

World GISs resources



**Global databases** 



```
Complex modelling techniques
```



Bases teóricas de los modelos de distribución o de nicho

# La distribución geográfica y las posibilidades e inconvenientes de extracción de información del nicho





### Species inhabit in favourable localities.....but not in all.



# Survey effort

Registros de espermatofitas en BIOTA (29-IX-2003)



# Observed richness





Do we need absence data to extrapolate distributions?



What is the meaning of the obtained representation?

		e	n 1		s á	n	¢.		r n	1	e (		ne		e ç	٤.	1.0	2	) q	e n	22	ú n	en re	y	e re	. p	ц. :	ц.	T	¢Ľ		fi.	y e	Į i	.0	•	E	1	y re	n	er,
e	r	e	ne	f	n j	a e	en	re	e	g en	i j	ó j	7 6	)	e		g i	n 1	re	çe	er	m	ne	r	en	e er	ŧē	h	4	• ;	y re i	n g	<b>c</b> n	y e	n q	é r	n	1	e	r	en
er.	ł	ie e	en	r	e	e ó		n		e		r	er.		ó	3																									
n			re	r	e e	r e	*		-	1													-	2									erer	1 1 1		e	¢	ón q	e f	- 1	n e
у		n ei	Ó	n e	•	er		T	X	ĥ	2	4	2	r	Δ	<b>1</b> ]	h	Δ	n	n	ic	'C	ir		5	CT	14	<b>ב</b>	<b>.</b> i	ρ	<b>C</b> <sup>(</sup>	)		e r	n r	e e	2 1	r,	en re	e 1	1.
á	e	n e	e		ne	ene	n			11			a					L				0		18	5	<b>P</b>	<b>)</b>			U			. e	r	, r	e e		e re	r	î	e
qe	e		re	ó	n e	nfr	e r	I		2	ón	e	,		reex		ón,		n	n d	òn	fëi	ſ	rn	fr	ón	e		e	e	, n	en	re		3	e e		en	e y,	r	n
	16	9	er		ó		n	er	nee	r,	e	rx	ne	9	n	er	ón	n e	e án		n	n	er n	nf	r	ón	re	e	e e		e r	e	e	n er	ó	n	1	en	re	е	e
en	y	e e	n e	r		en		q e	r err	r	у	enef	: ,	r	ener	qe	en	er	ner		e	e	e n	ó	у	r		2	e	n fé	e rn		e y	r	z	. e e	; e		nq	e e	e
en		r	i		e	en	e	en	en e	1	et	n	z	ón	r z n	e	e	re	r 1	n r	е,	e	én 1	ie e	r	fr n	r	r	e á		n re	e	n er	r	ġ	re	ó	q	e r	r	n
en		r	e e	ć																																					

Los sistemas naturales están compuestos por unos actores (las poblaciones o las especies, por ejemplo) que interactúan entre si y se relacionan con las características físicas y quimicas del medio. Estudiar y reconocer esas relaciones funcionales entre los elementos bióticos y abióticos de los sistemas naturales puede permitirnos manejar sosteniblemente dichos sistemas y reconocer cuando y en qué circunstancias se producen alteraciones en los *procesos* ecológicos con capacidad de alterar la diversidad biológica.

Tanto las direc

refieren prograt como el agrario

What are the not well-surveyed localities?

alies as at Use Contamilate do to Direct

nicas

sobre los flujos de energia y los ciclos de elementos y compuestos en los sistemas naturales. Se trata, por ejemplo, de regular los procesos que como la creación de infraestructuras, la utilización de biocidas, la sobreexplotación, la contaminación atmosférica o la transformación de los usos del suelo, inciden sobre la calidad del medio ambiente y, por tanto, sobre la diversidad biológica. Con ser necesarias, estas aproximaciones a la conservación no están basadas ni consideran la información sobre las especies. Las estrategias de conservación basadas en los procesos pueden y deben ser aplicadas en cualquier territorio y sus beneficios, por tener que vencer la inercia de los modelos económicos y productivos al uso, se manifestaran a medio y largo plazo. De este modo, aunque se deben aplicar políticas medioambientales tendentes a promover una utilización razonable de los recursos naturales, es también necesario afrontar la problemática concreta de conservar los actores biológicos que participan en los procesos ecológicos.

Distribución potencial y realizada

## **Potential Distribution**



# Equilibrium with the environmental conditions

## **Realized Distribution**



#### Role of dispersal limitations, biotic interactions and other contingent factors



### Potential Distribution



### **Background absences**









555



# Potential Distribution



### **Background absences**



Absences selected outside environmentally favourable regions.



near to RD

## How can we obtain relatively reliable absence data?

R-script

## YES

Simplified representation of an unknown reality

Y = f(X)What kind of predictors?  $\vec{\mathbf{Y}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix}, \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$ Presence Absence





## YES

Non-equilibrium with the environmental factors



When absences are selected far from the environmental conditions prevailing in the presences the predicted distribution area is higher.



## We need to use the "pure" influence of climate factors

SAM

## Models or colour maps?

MaxEnt Diva-Gis

## NOT



"Most ecologist would agree that niche is a central concept of ecology, even though we do not know exactly what it means"

*Real & Levin (1991)* 





## NOT

•Geographic distribution is an empirical evidence (uncertainty principle )

•Fundamental or realized niche cannot be quantified from geographic observations.



Occurrence or abundance are the result of the role played by multiple limiting factors that hinder the unveil of the niche.

Distributions do not reflect niches



Var 2







Methods in Ecology and Evolution



Methods in Ecology and Evolution

doi: 10.1111/j.2041-210X.2011.00141.x

#### Comparative interpretation of count, presence–absence and point methods for species distribution models

Geert Aarts<sup>1,2\*</sup>, John Fieberg<sup>3</sup> and Jason Matthiopoulos<sup>4,5</sup>

Ecology, 93(3), 2012, pp. 679-688 © 2012 by the Ecological Society of America

## Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model

#### ROBERT J. HUMANS

Department of Environmental Science and Policy, 1023 Wickson Hall, University of California, Davis, California 95616 USA

Abstract. Species distribution models are usually evaluated with cross-validation. In this procedure evaluation statistics are computed from model predictions for sites of presence and absence that were not used to train (fit) the model. Using data for 226 species, from six regions, and two species distribution modeling algorithms (Bioclim and MaxEnt), I show that this procedure is highly sensitive to "spatial sorting bias": the difference between the geographic distance from testing-presence to training-presence sites and the geographic distance from testing-absence (or testing-background) to training-presence sites. I propose the use of pairwise distance sampling to remove this bias, and the use of a null model that only considers the geographic distance to training sites to calibrate cross-validation results for remaining bias. Model evaluation results (AUC) were strongly inflated: the null model performed better than MaxEnt for 45% and better than Bioclim for 67% of the species. Spatial sorting bias and area under the receiver-operator curve (AUC) values increased when using partitioned presence data and random-absence data instead of independently obtained presence-absence testing data from systematic surveys. Pairwise distance sampling removed spatial sorting bias, yielding null models with an AUC close to 0.5, such that AUC was the same as null model calibrated AUC (cAUC). This adjustment strongly decreased AUC values and changed the ranking among species. Cross-validation results for different species are only comparable after removal of spatial sorting bias and/or calibration with an appropriate null model.

Key words: AUC; Bioclim; cross-validation; MaxEnt; model evaluation; niche model; pairwise distance sampling; spatial autocorrelation; spatial sorting bias; species distribution model.



Ecography 36: 864–867, 2013 doi: 10.1111/j.1600-0587.2013.00321.x © 2013 The Authors. Ecography © 2013 Nordic Society Oikos Subject Editor: Miguel Araújo. Accepted 22 February 2013

#### Inference from presence-only data; the ongoing controversy

#### **Trevor Hastie and Will Fithian**

T. Hastie (hastie@stanford.edu) and W. Fithian, Statistics Dept, Stanford Univ., CA 94305, USA.

Presence-only data abounds in ecology, often accompanied by a background sample. Although many interesting aspects of the species' distribution can be learned from such data, one cannot learn the overall species occurrence probability, or prevalence, without making unjustified simplifying assumptions. In this forum article we question the approach of Royle et al. (2012) that claims to be able to do this.



# Accuracy measurements

## Discrimination

the ability to separate the instances of presence from the instances of absence •Calibration

how well the estimated probability of presence represents the observed

proportion of presences



Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2013) 22, 508-516



#### Discrimination capacity in species distribution models depends on the representativeness of the environmental domain

Alberto Jiménez-Valverde<sup>1\*</sup>, Pelayo Acevedo<sup>1,2</sup>, A. Márcia Barbosa<sup>3,4</sup>, Jorge M. Lobo<sup>5</sup> and Raimundo Real<sup>1</sup>



Observed proportion of positive cases





ORIGINAL PAPER

### Threshold-dependence as a desirable attribute for discrimination assessment: implications for the evaluation of species distribution models

Alberto Jiménez-Valverde

