muy importante acercar los resultados de nuestras investigaciones no solo a los políticos, sino a la sociedad en general, que es la única con cierta capacidad para presionar a la voluntad política en un sentido u otro.

## CONCLUSIONES

---

### CONCLUSIONES

#### 10.1 METODOLOGÍA

1. La asimilación de las Tecnologías de la Información mejora sustancialmente las posibilidades de la investigación en biodiversidad, posibilitando la captura, procesamiento y análisis de grandes volúmenes de datos.
2. Para que esta asimilación sea productiva, es necesario que los investigadores conozcan las posibilidades y herramientas que la tecnología pone a su disposición.
3. Los programas informáticos de código abierto actualmente disponibles permiten ejecutar proyectos de modelado de gran envergadura con coste cero.
4. Los Modelos de Distribución de Especies tienen un amplio rango de aplicaciones en distintos campos de la Biología, especialmente en los campos de biodiversidad y conservación.
5. Una serie temporal de Modelos de Distribución de Especies en un contexto de cambio climático permite la localización de frentes de avance y retroceso de las poblaciones naturales, que puede aplicarse al establecimiento de redes de seguimiento y al diseño de planes de manejo y conservación.
6. No existe ninguna metodología estándar para evaluar modelos de distribución que esté plenamente aceptada. La evaluación de los modelos por especialistas en la distribución de las especies debería ser el camino a seguir.

7. La utilización de técnicas de ensamblado de modelos tiene más probabilidades de ofrecer buenos resultados que la aplicación de un único algoritmo. La técnica de ensamblado además ofrece la posibilidad de estudiar la variabilidad del conjunto de modelos, y aplicarla a una evaluación más robusta del resultado.
8. Los Modelos de Distribución de Especies presentan cierto número de incertidumbres que pueden limitarse añadiendo otros factores (además de las variables ambientales) que afectan a la distribución de las especies, como migración, fragmentación del hábitat, tolerancia de las especies o interacciones biológicas.
9. Actualmente hay más limitaciones en los datos ecológicos disponibles para calibrar los Modelos de Distribución que en las posibilidades que ofrecen la ciencia y la tecnología para perfeccionarlos.

#### 10.2 CAMBIO DE USOS DEL SUELO

1. Los Modelos de Distribución de Especies pueden aplicarse a la detección de conflictos entre conservación y desarrollo económico, facilitando la localización de poblaciones amenazadas por cambios en el uso del suelo.
2. La combinación de MaxEnt y GEOMOD permite predecir con gran precisión la expansión de invernaderos. Esta técnica puede aplicarse a la evaluación de riesgos para la biodiversidad en regiones en las que la agricultura bajo plástico es el motor de desarrollo económico.

#### 10.3 CAMBIO DEL CLIMA

1. Según las simulaciones climáticas, a lo largo del siglo XXI Andalucía podría volverse más árida y cálida, aunque con variaciones locales significativas según la región geográfica, el escenario y el intervalo de tiempo considerado.

2. El incremento de temperaturas en Sierra Nevada puede afectar negativamente al pastizal psicroxerófilo por desaparición total de su rango climático idóneo, que favorecería el ascenso altitudinal del enebro-piornal. En las cotas medias podría incrementarse la competencia por el espacio entre el encinar y el roble a favor del primero.
3. La idoneidad del hábitat en las poblaciones de especies raras y amenazadas de hábitats áridos de Andalucía Oriental puede disminuir drásticamente en los próximos años, aunque es necesario realizar nuevos estudios, porque hay pocos registros de presencia, y algunas de las especies tienen rangos de distribución geográfica de escala supra-continental.
4. Según una muestra de 198 especies vegetales, el Cambio del Clima podría extinguir entre un 13 y un 31 % de la flora andaluza según el escenario considerado.
5. El patrón de extinción previsto está perfectamente correlacionado con el incremento de la temperatura media máxima de verano y la disminución de la precipitación anual.
6. Los Espacios Naturales Protegidos de Andalucía Oriental estarían expuestos a una mayor erosión de su diversidad vegetal que los de Andalucía Occidental, que podrían incluso incrementar localmente su riqueza de especies.



Parte V

APÉNDICE



## EVALUACIÓN DE ENSAMBLADOS

---

Cada especie que se introduce en MODPLAN está representada por una muestra de presencias que tiene dos propiedades: **número de presencias** y **área de extensión**. MODPLAN parte aleatoriamente la muestra según una proporción determinada por el parámetro PORCENTAJE EVALUACIÓN en un conjunto de puntos de calibrado, que se utilizan para entrenar los algoritmos, y un conjunto de puntos de evaluación. Éstos son lo que MODPLAN utiliza, junto a 10000 puntos aleatorios, para calcular las métricas e evaluación. Posteriormente el sistema calibra un nuevo modelo con todos los puntos disponibles. De este modelo se extraen los valores de idoneidad y desviación de las presencias y puntos aleatorios (que describen distintas propiedades del modelo) para generar los datos del informe final. A este modelo se le asignan directamente los valores de las métricas de evaluación del modelo anterior.

La descrita es una metodología propuesta por Rencher (1995), que ya apareció comentada en el Capítulo 2. Sin embargo, el proceso que involucra, cuando se trabaja en un entorno de modelado masivo, incrementa el tiempo de computación aproximadamente un 25 %, un consumo de recursos importante cuando el tiempo de ejecución es un factor a tener en cuenta. Además, cuando se trabaja con tamaños de muestra pequeños, los valores obtenidos pueden ser poco representativos del ajuste real del modelo.

Tratando de superar estas dificultades, en este anexo presento un razonamiento que justifica (con matices) la evaluación de los modelos utilizando únicamente los valores de idoneidad y desviación de todos los puntos de calibrado, y por tanto, la eliminación de la fase inicial de partición de datos, calibrado y evaluación.



Para tratar de elaborar un criterio de evaluación tabulé los datos de evaluación y las propiedades de las muestras de todas las especies de trabajo del Capítulo 7. Las variables de interés identificadas fueron:

- Propiedades del modelo: *extensión* (área de extensión de los puntos de presencia), y *presencias* (número total de registros de presencia)
- Métricas de evaluación: *AUCg*, *AUCI* y *CIDp*
- Propiedades del modelo: *I\_p\_total* (promedios de idoneidad de todos los puntos de presencia sobre el modelo recalibrado) y *D\_p\_total* (promedios de desviación de los todos los puntos de presencia sobre el modelo recalibrado).

*Descarté la métrica IC por considerarla poco útil. Una prueba fallida...*

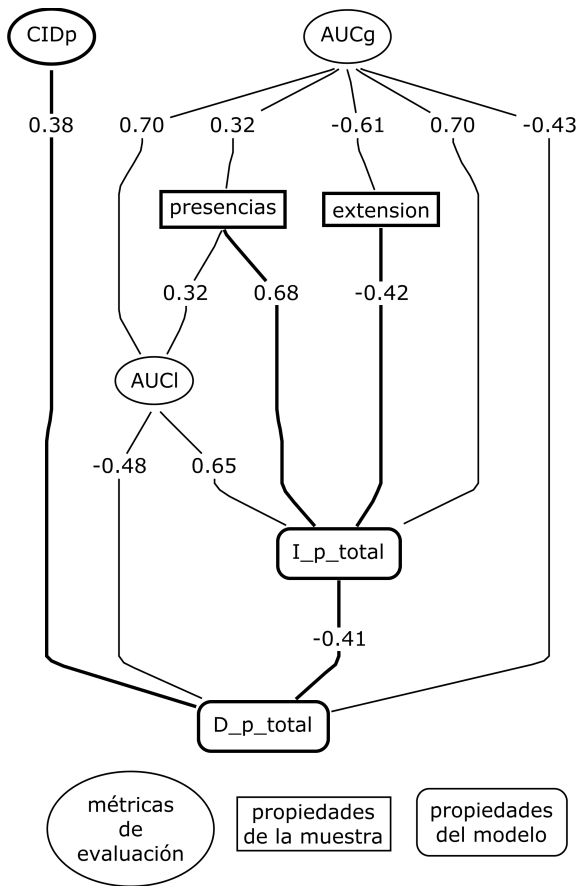
Para valorar las relaciones dentro del conjunto de variables exploré los datos mediante distintas herramientas (matrices de correlación, matrices de diagramas de dispersión, análisis de correlación de Spearman (con test de significación), normalización (funciones arcoseno y logaritmo, según el caso).

Con los resultados del análisis construí un diagrama de relaciones entre variables (1) como herramienta de partida para justificar y desarrollar una nueva métrica de evaluación basada en los valores de idoneidad y desviación de todos los puntos de presencia sobre el modelo.

*AUCg* y *AUCI* mostraron correlación entre ellas (0.70), pero ninguna con *CIDp*, que solo mostró correlación positiva (0.38) con *D\_p\_total* (ver Figura 1).

La métrica de evaluación *AUCg* resultó muy sensible a las propiedades de la muestra, con una correlación significativa ( $p < 0.005$ ) con *extensión* (-0.61) y *presencias* (0.32). *AUCI* mostró la misma correlación con *presencias*, pero ninguna correlación significativa con *extensión*, demostrando ser menos sensible a las características de la muestra.

El diagrama de relaciones revela la gran influencia de las propiedades de la muestra sobre los valores de idoneidad



**Figura 1:** Diagrama de relaciones ( $\rho$  de Spearman, significativos según  $p < 0.05$ ) entre métricas de evaluación, propiedades de la muestra, y valores de las presencias sobre el modelo. La sección en negrita es la que resulta de eliminar las métricas AUCg y AUCI del diagrama.

de los puntos de presencia, especialmente *presencias* (0.68). Sin embargo, la desviación de los puntos de presencia no presenta ninguna relación con *presencias* y *extensión*,

La métrica de evaluación AUCg depende significativamente (relación negativa) del área de presencia, porque las especies con mayor extensión tienen mayor probabilidad de que sus poblaciones estén alejadas en el espacio ambiental. El modelo resultante debe cubrir amplios gradientes que ofrecen superficie a los puntos aleatorios en el test AUC, bajando su valor. Esta cuestión ya la apuntaba [Lobo et al.](#)

(2008), y en acuerdo con este autor, AUCg solo se presenta como referencia, por ser la medida más parecida a la empleada en otros estudios. AUCl presenta limitaciones similares, y no he encontrado una ventaja que justifique su aplicación como sustituto de AUCg.

CIDp es el comienzo de otra historia. Independiente de las propiedades de la muestra (característica interesante), se basa en una idea sencilla: cuanto menor es la desviación en los puntos de presencia, mayor es el consenso de los modelos respecto a sus valores de idoneidad. Pero esta métrica tiene un problema grave: para un promedio de idoneidad concreto, existe un máximo teórico de desviación. En las zonas de valores altos de idoneidad, este máximo teórico tiene un valor bajo; en los valores medios llega a su máximo, y en los valores bajos solo es posible una baja idoneidad.

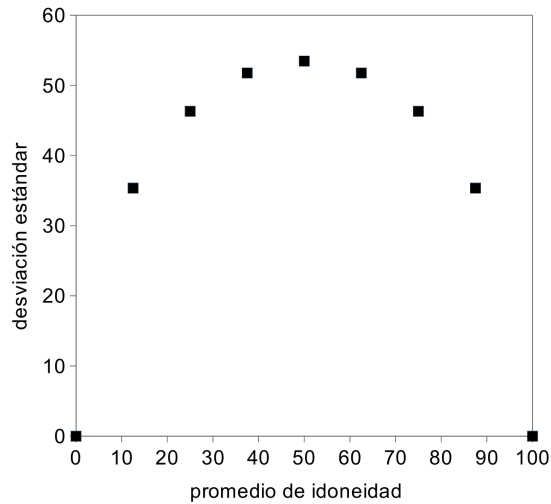
Este comportamiento puede comprenderse fácilmente a partir de la siguiente explicación. Si una celda tiene un promedio de idoneidad igual a 100 eso quiere decir que todos los modelos le han dado el mismo valor, y en ese punto la desviación estándar es cero. Suponiendo que a la misma celda 7 de los algoritmos le han dado un valor 100, y el algoritmo restante le ha dado un valor 0, la desviación estándar resultante es 35.36. El caso de máxima desviación (53.45) podemos encontrarlo cuando cuatro modelos dan a la celda valor 100, y cuatro modelos le dan valor 0. Calculando todos los valores máximos posibles de desviación correspondientes a cada nivel de idoneidad y dibujando los puntos en un gráfico puede verse que ambos, idoneidad y desviación, se relacionan mediante una bonita curva (ver Figura 2).

La relación se ajusta a una curva parabólica descrita por la siguiente ecuación:

$$D = 2,2651 \cdot I + 1,0648 \cdot I + (-0,010649 \cdot I^2) \quad (A.1)$$

Donde **D** es la desviación estándar e **I** es el promedio de idoneidad del hábitat.

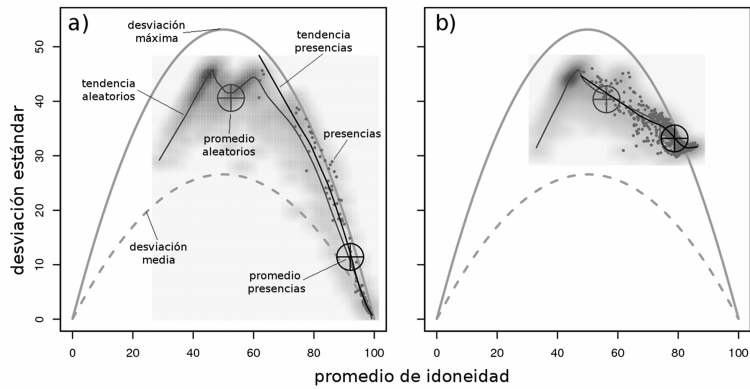
Esta curva describe el valor máximo que puede tomar la desviación estándar para cada valor de idoneidad.



**Figura 2:** Puntos que muestran la tendencia de la relación entre promedio de idoneidad y desviación estándar.

Conociendo el comportamiento de la curva, es posible llegar a la conclusión de que CIDp no es un buen indicador del funcionamiento del modelo, porque solo tiene valores en el límite superior de idoneidad. En este momento, puedo considerarla una métrica fracasada, pero que ha abierto puertas a una cuestión interesante.

Volviendo a la Figura 1, y ya descartadas las tres métricas de evaluación, solo quedan como recursos las propiedades del modelo, esto es, los valores de idoneidad y la desviación de sus celdas. Estas propiedades pueden calcularse para todo el modelo, tomando una muestra aleatoria de puntos que describiría el comportamiento del modelo en todo el espacio geográfico, y para las localidades de presencia de las especies. Comparando los valores de ambos conjuntos de puntos en el espacio descrito por la parábola de desviación estándar, es posible determinar hasta que punto el modelo de distribución es capaz de discriminar entre registros de presencia y puntos aleatorios. Mejor lo describiré con dos ejemplos en la 3.



**Figura 3:** Propuesta de evaluación visual de modelos. Las nubes oscuras representan los 10000 puntos aleatorios. La figura **a** muestra un modelo bien ajustado, en el que el promedio de las presencias y el promedio de los puntos aleatorios están bien diferenciados. La figura **b** muestra los resultados de un modelo mal ajustado, en el que el promedio de desviación de las presencias es muy alto, cercano al promedio de desviación de los puntos aleatorios.

## BIBLIOGRAFÍA

---

- Aarssen L.W., 1983. Ecological combining ability and competitive combining ability in plants: toward a general evolutionary theory of coexistence in systems of competition. *American Naturalist*, 122(6):707–731. (Citado en la página 14.)
- Aguilera F., 2006. Predicción del crecimiento urbano mediante sistemas de información geográfica y modelos basados en Autómatas Celulares. *Geofocus*, 6:81–112. (Citado en la página 56.)
- Anónimo, 1992. *Directive 92/43 of the Council of the European Community on the Conservation of Habitats and Wild Fauna and Flora*. (Citado en las páginas 41, 42, y 53.)
- Anónimo, 2003. *Ley de 8/2003 de 28 de octubre, de la Flora y la Fauna Silvestres*. (Citado en las páginas 41 y 43.)
- Araújo M.B. y New M., 2007. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, 22(1):42–47. (Citado en las páginas 99, 103, y 142.)
- Araújo M.B. y Williams P.H., 2000. Selecting areas for persistence using occurrence data. *Biological Conservation*, 96: 331–345. (Citado en la página 19.)
- Araújo M.B., Williams P.H., y Fuller R.J., 2002. Dynamics of extinction and the selection of nature reserves. *Proceedings of the Royal Society B: Biological Sciences*, 269(1504):1971–1980. (Citado en la página 40.)
- Araújo M.B., Cabeza M., Thuiller W., Hannah L., y Williams P.H., 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, 10:1618–1626. (Citado en la página 87.)
- Araújo M.B., Pearson R.G., Thuiller W., y Erhard M., 2005a. Validation of species climate impact models under climate

- change. *Global Change Biology*, 11(9):1504–1513. (Citado en las páginas 16, 99, y 191.)
- Araújo M.B., Whittaker R.J., Ladle R.J., y Erhard M., 2005b. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, 14(6):529–538. (Citado en la página 103.)
- Araújo M.B., Thuiller W., y Pearson R.G., 2006. Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, 33(10):1712–1728. (Citado en las páginas 12 y 87.)
- Arntzen J.W., 2006. From descriptive to predictive distribution models: A working example with Iberian amphibians and reptiles. *Frontiers in Zoology*, 3(8):1–11. (Citado en la página 45.)
- Austin M.P., 1980. Searching for a model for use in vegetation analysis. *Vegetatio*, 42(1-3):11–21. (Citado en las páginas 17 y 18.)
- Austin M.P., 1985. Continuum concept, ordination methods, and niche theory. *Annual review of ecology and systematics*, 16:39–61. (Citado en la página 8.)
- Austin M.P., 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, 200(1-2):1–19. (Citado en la página 99.)
- Austin M.P. y Smith T.M., 1989. A new model for the continuum concept. *Vegetatio*, 83(1-2):35–47. (Citado en la página 7.)
- Benito B.M. y Peñas J., 2008. *Greenhouses, land use change, and predictive models: MaxEnt and Geomod working together*, capítulo 11, páginas 297–317. Environmental Science and Engineering. Springer-Verlag, Berlin. (Citado en la página 37.)
- Benito B.M., Martínez-Ortega M.M., Muñoz L.M., Lorite J., y Peñas J., 2009. Assessing extinction-risk of endangered

- plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses. *Biodiversity and Conservation*, 18(9):2509–2520. (Citado en la página 37.)
- Benito-Garzón M., 2006. *El efecto del Cambio Climático sobre las distribuciones de los bosques ibéricos: pasado, presente y futuro*. Tesis Doctoral, Universidad Autónoma de Madrid. (Citado en las páginas 104 y 105.)
- Benito-Garzón M., Sánchez de Dios R., y Sáinz Ollero H., 2008. Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, 11(2):169–178. (Citado en la página 87.)
- Blanca G., Cueto M., Martínez-Lirola M.J., y Molero-Mesa J., 1998. Threatened vascular flora of Sierra Nevada (Southern Spain). *Biological Conservation*, 85:269–285. (Citado en las páginas 88, 89, 90, y 144.)
- Blanca G., López M.R., Lorite J., M.J. Martínez-Lirola, Molero-Mesa J., Quintas S., Ruiz M., Varo M.A., y Vidal S., 2000. *Flora amenazada y endémica de Sierra Nevada*. Universidad de Granada, Granada. (Citado en la página 144.)
- Broennimann O., Thuiller W., Hughes G., Midgley G., Alkemade J.M.R., y Guisan A., 2006. Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology*, 12(6):1079–1093. (Citado en la página 99.)
- Brown J.H., 1984. On the relationship between abundance and distribution of species. *American Naturalist*, 124(2): 255–279. (Citado en las páginas 7, 12, 14, 15, y 21.)
- Brown J.H. y Gibson A.C., 1983. *Biogeography*. Mosby, St. Louis. (Citado en las páginas 9 y 14.)
- Brunet J., von Oheimb G., y Diekmann M., 2000. Factors influencing vegetation gradients across ancient-recent woodland borderlines in southern Sweden. *Journal of Vegetation Science*, 11(4):515–524. (Citado en la página 164.)



- Brunet M., Casado M.J., Castro M., Galán P., López J. A., Martín J.M., Pastor A., Petisco E., Ramos P., Ribalaygua J., y Rodríguez E., 2007. Generación de escenarios regionalizados de cambio climático para España. Technical report, Ministry of Environment, Government of Spain. (Citado en la página 90.)
- Cabezudo B., Talavera S., Blanca G., Salazar C., Cueto M., Valdés B., Hernández J., Herrera C.M., Rodríguez C., y Navas D., 2005. Lista Roja de la Flora Vasculare de Andalucía. *Consejería de Medio Ambiente de la Junta de Andalucía, Sevilla*. (Citado en la página 43.)
- Carpenter J., Gillison G., y Winter A.N., 1993. DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, 2(6):667–680. (Citado en la página 44.)
- Castro J., 2000. *Dinámica de la regeneración de los pinares autóctonos de Pino silvestre (Pinus sylvestris L. var. nevadensis Chirst) de Sierra Nevada y Sierra de Baza*. Tesis Doctoral, Universidad de Granada, Granada. (Citado en la página 166.)
- Castro J., Zamora R., Hódar J.A., Gómez J.M., y Gómez-Aparicio L., 2004. Benefits of Using Shrubs as Nurse Plants for Reforestation in Mediterranean Mountains: A 4-Year Study. *Restoration Ecology*, 12(3):352–358. (Citado en la página 98.)
- Chase J.M. y Leibold M.A., 2003. *Ecological niches: linking classical and contemporary approaches*. University of Chicago Press, Chicago, Illinois, USA. (Citado en la página 12.)
- Chen G. y Peterson A.T., 2002. Prioritization of areas in China for biodiversity conservation based on the distribution of endangered bird species. *Bird Conserv. Int.*, 12:197–209. (Citado en la página 26.)
- Clark J.S., Fastie C., Hurtt G., Jackson S.T., Johnson C., King G.A., Lewis M., Lynch J., Pacala S., Prentice C., Schupp E.W., Webb III T., y Wyckoff P., 1998. Reid's paradox of

- rapid plant migration. *BioScience*, 48(1):13–24. (Citado en la página 163.)
- Clements F.E., 1936. Nature and structure of the climax. *J. Ecol.*, 24(1):252–284. (Citado en la página 7.)
- Collins S.L., Glenn S.M., y Roberts D.W., 1993. The hierarchical continuum concept. *Journal of Vegetation Science*, 4 (2):149–156. (Citado en la página 8.)
- Costa L.J., Wolfe G.C., Shepard C., Caldwell D.B., y Vitt J.P., 2008. Detecting the influence of climatic variables on species distributions: A test using GIS niche-based models along a steep longitudinal environmental gradient. *Journal of Biogeography*, 35(4):637–646. (Citado en la página 52.)
- Crisp M.D., Arroyo M.T.K., Cook L.G., Gandolfo M.A., Jordan G.J., McGlone M.S., Weston P.H., Westoby M., Wilf P., y Linder H.P., 2009. Phylogenetic biome conservatism on a global scale. *Nature*, 458(7239):754–756. (Citado en la página 14.)
- Davis M.B. y Shaw R.G., 2001. Range shifts and adaptive responses to quaternary climate change. *Science*, 292(5517): 673–679. (Citado en las páginas 98, 163, y 165.)
- Davis M.B., Shaw R.G., y Etterson J.R., 2005. Evolutionary responses to changing climate. *Ecology*, 86(7):1704–1714. (Citado en la página 14.)
- de Candolle A., 1855. *Géographie botanique raisonnée*. (Citado en la página 17.)
- Dixon R.K., Brown S., Houghton R.A., Solomon A.M., Trexler M.C., y Wisniewski J., 1994. Carbon pools and flux of global forest ecosystems. *Science*, 263(5144):185–190. (Citado en la página 55.)
- Donoghue M.J. y Smith S.A., 2004. Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philos. Trans. R. Soc. London Ser. B*, 359:1633–44. (Citado en la página 15.)

- Douglas T., Critchley D., y Park G., 1996. The deintensification of terraced agricultural land near Trevélez, Sierra Nevada, Spain. *Global Ecology and Biogeography Letters*, 5: 258–270. (Citado en la página 100.)
- Eastman J.R., 2006. Idrisi Andes. Guide to GIS and image processing. *Clark University, Worcester*, páginas 189–214. (Citado en la página 44.)
- Elith J., Graham C.H., Anderson R.P., Dudík M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.McC.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S., Wisz M.S., y Zimmermann N.E., 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29(2):129–151. (Citado en las páginas 44, 45, 50, 51, 60, y 91.)
- Elith S. y Barry J., 2006. Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43(3):413–423. (Citado en la página 44.)
- Elston S.T. y Buckland D.A., 1993. Empirical models for the spatial distribution of wildlife. *Journal of Applied Ecology*, 30(3):478–495. (Citado en la página 40.)
- Elton C., 1927. *Animal Ecology*. Segdwick and Jackson. London. (Citado en la página 10.)
- Engler R., Guisan A., y Rechsteiner L., 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2):263–274. (Citado en la página 51.)
- Engler R., Randin C.F., Vittoz P., Czaka T., Beniston M., Zimmermann N.E., y Guisan A., 2009. Predicting future distributions of mountain plants under climate change: Does dispersal capacity matter? *Ecography*, 32(1):34–45. (Citado en las páginas 160, 164, y 165.)

- Etterson J.R. y Shaw R.G., 2001. Constraint to adaptive evolution in response to global warming. *Science*, 294: 151–154. (Citado en la página 190.)
- Evangelista P.H., Kumar S., Stohlgren T.J., Jarnevich C.S., Crall A.W., Norman III J.B., y Barnett D.T., 2008. Modelling invasion for a habitat generalist and a specialist plant species. *Diversity and Distributions*, 14(5):808–817. (Citado en la página 26.)
- Farber O. y Kadmon R., 2003. Assessment of alternative approaches for bioclimatic modelling with special emphasis on the Mahalanobis distance. *Ecol Model*, 160:115–130. (Citado en la página 44.)
- Fernández J.M., 1996. El volcanismo calco-alcalino en el Parque Natural de Cabo de Gata-Níjar (Almería). Estudio volcanológico y petrológico. *Monografías Medio Natural n° 2. Sociedad Almeriense de Historia Natural - Consejería de Medio Ambiente de la Junta de Andalucía*. (Citado en la página 42.)
- Fielding A.H. y Bell J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1): 38–49. (Citado en la página 45.)
- Fitzpatrick M.C., Gove A.D., Sanders N.J., y Dunn R.R., 2008. Climate change, plant migration, and range collapse in a global biodiversity hotspot: the Banksia (Proteaceae) of Western Australia. *Global Change Biology*, 14(6):1337–1352. (Citado en las páginas 87, 91, 96, y 99.)
- Flato G.M. y Boer G.J., 2001. Warming asymmetry in climate change simulations. *Geophysical Research Letters*, 28:195–198. (Citado en la página 90.)
- Franklin J., 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, 19(4):474–499. (Citado en la página 32.)

- García D., Zamora R., Hódar J.A., y Gómez J.M., 1999. Age structure of *Juniperus communis* L. in the Iberian peninsula: Conservation of remnant populations in Mediterranean mountains. *Biological Conservation*, 87:215–220. (Citado en la página 98.)
- García-Aguilar J.M., 2000. Late Neogene to recent continental history and evolution of the Guadix-Baza Basin (SE Spain). *Rev. Soc. Geol. España*, 13(1):65–77. (Citado en la página 143.)
- Gastón K.J., 2003. *The structure and dynamics of geographic ranges*. Oxford University Press. (Citado en la página 16.)
- Gastón K.J. y Mound L.A., 1993. Taxonomy, hypothesis testing and the biodiversity crisis. *Proceedings of the Royal Society B: Biological Sciences*, 251(1331):139–142. (Citado en la página 26.)
- Gauch H.G.Jr. y Whittaker R.H., 1972. Coenocline simulation. *Ecology*, 53:446–451. (Citado en la página 8.)
- Ghalambor C.K., Mckay J.K., Carroll S.P., y Reznick D.N., 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, 21(3):394–407. (Citado en la página 98.)
- Giorgi F., Hurrell J.W., Marinucci M.R., y Beniston M., 1997. Elevation dependency of the surface climate change signal: A model study. *Journal of Climate*, 10:288–296. (Citado en la página 87.)
- Gleason H.A., 1926. The individualistic concept of the plant association. *Bull. Torrey Bot. Club*, 53(1):7–26. (Citado en las páginas 7 y 8.)
- Grabherr G., Gottfried M., y Pauli H., 1994. Climate effects on mountain plants [7]. *Nature*, 369:448. (Citado en las páginas 86, 88, y 96.)
- Grinnell J., 1917. The niche-relationships of the California Thrasher. *Auk*, 34:427–433. (Citado en la página 10.)

- Guisan A. y Thuiller W., 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9):993–1009. (Citado en las páginas 19 y 40.)
- Guisan A. y Zimmermann N.E., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135 (2-3):147–186. (Citado en las páginas 17, 18, 40, 43, 52, y 56.)
- Guisan A., Broennimann O., Engler R., Vust M., Yoccoz N.G., Lehmann A., y Zimmermann N.E., 2006. Using niche-based models to improve the sampling of rare species. *Conservation Biology*, 20(2):501–511. (Citado en la página 40.)
- Gómez J.M., García D., y Zamora R., 2003. Impact of vertebrate acorn- and seedling-predators on a Mediterranean *Quercus pyrenaica* forest. *Forest Ecology and Management*, 180:125–134. (Citado en la página 98.)
- Hale S.S. y Hollister J.W., 2009. Beyond data management: how ecoinformatics can benefit environmental monitoring programs. *Environmental Monitoring Assessment*, 150:227–235. (Citado en la página 25.)
- Hampe A. y Petit R.J., 2005. Conserving biodiversity under climate change: The rear edge matters. *Ecology Letters*, 8 (5):461–467. (Citado en las páginas 165 y 166.)
- Hanski I., 1991. Single-species metapopulation dynamics: concepts, models and observations. *Biological Journal of the Linnean Society*, 42(1-2):17–38. (Citado en la página 8.)
- Hengl T., Heuvelink G.B.M., y Stein A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 120(1-2):75–93. (Citado en la página 119.)
- Hernández-Bermejo J.E. y Clemente M., 1994. *Protección de la flora de Andalucía*. Junta de Andalucía. (Citado en la página 168.)
- Higgins S.I. and Nathan R. y Cain M.L., 2003. Are long-distance dispersal events in plants usually caused by

- nonstandard means of dispersal? *Ecology*, 84(8):1945–1956. (Citado en la página 164.)
- Hijmans R.J. y Graham C.H., 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, 12(12): 2272–2281. (Citado en la página 91.)
- Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., y Jarvis A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978. (Citado en la página 97.)
- Hirzel A.H., Hausser J., Chessel D., y Perrin N., 2002. Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? *Ecology*, 83(7): 2027–2036. (Citado en la página 44.)
- Hirzel A.H., Hausser J., y Perrin N., 2004. Biomapper 3.0. *Laboratory for Conservation Biology. University of Bern.* (Citado en la página 44.)
- Hirzel A.H., Hausser J., y Perrin N., 2006. Biomapper 3.2. *Laboratory for Conservation Biology, Department of Ecology and Evolution, University of Lausanne, Switzerland.* (Citado en las páginas 51 y 58.)
- Hogeweg P. y Hesper B., 1978. Interactive instruction on population interactions. *Computers in Biology and Medicine*, 8(4):319–327. (Citado en la página 23.)
- Honnay O., Verheyen K., Butaye J., Jacquemyn H., Bossuyt B., y Hermy M., 2002. Possible effects of habitat fragmentation and climate change on the range of forest plant species. *Ecology Letters*, 5(525-530). (Citado en la página 164.)
- Hughes L., 2000. Biological consequences of global warming: is the signal already apparent? *Trends in Ecology & Evolution*, 15:56–61. (Citado en la página 86.)
- Hulme M. y Sheard N., 1999. Climate Change Scenarios for the Iberian Peninsula. Technical report, Climatic Research Unit, Norwich, UK. (Citado en la página 95.)

- Humboldt A., 1807. *Essai sur la Geographie des plantes*. París. (Citado en la página 17.)
- Hutchinson E.G., 1978. Erratum: An introduction to population ecology. *Science*, 202(4374):1269. (Citado en la página 11.)
- Hutchinson G.E., 1957. Concluding remarks. *Cold Spring Harbor Symp. Q. Biol.*, 22:415–427. (Citado en las páginas 10 y 16.)
- IPCC, 2001. Third Special Report. Climate Change 2001: The Scientific basis. Technical report, IPCC, Cambridge University Press, Cambridge. (Citado en la página 87.)
- Jackson D.A., 1995. PROTEST: a procrustean randomization test of community environment concordance. *Ecoscience*, 2:297–303. (Citado en la página 63.)
- Jantz C.A., Goetz S.J., y Shelley M.K., 2004. Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. *Environmental Planning B: Planning and Design*, 31(2):251–271. (Citado en la página 56.)
- Jaynes E.T., 1957. Information theory and statistical mechanics. *Physical Review*, 106:620–630. (Citado en la página 91.)
- Johnson C.J. y Gillingham M.P., 2005. An evaluation of mapped species distribution models used for conservation planning. *Environmental Conservation*, 32(2):117–128. (Citado en la página 56.)
- Joly F., 1988. *La Cartografía*. Oikos-Tau. (Citado en la página 28.)
- Jones P.D., New M., Parker D.E., Martin S., y Rigor I.G., 1999. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, 37:173–199. (Citado en la página 86.)



- Kearney M., 2006. Habitat, environment and niche: What are we modelling? *Oikos*, 115(1):186–191. (Citado en la página 9.)
- Kirkpatrick M. y Barton N.H., 1997. Evolution of a species' range. *American Naturalist*, 150(1):1–23. (Citado en la página 14.)
- Kolasa J., 1989. Ecological systems in hierarchical perspective: breaks in community structure and other consequences. *Ecology*, 70(1):36–47. (Citado en la página 8.)
- Krishtalka L. y Humphrey P.S., 2000. Can natural history museums capture the future? *BioScience*, 50(7):611–617. (Citado en la página 24.)
- L. Kullman, 2002. Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, 90:68–77. (Citado en la página 86.)
- Lamont B.B. y Connell S.W., 1996. Biogeography of *Banksia* in southwestern Australia. *Journal of Biogeography*, 23:295–309. (Citado en la página 97.)
- Lenihan J.M., 1993. Ecological response surfaces for North American boreal tree species and their use in forest classification. *Journal of Vegetation Science*, 4(5):667–680. (Citado en la página 16.)
- Leohle C., 1983. Evaluation of theories and calculation tools in ecology. *Ecological Modelling*, 19(4):239–247. (Citado en la página 28.)
- Lischke H., Zimmermann N.E., Bolliger J., Rickebusch S., y Löffler T.J., 2006. TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling*, 199(4):409–420. (Citado en las páginas 164 y 165.)
- Liu C., Berry P.M., Dawson T.P., y Pearson R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28(3):385–393. (Citado en la página 121.)

- Lobo J.M., Jiménez-valverde A., y Real R., 2008. AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17(2): 145–151. (Citado en las páginas 161 y 213.)
- Loiselle B.A., Howell C.A., Graham C.H., Goerck T., Brooks J.M., Smith K.G., y Williams P.H., 2003. Avoiding pitfalls of using species distribution models in conservation planning. *Conserv Biol*, 17:1591–1600. (Citado en la página 51.)
- Lorite J., Gómez F., Mota J.F., y Valle F., 2007. Orophilous plant communities of Baetic range in Andalusia (southeastern Spain): Priority altitudinal-island for conservation. *Phytocoenologia*, 37(3-4):625–644. (Citado en las páginas 89 y 100.)
- Lorite J., Salazar C., Peñas J., y Valle F., 2008. Phytosociological review on the forests of *Quercus pyrenaica* Willd. *Acta Botanica Gallica*, 155(2):219–233. (Citado en la página 89.)
- Mann M.E., 2007. Climate over the past two millennia. *Annual Review of Earth and Planetary Sciences*, 35:111–136. (Citado en las páginas 86 y 87.)
- Marmion M., Hjort J., Thuiller W., y Luoto M., 2009. Statistical consensus methods for improving predictive geomorphology maps. *Computers and Geosciences*, 35(3):615–625. (Citado en las páginas 103, 117, 142, y 158.)
- Martínez-Fernández J. y Esteve M.A., 2004. Assessing the sustainability of Mediterranean intensive agricultural systems through the combined use of dynamic system models, environmental modelling and geographical information systems. *Quaddus MA, Siddique MAB (eds) Handbook of Sustainable Development Planning. Studies in Modelling and Decision Support*, páginas 215–248. (Citado en la página 41.)
- Martínez-Meyer E., Peterson A.T., y Hargrove W.W., 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity.

- Global Ecology and Biogeography*, 13(4):305–314. (Citado en la página 13.)
- Melendo M., Giménez E., Cano E., Gómez-Mercado F., y Valle F., 2003. The endemic flora in the south of the Iberian Peninsula: taxonomic composition, biological spectrum, pollination, reproductive mode and dispersal. *Flora*, 198: 260–276. (Citado en la página 168.)
- Miranda J.D., Padilla F.M., Lázaro R., y Pugnaire F.I., 2009. Do changes in rainfall patterns affect semiarid annual plant communities? *Journal of Vegetation Science*, 20:269–276. (Citado en la página 193.)
- Mitchell T.D. y Hulme M., 1999. Predicting regional climate change: Living with uncertainty. *Progress in Physical Geography*, 23:57–78. (Citado en la página 99.)
- Mota J.F., Peñas J., Castro H., Cabello J., y Guirado J.S., 1996. Agricultural development vs biodiversity conservation: The Mediterranean semiarid vegetation in El Ejido (Almería, southeastern Spain). *Biodiversity and Conservation*, 5 (12):1597–1617. (Citado en la página 41.)
- Mota J.F., Pérez-García F.J., Peñas J., Cabello J., y Cueto M., 2003. *Flora Amenazada de la provincia de Almería*, capítulo La flora amenazada de Almería en tablas y fichas. Especies en peligro: casos estudiados, páginas 227–245. Universidad de Almería-IEA, Almería. (Citado en la página 42.)
- Mota J.F., Cabello J., Cerrillo M.I., y Rodríguez-Tamayo M.L., 2004. *Los subdesiertos de Almería: naturaleza de cine*. Consejería de Medio Ambiente. Junta de Andalucía. (Citado en la página 42.)
- Murphy J., 2000. Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology*, 20:489–501. (Citado en la página 87.)
- Médail F. y Quèzel P., 1999. Biodiversity Hotspots in the Mediterranean Basin: Setting Global Conservation Priorities. *Conservation Biology*, 13:1510–1513. (Citado en las páginas 88 y 141.)

- Ninyerola M., Pons X., y Roure J.M., 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, 20(14):1823–1841. (Citado en las páginas 44 y 91.)
- Nogués-Bravo D. y Araújo M.B., 2008. Climate change in Mediterranean mountains during the 21st century. *Ambio*, 37:149–158. (Citado en las páginas 87, 88, y 95.)
- Parmesan C., 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, 37:637–669. (Citado en la página 163.)
- Pearman P.B., Guisan A., Broennimann O., y Randin C.F., 2008. Niche dynamics in space and time. *Trends in Ecology & Evolution*, 23(3):149–158. (Citado en las páginas 13, 14, y 98.)
- Pennington D.D., Higgins D., Peterson A.T., Jones M.B., Ludascher B., y Bowers S., 2007. *Workflows for e-Science*, capítulo Ecological Niche Modelling Using the Kepler Workflow System. Springer, Londres. (Citado en la página 105.)
- Peters D. y Stockwell D., 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13 (2):143–158. (Citado en la página 44.)
- Peterson A.T., Soberón J., y Sánchez-Cordero V., 1999. Conservatism of ecological niches in evolutionary time. *Science*, 285(5431):1265–1267. (Citado en las páginas 13 y 14.)
- Peterson A.T., Sanchez-Cordero V., Soberon J., Bartley J., Buddemeier R.W., y Navarro-Siguenza A.G., 2001. Effects of global climate change on geographic distributions of Mexican Cracidae. *Ecological Modelling*, 144(1):21–30. (Citado en la página 164.)
- Peterson A.T., Ortega-Huerta M.A., Bartley J., Sánchez-Cordero V., Soberón J., Buddemeler R.H., y Stockwell

- D.R.B., 2002. Future projections for Mexican faunas under global climate change scenarios. *Nature*, 416(6881): 626–629. (Citado en la página 13.)
- Peterson A.T., Papes M., y Kluza D.A., 2003. Predicting the potential invasive distributions of four alien plant species in North America. *Weed Science*, 51(6):863–868. (Citado en la página 26.)
- Peterson A.T., Papes M., y Soberón J., 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213:63–72. (Citado en la página 159.)
- Phillips S.J. y Dudík M., 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(31(2)):161–175. (Citado en las páginas 91 y 159.)
- Phillips S.J., Anderson R.P., y Schapire R.E., 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(190):231–259. (Citado en las páginas 44, 50, 51, 60, 61, 64, y 91.)
- Polechová J. y Storch D., 2009. *Encyclopedia of Ecology*, capítulo Ecological Niche. Elsevier. (Citado en la página 12.)
- Pontius R.G.Jr. y Chen H., 2006. GEOMOD Modeling. *Idrisi Andes Help Contents, Massachusetts, Clark University*. (Citado en la página 62.)
- Pontius R.G.Jr. y Pacheco P., 2004. Calibration and validation of a model of forest disturbance in the Western Ghats, India 1920–1990. *GeoJournal*, 61(4):325–334. (Citado en la página 56.)
- Pontius R.G.Jr., Cornell J.D., y Hall C.A.S., 2001. Modeling the spatial pattern of land-use change with GEOMOD2: Application and validation for Costa Rica. *Agriculture, Ecosystems and Environment*, 85(1-3):191–203. (Citado en las páginas 56 y 61.)
- Posillico M., Meriggi A., Pagnin E. and Lovari S., y Russo L., 2004. A habitat model for brown bear conservation

- and land use planning in the central Apennines. *Biological Conservation*, 118(2):141–150. (Citado en la página 56.)
- Preston J.T., Knick K.L., y Rotenberry S.T., 2006. GIS-based niche modeling for mapping species habitat. *Ecology*, 87(6):1458–1464. (Citado en la página 40.)
- Prinzing A., Durka W., Klotz S., y Brandl R., 2001. The niche of higher plants: Evidence for phylogenetic conservatism. *Proceedings of the Royal Society B: Biological Sciences*, 268(268):2383–2389. (Citado en las páginas 14 y 98.)
- Pulliam R., 2000. On the relationship between niche and distribution. *Ecology Letters*, 3:349–361. (Citado en la página 19.)
- R-Team, 2009. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2009. URL <http://www.R-project.org>. (Citado en la página 91.)
- Rencher A.C., 1995. *Methods of multivariate analysis*. Wiley, New York. (Citado en la página 211.)
- Rivas-Martínez S., Asensi A., Díez-Garretas B., Molero-Mesa J., y Valle F., 1997. Biogeographical synthesis of Andalusia (southern Spain). *Journal of Biogeography*, 24:915–928. (Citado en la página 167.)
- Roeckner E., Arpe K., Bengtsson L., Christoph M., Claussen M., Dumenil L., Esch M., Giorgetta M., Schlese U., y Schulzweida U., 1996. Report 218: The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. Technical report, Max Planck Institute, Hamburg, Germany. (Citado en la página 90.)
- Rogers R. y Johnson P.S., 1998. Approaches to modeling natural regeneration in oak-dominated forests. *For. Ecol. Manage*, 106:45–54. (Citado en la página 29.)
- Root T.L., Price J.T., Hall K.R., Schneider S.H., Rosenzweig C., y Pounds J.A., 2003. Fingerprints of global warming

- on wild animals and plants. *Nature*, 421:57–60. (Citado en la página 86.)
- Rorison I.H., 1973. *The effect of extreme soil acidity on the nutrient uptake and physiology of plants, in Acid Sulphate Soil*. Dost, H. (Citado en la página 18.)
- Sala O.E., Chapin III F.S., Armesto J.J., Berlow E., Bloomfield J., Dirzo R., Huber-Sanwald E., Huenneke L.F., Jackson R.B., Kinzig A., Leemans R., Lodge D.M., Mooney H.A., Oesterheld M., Poff N.L., Sykes M.T., Walker B.H., Walker M., y Wall D.H., 2000. Global biodiversity scenarios for the year 2100. *Science*, 287(5459):1770–1774. (Citado en la página 40.)
- Salisbury E.J., 1926. The geographical distribution of plants in relation to climatic factors. *The Geographical Journal*, 57: 312–355. (Citado en la página 17.)
- Segurado P. y Araújo M.B., 2004. An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31(10):1555–1568. (Citado en las páginas 44, 46, y 50.)
- Soberón J., 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, 10(12): 1115–1123. (Citado en las páginas 9, 11, 16, y 19.)
- Soberón J. y Peterson A.T., 2004. Biodiversity informatics: Managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1444):689–698. (Citado en la página 26.)
- Soberón J. y Peterson A.T., 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2:1–10. (Citado en las páginas 17 y 20.)
- Soule M.E., 1991. Conservation: Tactics for a constant crisis. *Science*, 253(5021):744–750. (Citado en la página 40.)
- Svenning J.C. y Skov F., 2004. Limited lling of the potential range in European tree species. *Ecol. Lett.*, 7:565–573. (Citado en la página 16.)

- Sánchez-Cordero V. y Martínez-Meyer E., 2000. Museum specimen data predict crop damage by tropical rodents. *Proceedings of the National Academy of Sciences of the United States of America*, 97(13):7074–7077. (Citado en la página 26.)
- Takakura T. y Fang W., 2002. *Climate Under Cover*. Springer. (Citado en las páginas 41 y 55.)
- Team GRASS Development, 2009. *Geographic Resources Analysis Support System (GRASS GIS) Software*. Open Source Geospatial Foundation, USA, 2009. URL <http://grass.osgeo.org>. (Citado en la página 91.)
- Theurillat J.P. y Guisan A., 2001. Potential impact of climate change on vegetation in the European alps: A review. *Climatic Change*, 50:77–109. (Citado en las páginas 86 y 98.)
- Thomas C.D. y Franco A.M.A. and Hill J.K., 2006. Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution*, 21(8):415–416. (Citado en la página 100.)
- Thomas C.D., Cameron A., Green R.E., Bakkenes M., Beaumont L.J., Collingham Y.C., Erasmus B.F.N., Ferreira M., Grainger A., Hannah L., Hughes L., Huntley B., Van Jaarsveld A.S., Midgley G.F., Miles L., Ortega-Huerta M.A., Peterson A.T., Phillips O.L., y Williams S.E., 2004. Extinction risk from climate change. *Nature*, 427(6970). (Citado en la página 13.)
- Thuiller W., 2003. BIOMOD - Optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9(10):1353–1362. (Citado en las páginas 44 y 105.)
- Thuiller W., Lavorel S., y Araújo M.B., 2005a. Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, 14(4):347–357. (Citado en la página 40.)
- Thuiller W., Lavorel S., Araújo M.B., Sykes M.T., y Prentice I.C., 2005b. Climate change threats to plant diversity in



- Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12):8245–8250. (Citado en las páginas 40, 86, y 96.)
- Thuiller W., Albert C., Araújo M.B., Berry P.M., Cabeza M., Guisan A., Hickler T., Midgley G.F., Paterson J., Schurr F.M., Sykes M.T., y Zimmermann N.E., 2008. Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, 9(3-4):137–152. (Citado en las páginas 164, 165, 166, y 190.)
- Thuiller W., Lafourcade B., Engler R., y M.B. Araújo, 2009. BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, 32:369–373. (Citado en la página 105.)
- Valle F., 2004. *Datos botánicos aplicados a la gestión del medio natural andaluz I: Bioclimatología y Biogeografía*. Consejería de Medio Ambiente, Junta de Andalucía, Sevilla. (Citado en la página 143.)
- Varcoe V.J., 1978. A note on the computer simulation of crop growth in agricultural land evaluation. *Soil Use Manage*, 6:37–54. (Citado en la página 29.)
- Verheyen K. y Hermy M., 2001. The relative importance of dispersal limitation of vascular plants in secondary forest succession in Muizen Forest Belgium. *Journal of Ecology*, 89(5):829–840. (Citado en la página 164.)
- Vittoz P. y Engler R., 2007. Seed dispersal distances: a typology based on dispersal modes and plant traits. *Botanica Helvetica*, 117:109–124. (Citado en las páginas 160, 164, y 171.)
- Walter H., 1953. Das Gesetz der relativen Standortskonstanz: Das Wesen der Pflanzengesellschaften. *Ber. Dtsch. bot. Ges.*, 66:228–236. (Citado en la página 18.)
- Walther G.R., 2004. Plants in a warmer world. *Perspectives in Plant Ecology Evolution and Systematics*, 6(3):169–185. (Citado en la página 164.)

- Ward D.F., 2007. Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions*, 9(6):723–735. (Citado en las páginas 40 y 51.)
- Wardle P. y Coleman M.C., 1992. Evidence for rising upper limits of four native New Zealand forest trees. *New Zealand Journal of Botany*, 30:303–314. (Citado en la página 86.)
- Whittaker R.H., 1965. Dominance and diversity in land plant communities. *Science*, 147(3655):250–260. (Citado en la página 17.)
- Wiens J.J., 2004. What is speciation and how should we study it? *American Naturalist*, 163(6):914–923. (Citado en la página 15.)
- Wiens J.J. y Graham C.H., 2005. Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, 36:519–539. (Citado en las páginas 13 y 15.)
- Williams M.B. y Araújo P.H., 2000. Selecting areas for species persistence using occurrence data. *Biological Conservation*, 96(3):331–345. (Citado en la página 40.)
- Williams P.H., Hannah L., Andelman S., Midgley G., Araújo M.B., Hughes G., Manne L., Martinez-Meyer E., y Pearson C., 2005. Planning for climate change: identifying minimum-dispersal corridors for the Cape Proteaceae. *Conservation Biology*, 19(4):1063–1074. (Citado en la página 87.)
- Woodward F.I., 1987. *Climate and Plant Distribution*. University Press, Cambridge. (Citado en la página 17.)
- Zaniewski A.E., Lehmann A., y Overton J.M., 2002. Predicting species spatial distributions using presence-only data: A case study of native New Zealand ferns. *Ecological Modelling*, 157(2-3):261–280. (Citado en las páginas 51 y 52.)



## DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD

---

Yo, **Blas Manuel Benito de Pando**, Investigador Predoctoral del Departamento de Botánica de la Universidad de Granada DECLARO:

Que la investigación realizada y el contenido de la presente Memoria de Tesis Doctoral es de creación propia, y no contiene material que haya sido previamente publicado o escrito por otra persona excepto donde se reconoce como tal. Declaro igualmente que hago justo reconocimiento en la tesis a las personas que contribuyeron con su trabajo y, finalmente, declaro que esta tesis es producto de mi propio trabajo con el apoyo permitido de terceros en cuanto a la concepción del proyecto, al estilo de la presentación o a la expresión escrita.

*Granada, 14 de septiembre de 2009*

---

Blas M. Benito de Pando