Cortesia Emilio Padoa-Schioppa e Francesco Ficetola

# L'uomo e la geografia

**Le attività umane sono cruciali nel determinare la biodiversità delle isole, come ha mostrato uno studio sui rettili del Mediterraneo**

di Emilio Padoa-Schioppa e Francesco Ficetola

Viviamo in un periodo in cui l'umanità è in grado di modificare gli equilibri complessivi della biosfera, un'epoca indicata anche con il termine Antropocene (*si veda il box a p. 99*). La pressione antropica sta causando, tra le altre cose, l'estinzione precoce di molte forme di vita. La perdita di biodiversità ha raggiunto livelli tanto allarmanti che i biologi della conservazione parlano di sesta estinzione di massa, paragonando la situazione attuale ad altri momenti del passato in cui vi furono estinzioni rapide, estese e che sconvolsero gli equilibri della Terra (per esempio la scomparsa dei dinosauri). La distruzione e la trasformazione di habitat e paesaggi naturali e la diffusione delle specie invasive sono tra le principali cause di questa estinzione di massa.

Non è però semplice documentare direttamente un'estinzione. Non sempre durante sopralluoghi e spedizioni si riesce a individuare una specie rilevata in passato, ma ciò non significa che la specie

AMANTE DELLE ISOLE. Un esemplare di lucertola campestre (*Podarcis sicula*), una delle specie di lucertola più diffuse sulle isole italiane.

# della vita

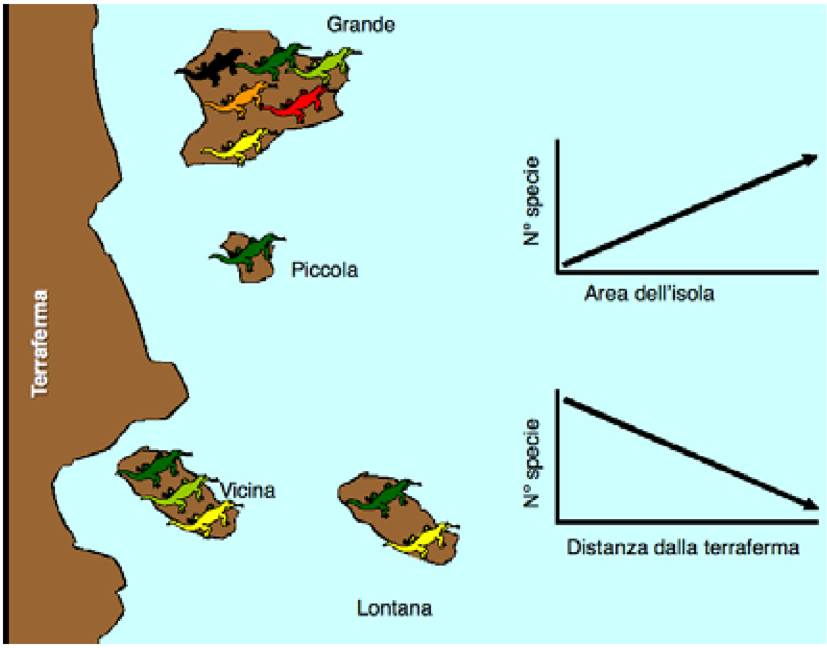
### IN SINTESI

- Fino a poco tempo fa si ipotizzava che il numero di specie insulari dipendesse soprattutto dalle dimensioni dell'isola e dalla sua distanza dalla terraferma.
- Ora uno studio sulla distribuzione dei rettili delle isole mediterranee ha dimostrato che l'antropizzazione del territorio influenza in modo fondamentale la biodiversità.
- Questo risultato mostra ancora una volta che la nostra epoca è caratterizzata dalla capacità dell'uomo di stravolgere gli equilibri della biosfera.

in questione sia veramente estinta: è capitato che qualche specie sia stata poi riscoperta. Tra l'altro ci sono estinzioni continue non rilevate, visto che numerose specie non sono state ancora descritte. Per queste ragioni, la reale portata dell'azione antropica viene stimata sulla base dei dati empirici disponibili. Le isole, con la loro estensione limitata e le specie uniche che le abitano, sono meravigliosi laboratori naturali per costruire ipotesi sui fattori che determinano la diffusione degli organismi, in ambienti in cui il numero di variabili può essere tenuto maggiormente sotto controllo.

All'interno dell'ecologia si è sviluppata una disciplina, la biogeografia insulare, che studia la distribuzione di piante e animali sulle isole. Negli anni sessanta gli ecologi statunitensi Robert MacArthur ed Edward O. Wilson avevano proposto un modello per cui il numero di specie in un'isola dipende da due fattori: le dimensioni dell'isola e la distanza dalla terraferma. Quando una specie colonizza un'isola, le dimensioni determinano la probabilità di stabilirvisi con successo oppure estinguersi: in isole di dimensioni maggiori, gli organismi possono stabilire popolazioni più numerose e trovare una maggiore varietà di microhabitat.

Quindi isole più grandi ospitano un nume-



MODELLO E REALTÀ. In alto il modello dell'equilibrio insulare proposto da MacArthur e Wilson, secondo cui un'isola grande o vicina al continente ospita più specie rispetto a una piccola o distante. Qui sopra una lucertola maltese, endemica delle isole maltesi e di due isole Pelagie.

Danilo Sossi (*grafico*); cortesia Emilio Padoa-Schioppa e Francesco Ficetola (*lucertola maltese*)

ro maggiore di specie e comunità più complesse e articolate. Inoltre, più un'isola è vicina alla terraferma maggiore è la probabilità che alcuni organismi attraversino il tratto di mare, aumentando il numero di specie presenti. Nel complesso, comunque, le dimensioni dell'isola sono considerate il fattore che determina maggiormente il numero di specie presenti. Dopo la pubblicazione della teoria di MacArthur e Wilson, i loro allievi hanno cominciato a studiare e sperimentare la validità dell'ipotesi. In particolare Daniel Simberloff ha fatto un esperimento complesso su alcuni isolotti al largo della Florida, osservando il numero di specie animali e vegetali e provvedendo poi a rimuovere tutte le specie animali per controllare le dinamiche di ricolonizzazione: i risultati dell' esperimento hanno confermato il modello di MacArthur e Wilson.

Numerosi altri studi su diversi gruppi tassonomici (uccelli, mammiferi, rettili, anfibi, piante e invertebrati) hanno prodotto risultati simili, portando alla formulazione del principio secondo cui in un'isola il numero di specie è direttamente proporzionale alla sua area e inversamente proporzionale alla distanza dal continente (*si veda lo schema in alto in questa pagina*), un modello che si è imposto come una delle principali basi teoriche in ecologia.

A questo semplice schema si è poi aggiunto che occorre considerare anche la varietà di ambienti in un'isola, visto che, a parità di superficie, un'isola con più ambienti (per esempio con un'ampia variazione altitudinale) dovrebbe ospitare un numero più elevato di specie.

L'azione dell'uomo può influenzare fortemente gli ecosistemi insulari: a causa delle loro piccole dimensioni, questi ambienti sono molto fragili, e l'uomo può determinare rapidamente l'estinzione delle specie che le abitano. Alcune delle estinzioni più note sono avvenute proprio su isole, come la scomparsa del dodo a Mauritius. Ma la presenza umana è stata raramente inserita nei modelli predittivi della biogeografia insulare. Qual è l'effetto dell'azione antropica sul numero di specie in un'isola? Il nostro studio sugli arcipelaghi del Mediterraneo ha cercato di analizzare se le attività umane possano modificare le relazioni biogeografiche classiche, come il rapporto tra area e numero di specie.

## Rettili e isole nel Mediterraneo

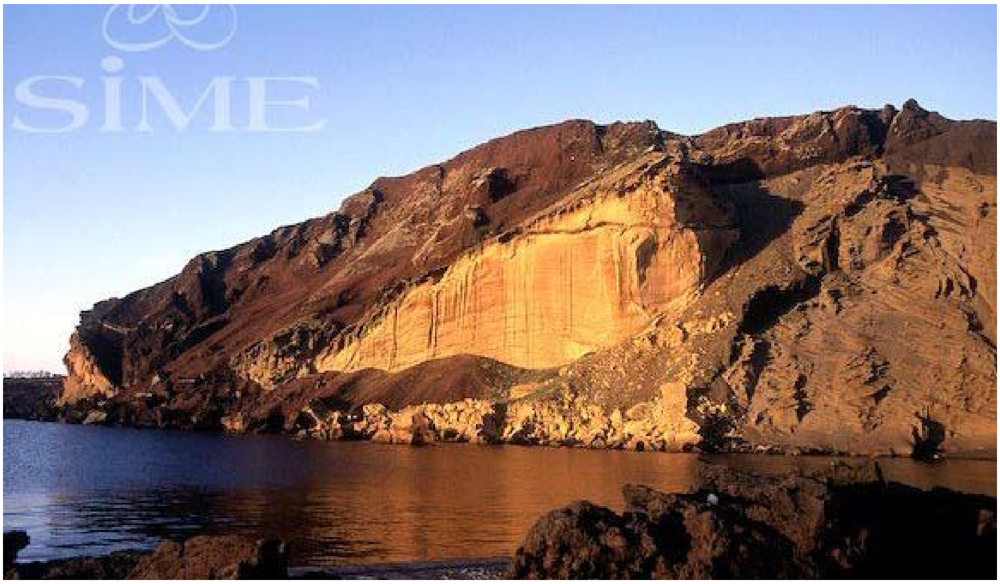
Nel bacino del Mediterraneo vi sono numerose isole, alcune delle quali molte ben studiate dal punto di vista biologico: è quindi una regione ideale per analisi biogeografiche. Quest'area ha però ospitato più di 3000 anni di storia umana, e le sue isole ne sono state protagoniste. Abbiamo visto il Mediterraneo come un laboratorio unico per capire come le attività umane interferiscano con i processi naturali che determinano la distribuzione delle specie. Il Mediterraneo è inoltre uno dei centri di biodiversità dei rettili: l'Unione internazionale di conservazione della natura mostra che i paesi del bacino ospitano la maggior ricchezza di rettili in Europa.

Numerose di queste specie sono minacciate di estinzione, anche perché molte sono endemismi insulari, cioè sono presenti solo su alcune isole e assenti dalla terraferma. La lucertola maltese (*Podarcis filfolensis*), per esempio, vive solo nell'arcipelago maltese e in due delle isole Pelagie, Lampione e Linosa; e la lucertola di Bedriaga (*Archeolacerta bedriagae*) vive solo in Sardegna, Corsica e in alcune isole satellite. Ma le specie endemi-

## La Terra ai tempi dell'Antropocene

Il termine Antropocene è stato coniato e diffuso da Paul Crutzen, premio Nobel per la chimica nel 1995 grazie ai suoi studi sull'ozono stratosferico. Con questo termine Crutzen indica che a partire dalla Rivoluzione industriale la Terra è entrata in una nuova fase geologica, dominata dall'uomo. Le azioni antropiche, in effetti, sono in grado di modificare gli equilibri del pianeta. L'evento più evidente è l'alterazione della composizione dell'atmosfera, dovuta alle emissioni di gas serra che determinano i cambiamenti climatici, riscaldamento globale in testa. Insieme al riscaldamento globale si osservano alterazioni nei cicli biogeochimici (come quello dell'azoto e quello del fosforo), nella chimica dell'atmosfera (come il buco nello strato di ozono), nella chimica delle acque e dei suoli. L'uomo agisce anche come forza geomorfologica, in grado di smuovere più terra di quanto facciano i fiumi con la loro attività di erosione. Inoltre altera radicalmente la superficie terrestre, distruggendo e cambiando habitat, ecosistemi e paesaggi a ritmi sempre più elevati e in modo diffuso in quasi tutta la biosfera.

A livello di biodiversità, infine, sono sempre più frequenti le estinzioni locali o globali collegabili alle attività antropiche. Le principali cause di scomparsa della biodiversità sono distruzione degli habitat, diffusione di specie alloctone invasive, inquinamento, crescita della popolazione umana ed eccesso di caccia e raccolta. E ovviamente tutti questi fattori possono agire in maniera sinergica: per esempio i cambiamenti climatici accelerano il declino di alcune specie già minacciate dalle attività umane.



POCO ANTROPIZZATA. L'isola di Linosa, nel canale di Sicilia, è poco antropizzata ed è stata studiata dagli autori per essere confrontata con altre isole mediterranee a elevato tasso di antropizzazione e urbanizzazione.

Matteo Carassale/SIME

che non sono limitate alle lucertole. In Sicilia vive una specie di testuggine palustre (*Emys trinacris*) distinta da quelle presenti nel resto d'Europa.

Il tasso di endemismo è elevato in Macaronesia (cioè le isole Canarie, Azzorre e Madeira, che pur trovandosi nell'Oceano Atlantico vengono spesso associate alla regione mediterranea), in cui è addirittura presente un genere di lucertole endemico (*Gallotia*) che comprende sette specie ed è un interessante esempio di radiazione adattativa. D'altra parte, millenni di storia umana hanno agito in maniera diversa sulle isole: alcune sono molto antropizzate e urbanizzate, alcune ospitano insediamenti antropici minimi e altre sono disabitate, rendendole perfette per i nostri obiettivi.

Sfruttando la letteratura scientifica disponibile, quali gli atlanti di distribuzione della fauna, abbiamo costruito una banca dati che racchiude informazioni sui rettili che vivono in 212 isole del Mediterraneo occidentale e della Macaronesia. Per ciascuna isola abbiamo elencato le specie di rettili terrestri, valutando se fossero autoctone o introdotte. Abbiamo poi misurato alcune caratteristiche geografiche (area, distanza dalla terraferma e altitudine massima) e alcune misure di pressione antropica (numero di abitanti e presenza di aeroporti). La presenza di aeroporti ci è sembrata un buon indicatore, facile da misurare e omogeneo in tutta l'area, dell'intensità dei flussi turistici ed economici delle isole.

**GLI AUTORI**

**EMILIO PADOA-SCHIOPPA** e **FRANCESCO FICETOLA** lavorano al Dipartimento di scienze dell'ambiente e del territorio dell'Università degli Studi di Milano-Bicocca, e si occupano di conservazione del paesaggio e della biodiversità. Padoa-Schioppa è ricercatore confermato in ecologia, Ficetola è assegnista di ricerca.

Silvio Fiore/La Presse (*Malta*)

## La teoria dell'informazione nei modelli ecologici

L'ecologia ha spesso sofferto di un complesso di inferiorità rispetto ad altre discipline scientifiche: in fisica, in chimica o anche in alcune aree della biologia è possibile fare esperimenti ripetibili per testare le ipotesi scientifiche. Questa procedura è però impraticabile sia dal punto di vista logistico sia da quello etico per gli studi di ecologia o biogeografia che cercano di capire le cause della distribuzione delle specie: nel nostro caso, non avremmo potuto immaginare di creare 200 isole artificiali, introdurvi un certo numero di specie e poi l'uomo. Per studi del genere la distribuzione delle specie è tradizionalmente analizzata con un processo induttivo (prima si fanno le osservazioni e poi si propone un'ipotesi), ma questo procedimento può rendere complessa la falsificazione delle ipotesi scientifiche.

Negli ultimi decenni, però, sono stati sviluppati raffinati metodi statistici, basati sulla teoria dell'informazione, che permettono di valutare quale ipotesi scientifica sia meglio suffragata dai dati disponibili. In pratica si segue un processo ipotetico deduttivo (proposta, confronto e falsificazione di ipotesi alternative), che permette di valutare la probabilità che una data ipotesi sia vera:

1) vengono formulate ipotesi a priori sui processi che possono determinare la distribuzione delle specie;
2) vengono costruiti modelli statistici che corrispondono alle diverse ipotesi;
3) i dati raccolti sono inseriti nei modelli, che sono confrontati sulla base di criteri statistici.

A questo punto è possibile stabilire qual è il modello più probabile, sulla base dei dati raccolti, e quindi quali processi biologici hanno verosimilmente determinato la distribuzione delle specie.

Sfruttando la teoria dell'informazione (*si veda il box qui a fianco*) abbiamo confrontato tra loro tre possibili modelli: il modello geografico, secondo cui le caratteristiche geografiche naturali determinano il numero di specie presenti sulle isole; il modello antropico, in cui la pressione umana spiega il numero di specie; e un modello congiunto, in cui il ruolo chiave è svolto da un effetto combinato delle caratteristiche geografiche e antropiche. Queste analisi sono state fatte sia per le specie autoctone sia per quelle introdotte dall'uomo.

### Rettili nativi e rettili alloctoni

Sulle isole considerate abbiamo individuato da 1 a 15 specie native di rettili. Come previsto, il numero di specie è maggiore nelle isole di maggior superficie (Sicilia, Sardegna e Corsica) e con rilievi, e diminuisce nelle isole più distanti dal continente. Le caratteristiche geografiche però non sono sufficienti a spiegare la ricchezza in specie. Anche l'antropizzazione è importante: a parità di superficie, le isole meno antropizzate hanno più specie. Malta, per esempio, è relativamente grande (250 chilometri quadrati) ma ha una pressione antropica molto elevata, e ospita sei specie autoctone; al largo della Sardegna ci sono diverse isolette (Tavolara, Molara o Asinara) di pochi chilometri quadrati ma poco antropizzate: ospitano fino a 9-11 specie native. Anche a Maiorca e Minorca vivono meno

NON SOLO LUCERTOLE. La natrice viperina (*in alto*) è stata introdotta sulle Baleari in epoca romana, dove ha causato gravi danni alla biodiversità. Il ramarro (*in basso*) è un grosso sauro che vive in Europa continentale e in alcune isole. A fronte, Malta, isola con elevato tasso di antropizzazione, che influisce in modo negativo sulla biodiversità.

Cortesia Emilio Padoa-Schioppa e Francesco Ficetola (*natrice viperina e ramarro*)

specie rispetto a quanto ci aspetteremmo dall'analisi di superficie e distanza dal continente.

Dalle analisi emerge quindi che il rapporto tra superficie e numero di specie non è lineare, come previsto dalla teoria. I dati sono descritti meglio da una linea spezzata: il numero di specie cresce rapidamente con l'area nelle isole con superficie fino 1,5 chilometri quadrati, ma la crescita si appiattisce nelle isole più grandi. La superficie di 1,5 chilometri quadrati è proprio il punto in cui l'impatto antropico aumenta improvvisamente, e la maggior parte delle isole sopra queste dimensioni è abitata.

Nel complesso, abbiamo costruito un modello secondo cui l'azione antropica modifica la relazione tra superficie e ricchezza in rettili. L'area ha un impatto diretto e positivo sul numero di specie: a parità di altri fattori, le isole più grandi sono le più ricche di specie. Ma l'uomo non si stabilisce a caso sulle isole, e l'area ha anche un effetto importante sul grado di antropizzazione: le isole più grandi sono anche quelle con più popolazione e scambi commerciali. A sua volta, l'antropizzazione influisce negativamente sulla ricchezza di specie. In pratica, le attività umane hanno modificato la relazione tra superficie e specie presenti.

Sulle isole inoltre abbiamo contato da 0 a 7 specie alloctone, di cui molte introdotte in tempi storici anche recenti dai continenti europeo e africano o da altre isole del Mediterraneo. Ne sono esempi il camaleonte (*Chamaleo chamaleon*), introdotto a Malta in tempi recenti, e la natrice viperina (*Natrix maura*), un serpente introdotto a Maiorca e Minorca in epoca romana. Alcune specie hanno causato danni incalcolabili alla biodiversità delle isole. Proprio la natrice viperina ha portato il rospo ostetrico di Maiorca (*Alytes muletensis*) sull'orlo dell'estinzione. Le nostre analisi mostrano che i fattori antropici sono essenziali per spiegare la distribuzione delle specie alloctone: le isole più popolate e con più scambi hanno subito un maggior numero di introduzioni, e quindi vi si sono stabilite più specie. Ma i soli fattori antropici non bastano e occorre considerare anche i fattori geografici. Le dimensioni dell'isola sono cruciali per il successo dell'insediamento: un'isola grande può offrire un maggior numero di nicchie ecologiche dove stabilirsi, e quindi ospitare più specie alloctone.

## L'imprescindibile ruolo dell'uomo

I nostri risultati mostrano che l'azione umana ha un'enorme influenza sulla distribuzione delle specie native e invasive. Isole con una grande pressione antropica ospitano meno specie native di quanto ci aspetteremmo considerando solo fattori geografici; per contro, aumentano le specie invasive.

Questo lavoro propone due spunti di riflessione. Da un lato, la distribuzione delle specie mostra che è valida la regola del «piccolo è bello»: oggi alcune piccole isole hanno un valore molto importante per la conservazione della biodiversità, visto che su di esse sopravvivono specie estinte sulle isole principali a causa dell'antropizzazione. La lucertola delle Eolie (*Podarcis raffonei*), per esempio, è scomparsa in tutte le isole principali e sopravvive solo su qualche isolotto o scogli disabitati.

Inoltre ci siamo resi conto che occorre sviluppare modelli che considerino l'azione dell'uomo. L'ecologia è una scienza relativamente giovane, e solo negli ultimi anni sta dibattendo sull'esistenza di leggi generali al suo interno. In una recente sintesi, Walter Dodds ne ha individuate diverse, e la terza recita: «L'uomo domina su ogni cosa». È il riconoscimento dell'Antropocene, e i nostri risultati vanno in questa direzione. È il momento di considerare la pressione antropica come un fattore che rivaleggia con i processi naturali e geografici nel determinare la biodiversità di un territorio. ■

## ➥ Letture

**The Status and Distribution of Reptiles and Amphibians of the Mediterranean Basin.** Cox N., Chanson J.S. e Stuart S.N., IUCN, Gland, Svizzera e Cambridge, 2006.

**Benvenuti nell'Antropocene.** Crutzen P., Mondadori, Milano, 2005.

**Laws, Theories and Patterns in Ecology.** Dods W.K., University of California Press, 2009.

**Human Activities Alter Biogeographical Patterns of Reptiles on Mediterranean Islands.** Ficetola G.F. e Padoa-Schioppa E., in «Global Ecology and Biogeography», Vol. 18, n. 2, pp. 214-222, novembre 2008.

**Island biogeography.** Whittaker R.J. e Fernández-Palacios J.M., Oxford University Press, 2007.

**Il futuro della vita.** Wilson E.O., Codice Edizioni, Torino, 2004.

*Diversity and Distributions, (Diversity Distrib.)* (2009) **15**, 108–116

BIODIVERSITY RESEARCH

# From introduction to the establishment of alien species: bioclimatic differences between presence and reproduction localities in the slider turtle

Gentile Francesco Ficetola[1]*, Wilfried Thuiller[2] and Emilio Padoa-Schioppa[1]

[1]*Dipartimento di Scienze dell'Ambiente e del Territorio, Università degli Studi di Milano-Bicocca, Piazza Della Scienza 1, 20126 Milano, Italy,* [2]*Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université J. Fourier, BP 53, 38041 Grenoble Cedex 9, France*

*Correspondence: Gentile Francesco Ficetola, Dipartimento di Scienze dell'Ambiente e del Territorio, Università degli Studi di Milano-Bicocca, Piazza Della Scienza 1, 20126 Milano, Italy. E-mail: francesco.ficetola@unimi.it

## ABSTRACT

**Aim** Understanding the factors determining the transition from introduction of aliens to the establishment of invasive populations is a critical issue of the study of biological invasions, and has key implications for management. Differences in fitness among areas of introduction can define the zones where aliens become invasive. The American slider turtle *Trachemys scripta* has been introduced worldwide, and has negative effects on freshwater communities, but only a subset of introduced populations breed successfully. We used species distribution models to assess the factors influencing the slider distribution in Italy, by analysing bioclimatic features that can cause the transition from presence of feral adults to breeding populations. We also evaluated whether climate change might increase the future suitability for reproduction.

**Location** Central and Northern Italy.

**Methods** The distribution of slider turtle was obtained from the literature, unpublished reports and field surveys. We used Maxent to build bioclimatic models.

**Results** Reproductive populations are associated to a clear bioclimatic envelope with warmer climate, more solar radiation and higher precipitations than populations where reproduction is not observed. Several Mediterranean areas currently have climatic features suitable for sliders. Scenarios of climate change predict the expansion of these areas. In the near future (2020), the proportion of populations in areas suitable for reproduction will dramatically increase.

**Main conclusion** Our study shows that bioclimatic differences can determine the areas where aliens become invaders. Management should be focused to these source areas. However, climate change can increase fitness in the future, and therefore the interactions between climate change and fitness can boost the invasiveness of this alien species.

**Keywords**

**Bioclimatic envelope, biological invasions, climate change, fitness, MAXENT, problematic alien species, reproduction, species distribution models.**

## INTRODUCTION

Alien invasive species (AIS) are a major cause of biodiversity loss (Strayer *et al.*, 2006; Ricciardi, 2007). They can negatively affect native species through predation and competition, can diffuse pathogens, can modify ecosystem functioning and abiotic features of the environments (Strayer *et al.*, 2006; Ricciardi, 2007). The prevention and control of AIS is thus a priority for conservation (Hulme, 2006).

Biological invasions can be described as a multistep process, comprising three major phases: initial dispersal (i.e. an organism moves long distances to areas outside its native range, for example through human assisted dispersal); establishment of self-sustaining populations within the non-native range; and invasion of the new range (Richardson *et al.*, 2000; Puth & Post, 2005). However, when species are invasive, they have strongly positive demographic trends and are often numerous, therefore their management can be extremely difficult and expensive (Hulme, 2006). For these

DOI: 10.1111/j.1472-4642.2008.00516.x

reasons, the first stages of invasions are the most critical for conservation (Puth & Post, 2005; Hulme, 2006). Preventing new introductions is currently considered the most effective management tool (Leung *et al.*, 2002; Keller *et al.*, 2008). However, some species will become introduced even in the presence of strict protocols of prevention and screening. In these cases, a rapid response may still stop the invasion (Hulme, 2006).

Unfortunately, studying and managing the first steps of invasions (i.e. before that species become invasive) is often difficult. First of all, only a small subset of introduced species become invasive (approximately 1%: Williamson & Fitter, 1996; but see also Suarez *et al.*, 2005), and the active control and monitoring of all non-native species may be not feasible. Moreover, many AIS often undergo a clear 'lag phase' before they show any sign of becoming invasive (Crooks & Soulé, 1999). During lag phases, species are usually rare; the low detection rate limits our power to detect presence, assess demographic changes and to find the factors determining the transition from establishment to invasion (Hulme, 2006). Molecular genetics can provide insights into the demography of early stages of invasions (Lindholm *et al.*, 2005; Ficetola *et al.*, 2008), but these studies are usually performed *post hoc*, i.e. when species are already invasive.

Habitat modelling is a powerful approach to evaluate the factors determining species invasions. Through modelling, data on the distribution of AIS can be used to find the environmental features determining species distribution, and therefore to build maps for risk assessment (Thuiller *et al.*, 2005; see also Elith *et al.*, 2006). Modelling is usually based on distribution within the native range (Roura-Pascual *et al.*, 2004; Thuiller *et al.*, 2005; Ficetola *et al.*, 2007b), on records in the invasive range (Bossenbroek *et al.*, 2007; Ward, 2007; Nielsen *et al.*, 2008) or both (Broennimann *et al.*, 2007). Nevertheless, the presence of a species does not necessarily imply that it has positive fitness in an environment. This can be extremely important for species that are introduced multiple times, sometimes in suboptimal areas. Particularly in long-lived species, adults can survive long periods in suboptimal habitats, and may even attain high densities if massively released, but can fail to achieve key steps of their life cycle, such as reproduction. This can determine human-mediated source–sink dynamics. The distinction between localities of presence, and localities where non-native species have positive fitness, can be the key to understanding the causes of the transition among the different steps of invasions. However, comparisons of fitness among areas are rarely included in models of distribution of non-native species.

Turtles are long-lived organisms, and adults can survive for decades in suboptimal habitats where environmental features are unsuitable for breeding (Gariboldi & Zuffi, 1994; Spinks *et al.*, 2003). Therefore, even if they can negatively affect native populations, the presence of non-native turtles does not necessarily imply that a species is invasive and colonizing new habitats. The long life cycle of turtles determines slow population dynamics (Congdon *et al.*, 1993), making turtles particularly suitable for the study of transition between the different steps of invasions, and to find the factors influencing these transitions. The slider turtle *Trachemys scripta* Schoepffs, 1972 is a native of Eastern Northern and Central America, but has been introduced as a pet in some 30 countries around the world, with several million individuals sold during the past decades (Lever, 2003). Young sliders are sold at a size of just a few centimetres, but can grow quickly, and are released by owners in natural and seminatural wetlands. The importation of *T. scripta elegans* has been banned in the European Union (Commission Regulation 349/2003), although other subspecies are still sold, and individuals traded before the ban continue to be released in natural and seminatural wetlands. Sliders are considered a potential threat to European freshwater ecosystems. They compete for food and basking places with the threatened European pond turtle *Emys orbicularis*, and can increase its mortality (Cadi & Joly, 2003, 2004). Moreover, at high densities sliders can modify wetland vegetation and the communities of macroinvertebrates and amphibians (Teillac-Deschamps & Prevot-Julliard, 2006). However, not all the European slider 'populations' are reproductively active. Reproduction has been observed only in a limited number of localities of southern Europe (Spain (Pleguezuelos, 2002); France (Cadi *et al.*, 2004); Italy (Sindaco *et al.*, 2006)), while most feral sliders are assumed to live in areas too cold for a successful reproduction. Nevertheless, we lack large scale and quantitative analysis of environmental features discriminating between reproductive and non-reproductive populations. An objective assessment of these features would identify populations that have the highest risk of becoming invasive, and thus the priority areas for management actions such as eradication.

The aim of this study was twofold. First, we assessed the factors influencing the distribution of the slider turtle in Italy, by analysing bioclimatic features that can cause the transition from presence of feral adults to breeding populations. Ongoing climate change is quickly modifying environmental conditions, and can affect invasion dynamics (Thuiller *et al.*, 2007). We therefore evaluated whether climate change might increase the suitability for reproduction, and boost the risk of establishment and invasion of this species in the near future.

## METHODS

### Distribution data

To obtain reliable distribution data over extensive areas throughout Italy, we combined data from the literature, from regional herpetological atlas (Lapini *et al.*, 1999; Bologna *et al.*, 2000; Bologna *et al.*, 2003; Fiacchini, 2003; Bernini *et al.*, 2004; Ragni *et al.*, 2006; Vanni & Nistri, 2006), the Italian fauna data base (Ruffo & Stoch, 2005), the ongoing Italian herpetological monitoring (Societas Herpetologica Italica, 2008), direct field surveys and personal communications from field herpetologists. We defined 'reproducing populations' to be populations where hatchling emergence has been observed (e.g. Lapini *et al.*, 1999; Ficetola *et al.*, 2003) excluding localities where females lay eggs but hatchling emergence has never been observed.

Our study was focused on Central and Northern Italy, because most of observations came from this area (Fig. 1). Observations in Southern Italy were scant, and this was probably related to

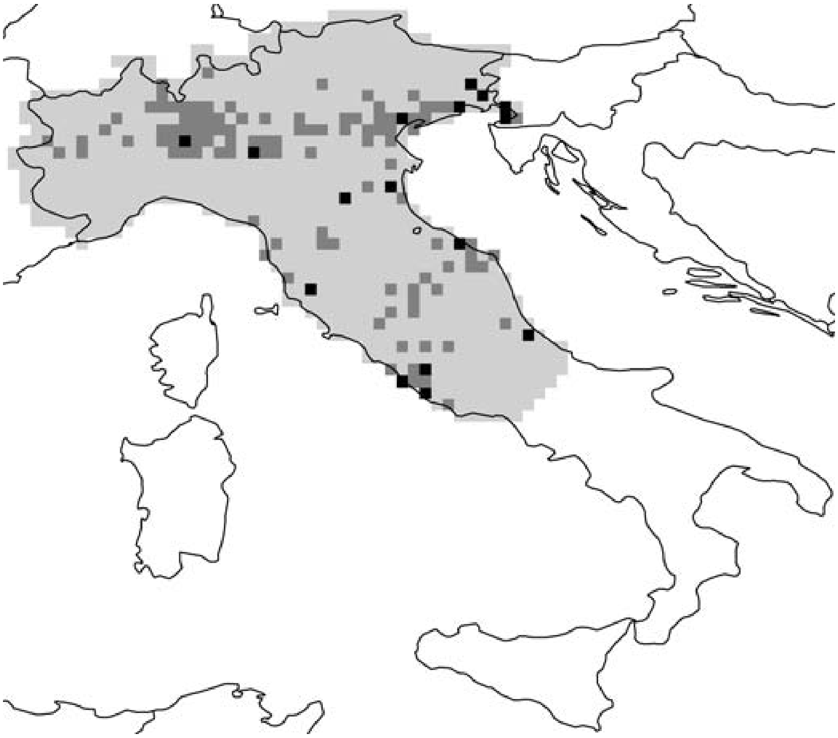


Figure 1 Study area in Italy (pale grey), and distribution of localities with feral (dark grey) and reproducing populations (black) of *Trachemys scripta*.

both reduced frequency of introduction and lower monitoring effort in these areas (see Sindaco *et al.*, 2006).

We did not use data of distribution of the slider turtle within the native range, since *T. scripta* has complex systematic with strong differentiation among subspecific entities (Stephens & Wiens, 2003). Individuals with different origin might have different climatic tolerance, but the subspecies and the origin of individuals introduced in Europe are often unknown and changed in time. In these conditions, models based on the introduced range only can be more informative (Steiner *et al.*, 2008). Therefore, our analysis describes the realized niche of introduced populations.

## Environmental data

For the climatic parameters, the CRU CL 2.0 global data set at 10′ × 10′ served as the base data set (New *et al.*, 2002). At the latitude of the study area, each 10′ × 10′ cell corresponds to a rectangle of approximately 13 × 18 km. Two temperature variables (winter temperature: average temperature in the coldest month; summer temperature: average temperature in the warmest month) and annual solar radiation (Wh/m$^2$/day) described the species thermal tolerance and the availability of thermal energy. Summed annual precipitation described the water availability. To avoid the multicollinearity issue, we did not include other climatic variables (such as annual temperature and seasonal precipitation) that were strongly correlated to linear combinations of the four climatic variables used. Furthermore, we used the human footprint at 10′ × 10′, a measure of human influence on global surface, combining data of population density, land transformation, human access and presence of infrastructures. This information was based on nine geographical data sets including satellite images, vector maps and census data (Sanderson *et al.*, 2002). Human footprint was used because the slider turtle and many other alien species are often associated with human modified landscapes (Ficetola *et al.*, 2004; Leprieur *et al.*, 2008).

Four climate-change scenarios were derived from HadCM3 (Hadley Centre for Climate Prediction and Research's General Circulation Model) for the period of 2010 to 2039 (referred to as the 2020 scenario) to obtain estimates of likely climatic conditions in the near future. The different global circulation model was run using four IPCC SRES (Intergovernmental Panel on Climate Change, Special Report on Emission Scenarios) storyline runs, reflecting different assumptions about demographic changes, socioeconomic and technological development (Nakicenovic & Swart, 2000). These include A1, A2, B1 and B2, ranging from fossil-fuel intensive to alternative futures involving rapid adoption of new technologies. This range of scenarios gives some idea of the range of greenhouse gas emission pathways that might be taken during the next decades.

## Data analysis

Environmental suitability was modelled using Maxent 3.1 (Phillips *et al.*, 2006; Phillips & Dudík, 2008). Maxent is a machine learning method that estimates the distribution of a species by finding the probability distribution of maximum entropy (i.e. that is closest to uniform) subject to constraints representing our incomplete information about the distribution. The constraints are that the expected value of each environmental variable should match its average over sampling locations derived from environmental layers (Phillips *et al.*, 2006). The model evaluates the suitability of each grid cell as a function of environmental variables at that cell. Some advantages of Maxent are that it requires presence-only data, deterministic algorithms have been developed that always converge to the optimal probability distribution, and can calculate the relative importance of different environmental variables (Phillips *et al*., 2006). We used a logistic output of Maxent, with suitability values ranging from 0 (unsuitable habitat) to 1 (optimal habitat) (Phillips & Dudík, 2008). Following Pearson *et al.* (2007), we used the 10th percentile training presence as a suitability threshold, i.e. we assumed that a cell is suitable if its suitability score is greater than the 10th percentile of training presence points. We developed models using linear, quadratic and hinge functions (Phillips & Dudík, 2008). In recent comparisons among several techniques of prediction of species distribution, Maxent was among the most effective methods using presence-only data, and showed a particularly good performance when analysing data sets with a small number of presence records (Elith *et al*., 2006; Hernandez *et al*., 2006). The reliability of the results of Maxent has been confirmed by its good capacity to predict novel presence localities for poorly known species (Pearson *et al*., 2007) and the outcome of introductions of invasive species outside the native range (Ficetola *et al*., 2007b).

Using Maxent, we built two models to compare the bioclimatic envelopes corresponding to different levels of fitness. First, we built a model using presence data. This model describes the environmental features needed for the presence and survival of adults, independently from reproduction. The second model used data on the distribution of breeding localities, and described the environmental features required for successful reproduction (i.e. for the establishment of naturalized populations). All models included the five environmental variables as predictors.

We used null-models to test for significance of Maxent models (Raes & ter Steege, 2007). For each model based on turtle distribution, we generated 99 null-distributions of random points in the study area; the number of random points per distribution was equal to the actual number of presence points. Next, we used Maxent to create models relating the null-distributions to the environmental layers. We then compared the area under the curve of the receiver operator plot (AUC) (Manel *et al*., 2001) of the randomly generated models with the AUC of the models generated using the actual distribution data. These randomly generated models can be thus used as a null-hypothesis against which to test the significance of species distribution models. If the AUC of the slider turtle models was significantly higher than the AUC of randomly generated models, it was considered as evidence that the species distribution model performs significantly better than expected by chance (Raes & ter Steege, 2007).

We used the jackknife procedure developed by Pearson *et al.* (2007) to evaluate the predictive performance of our model, i.e. the ability to correctly predict new localities of reproduction. Each observed locality of reproduction was removed once from the data set and a model was built using the remaining $n - 1$ localities. The predictive performance was then evaluated based on the ability of each model to correctly predict the locality excluded from the training data set, using the 10th percentile training presence threshold (Pearson *et al*., 2007). This approach is valid only for small data sets (less than 25 calibration points: Pearson *et al*., 2007), and was therefore applied to the analysis of reproduction.

To compare bioclimatic envelopes of presence points and of reproducing populations, we compared the response curves obtained using Maxent. Maxent curves provide only qualitative estimates of differences in niches, also because confidence intervals are not available. Nevertheless, they can provide useful insights into differences among distribution models (Martínez-Freiría *et al*., 2008).

Finally, we used the four 2020 scenarios to project the predicted suitability to the future climatic conditions. Estimations of the future human footprint are not available. However, it is unlikely that human footprint in Europe will decrease in the near future, and slider turtles are positively associated with human footprint (see results). Therefore, models assuming a constant human footprint are conservative in respect of the future suitability.

## RESULTS

Feral populations of slider turtle were present in 121 $10' \times 10'$ pixels, corresponding to 14% of the study area (Fig. 1). However, reproduction was observed only in a small subset of localities (16 pixels) (Fig. 1).

Bioclimatic features were different among pixels with and without observed reproduction. Pixels with reproduction had higher summer (unequal variance *t*-test: $t_{57} = 2.687$, $P = 0.009$) and winter temperature ($t_{20} = 2.686$, $P = 0.014$) and more annual radiation ($t_{21} = 2.462$, $P = 0.023$) than pixels where reproduction was not observed. We did not observe significant differences in annual precipitation and human footprint (both $P > 0.75$).

### Suitability models

The variables most important to explain the presence of feral populations were summer temperature (61.1% of explained variation accounted), and human footprint (16% of variation); annual precipitation and solar radiation accounted for 11.2 and 10.5% of variation, respectively, while winter temperature explained only a minor portion of variance. The presence of feral populations of slider turtles was positively related to summer temperature, annual precipitation and human footprint, while the relationship with solar radiation showed a less clear, non-linear pattern (Fig. 2). The areas with the highest probability of containing feral populations were mostly in Northern Italy, close to the largest cities such as Milan, Turin and Venice, and in Central Italy around the cities of Rome and Florence (Fig. 3). The AUC of this model was 0.828, and was significantly higher than the null-model AUC (median = 0.603; 95% CI = 0.564–0.644). This indicates a good fit of the model.

The model describing the features of reproduction areas was quite different. The most important variable to explain the presence of reproducing populations was summer temperature (81.5% of variation). Annual radiation and precipitation explained 7.4% and 6.5% of variation, respectively, while human footprint and winter temperature explained only a minor

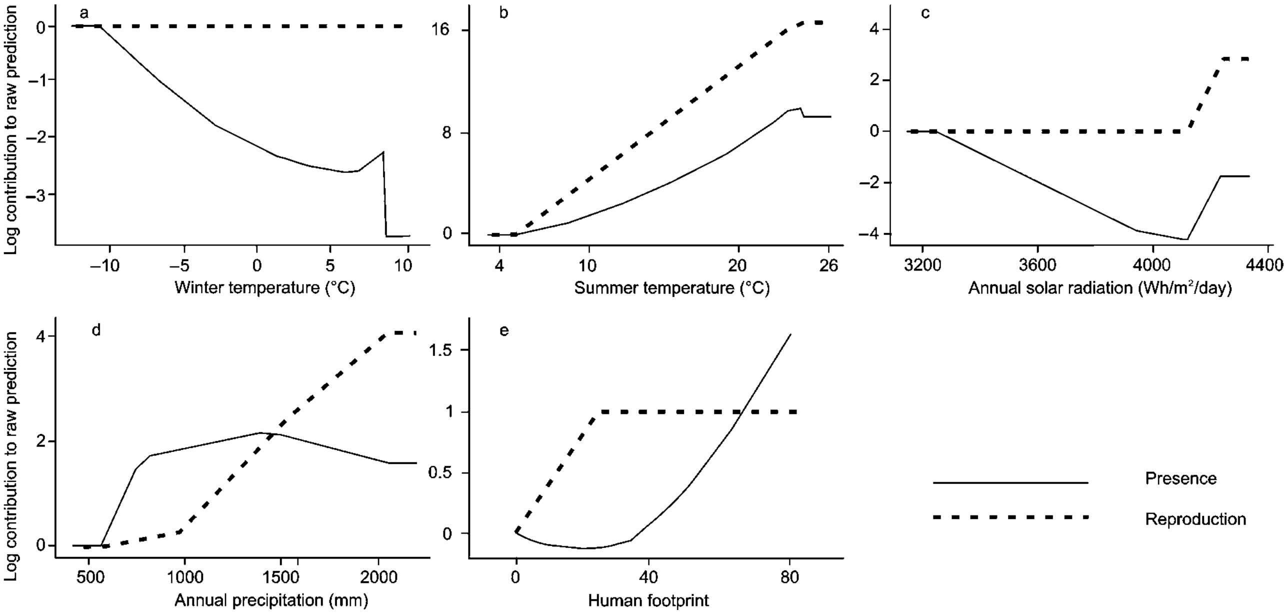


Figure 2 Results of environmental suitability models: relationships between environmental features and presence (continuous line) or reproduction (broken line) of *Trachemys scripta*.

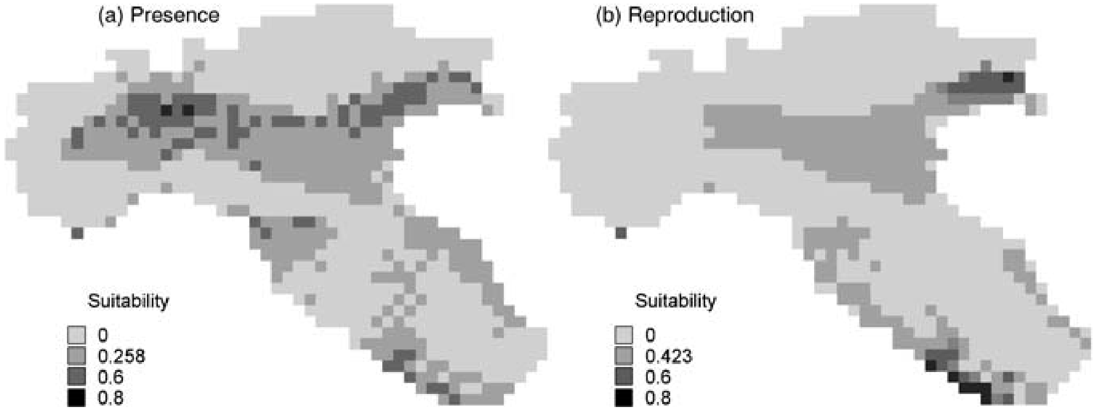


Figure 3 Results of environmental suitability models: predicted probability of (a) presence and (b) reproduction of *Trachemys scripta*. The different suitability thresholds (0.258 and 0.423) correspond to the 10th percentile training presence thresholds of the presence and reproduction models, respectively.

portion of variance. The reproduction of slider turtles was related to areas with high summer temperature, solar radiation and annual precipitation (Fig. 2). The AUC of this model was 0.868, and was significantly higher than the AUC of null-model (median = 0.725, 95% CI = 0.640–0.803). This indicates a good fit of the model. The jackknife procedure showed that the model had a good capacity to predict new localities of reproduction (predictive success = 0.75, average probability of success under randomness = 0.32, $P = 0.0005$).

The areas most suitable for reproduction did not correspond exactly to the areas where the probability of presence of feral populations is highest (Fig. 3). Probability of reproduction was highest in the Mediterranean and coastline areas. Using the 10th percentile training presence threshold, 42% of the study area was suitable for the presence of slider turtles, while only 27% was suitable for reproduction (Fig. 3).

The comparison of response curves confirmed the differences between the bioclimatic niches obtained using presence and reproduction data (Fig. 2). Slider turtle reproduction showed a more positive relationship with solar radiation, summer and winter temperature than slider turtle presence.

## Future suitability

The projection of suitability for reproduction using the bioclimatic features of the 2020 scenarios showed a clear increase of suitable areas (Fig. 4, see Appendix S1 in Supporting Information). Despite minor differences among the four 2020 scenarios (Fig. 4,

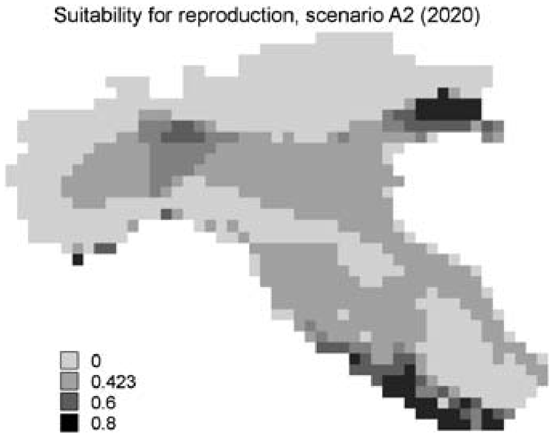


Figure 4 Projected suitability for reproducing populations of *Trachemys scripta* under future climatic conditions (2020, scenario A2). Results of projections using different scenarios (A1, B1 and B2) are extremely similar and are shown in Appendix S1.

Appendix S1), all of them predicted an expansion of suitable areas far from the coastline, in the Northern regions and in most of areas where feral populations are currently present (compare with Fig. 1). For example, following the A2 scenario, 87% of the 121 pixels where feral populations are currently present will be suitable for reproduction in 2020.

## DISCUSSION

Under present-day environmental conditions, the bioclimatic envelope entailed by the presence of feral populations of slider turtle is markedly different from the envelope of populations where reproduction actually occurs. This indicates that most of the feral populations are introduced and survive in suboptimal environment, where bioclimatic conditions currently are not suitable for reproduction. However, lack of reproduction in many feral populations (Luiselli *et al.*, 1997; Bringsøe, 2001; Prévot-Julliard *et al.*, 2007) does not mean that the slider turtle will quickly become extinct before establishment. These long-lived species can persist for decades and thus influence the native species even in absence of actual reproduction. Our analysis clearly shows that the environmental features required for reproduction are currently present in several areas of Mediterranean Europe (Fig. 3), and that ongoing climate change will likely expand the areas of suitability in the near future (Fig. 4, Appendix S1).

The distribution of reproductive populations delineates a bioclimatic envelope corresponding to areas with high summer temperature, intense solar radiation and rather high annual precipitation (Fig. 2). The identification of a clear envelope, distinct from the one of non-reproductive populations, shows that reproduction does not occur in a random subset of presence localities, but is instead related to well-defined environmental parameters. The association with high temperature and solar radiation likely describes physiological limitations and the large amount of thermal energy required for embryo development. For example, the optimum temperature for slider activity is 25–26 °C, and the optimum temperature for embryo development is above 25 °C (Cagle, 1950; Morreale & Gibbons, 1986; Cadi *et al.*, 2004). Our model describes very well these requirements, and maximum suitability for both presence and reproduction was at about 26 °C (Fig. 2b). Moreover, embryos require a wet substrate for correct development (Tucker & Paukstis, 2000). Therefore, the most arid Mediterranean areas are not suitable for reproduction, which explains the positive relationship with annual precipitation. High precipitation can be also associated to the presence of wetlands, with obvious positive effects on populations of a freshwater turtle. Precipitation, temperature and solar radiation are key drivers of the energy/water balance in ectotherms, and can therefore be important also for other invaders (e.g. reptiles, amphibians and crustaceans).

Human footprint was only positively correlated to species presence but not to species reproduction. This is intuitive, as species introduction into the environment occurs through pet release or escape and pets are most likely to be kept in human-modified areas, but human modifications of landscape do not have a positive effect on turtle reproduction.

The smallest predicted range for reproduction (compare Fig. 3a,b) suggests that adult sliders can survive under a wide range of environmental conditions, but they have more stringent requirements for critical phases of life history, such as reproduction. The AUC of models of reproductive populations (0.868) was higher than the AUC of localities of presence (0.828). This further confirms that reproductive populations have a small but well-defined niche, because distribution models of generalist species often have lower AUC (Brotons *et al.*, 2004; Allouche *et al.*, 2006). Species distribution models usually rely on occurrence data, without any knowledge of the actual fitness of the population in those areas. However, a more complete modelling of ecological niche should define the environmental conditions where the fitness of individuals is greater than one (Guisan & Thuiller, 2005; Kearney, 2006). Simply considering distribution records does not allow full delineation of a species fundamental niche (Kearney, 2006), and can predict suitability into too large areas. Moreover, the projection of bioclimatic models in the future, or into new geographical areas, is a critical phase of the application of bioclimatic models to conservation (Guisan & Thuiller, 2005). The differences between the suitability obtained using presence records, and the one obtained using fitness measures could be amplified when bioclimatic niches are projected to new conditions.

For these reasons, using fitness measures when available instead of occurrence data can lead to great improvement in the quality of species distribution models. The improvement can be particularly valuable for alien species, for species with high vagility that are often observed in suboptimal environments, and for species with strong source-sink dynamics. The availability of modelling techniques having good performance even with a few presence localities, such as Maxent (Hernandez *et al.*, 2006), can be particularly useful, because collecting data on species presence/absence is clearly more cost-effective than comparing fitness among localities.

### From introduction to establishment: conservation implications

Only a subset of feral populations is currently breeding, and their reproductive success is probably not very high. Nevertheless, the small number of localities where reproduction has been ascertained should not be used to underestimate the risk of establishment of slider turtle populations. First, detecting reproduction of freshwater turtles can be difficult, because juveniles are less detectable than adults (Zuffi, 2000) and small turtles are still released in natural wetlands by unconcerned people. The actual number of reproductive populations is probably larger than that reported here (see also Cadi *et al.*, 2004). Second, some peculiar features of introduced sliders may increase their recruitment potential. Turtles sold as pets are reared in farms with artificially high temperature, to accelerate the development rate. As sex determination in the slider is temperature dependent, the sex ratio of introduced turtles is unbalanced, with a prevalence of females (Cadi *et al.*, 2004). If environmental conditions are suitable, this high proportion of females may increase the number of recruits of introduced populations (see Girondot *et al.*, 1998).

Most importantly, ongoing global climate change will probably boost the fitness of introduced individuals in the near future. Long lived animals can survive decades in suboptimal habitats. As the longevity of sliders in nature is at least 30 years (Gibbons & Semlitsch, 1982), the individuals that are released now can survive for 20 years or so in suboptimal areas, where bioclimatic features are currently unsuitable for reproduction. These feral individuals may successfully reproduce during particularly warm years, or in the near future, when suitability will increase (compare Figs 1 and 4 and the Appendix S1). The successful reproduction of these populations corresponds to the transition from introduced individuals to established alien populations. In many cases, alien species became invasive and problematic after lag phases that may last decades, and then suddenly explode (Crooks & Soulé, 1999). Climate change can be a key factor triggering this process.

This is therefore a case where species distribution models have direct and practical implications for the management. Immediate management actions (such as the eradication of feral individuals) have high probability of success. Currently, many non-native populations of the slider are non-reproductive. These feral sliders can have negative effects on native biodiversity (Cadi & Joly, 2003, 2004; Teillac-Deschamps & Prevot-Julliard, 2006), but their eradication can be successful even with moderate effort, since current recruitment is zero, and trapping techniques have a high success rate on adult freshwater turtles (Fowler & Avery, 1994). However, if the species becomes naturalized, eradication will be much more difficult and costly. Management efforts should first focus on the areas where suitability for reproduction is highest (Fig. 3b), because these areas are sources of biological invasion. It is also important to act promptly, to remove individuals from areas where suitability is predicted to increase in the near future. Off course, these management actions will be successful only if education programs will help stopping the release of new turtles in natural environments (Ficetola *et al.*, 2007a).

It is widely accepted that the ongoing global climate change can exacerbate the issues of biological invasions, but we are far from a full understanding of the mechanisms facilitating the response of AIS to the change (Thuiller *et al.*, 2007; Richardson & Pysek, 2008). Differences in fitness among areas of the introduced ranges can strongly interact with climatic change, complicating the invasion dynamics, and should be considered in models trying to predict the future of invasions.

## ACKNOWLEDGEMENTS

We thank all field herpetologists providing unpublished observations, particularly A. Romano, L. Lapini and U. Fusini. G. F. Ficetola was founded by a grant of EVK2-CNR on biodiversity conservation. W. Thuiller was partly funded by the EU FP6 MACIS specific targeted project (Minimization of and adaptation to climate change: impacts on biodiversity no. 044399) and EU FP6 ECOCHANGE integrated project (Challenges in assessing and forecasting biodiversity and ecosystem changes in Europe). The comments of M. Burgman, J. Ott and two anonymous reviewers improved an earlier version of the manuscript.

## REFERENCES

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Bernini, F., Bonini, L., Ferri, V., Gentilli, A., Razzetti, E. & Scali, S. (2004) *Atlante degli Anfibi e dei Rettili della Lombardia*. Provincia di Cremona, Cremona, Lombardy, Italy.

Bologna, M.A., Capula, M. & Carpaneto, G.M. (2000) *Anfibi e rettili del Lazio*. Fratelli Palombi Editori, Rome, Italy.

Bologna, M.A., Capula, M., Carpaneto, G.M., Cignini, B., Marangoni, C., Venchi, A. & Zapparoli, M. (2003) *Anfibi e rettili a Roma – Atlante e guida delle specie presenti in città*. Comune di Roma, Stilgrafica srl, Rome, Italy.

Bossenbroek, J.M., Johnson, L.E., Peters, B. & Lodge, D.M. (2007) Forecasting the expansion of zebra mussels in the United States. *Conservation Biology*, **21**, 800–810.

Bringsøe, H. (2001) *Trachemys scripta* (Schoepff, 1792) – Buchstaben-Schmuckschildkröte. *Handbuch der Reptilien und Amphibien Europas. schildkröten (Testudines)* (ed. by U. Fritz), pp. 525–583. AULA, Wiebelsheim, Germany.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T.P. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.

Brotons, L., Thuiller, W., Araujo, M.B. & Hirzel, A.H. (2004) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437–448.

Cadi, A., Delmas, V., Prévot-Julliard, A.-C., Joly, P. & Girondot, M. (2004) Successful reproduction of the introduced slider turtle (*Trachemys scripta elegans*) in the South of France. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **14**, 237–246.

Cadi, A. & Joly, P. (2003) Competition for basking places between the endangered European pond turtle (*Emys orbicularis galloitalica*) and the introduced red-eared slider (*Trachemys scripta elegans*). *Canadian Journal of Zoology*, **81**, 1392–1398.

Cadi, A. & Joly, P. (2004) Impact of the introduction of the read-eared slider (*Trachemys scripta elegans*) on survival rates of the European pond turtle (*Emys orbicularis*). *Biodiversity and Conservation*, **13**, 1511–2518.

Cagle, F.R. (1950) The life history of the slider turtle, *Pseudemys scripta troostii* (Hoolbrook). *Ecological Monographs*, **20**, 31–54.

Congdon, J.D., Dunham, A.E. & Van Loben Sels, R.C. (1993) Delayed sex maturity and demographics of Blanding's turtles (*Emydoidea blandingii*): implications for conservation and management of long-lived organisms. *Conservation Biology*, **7**, 826–833.

Crooks, J. & Soulé, M.E. (1999) Lag times in population explosions of invasive species: causes and implications. *Invasive species and biodiversity management* (ed. by O.T. Sandlund, S.J. Schei and A. Viken), pp. 103–125. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Elith, J., Graham, C.H., Anderson, R.P., Dudýk, 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.M., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Fiacchini, D. (2003) *Atlante degli Anfibi e dei Rettili della Provincia di Ancona*. Nuove Ricerche Editore, Ancona, Italy.

Ficetola, G.F., Monti, A. & Padoa-Schioppa, E. (2003) First record of reproduction of *Trachemys scripta* in the Po Delta. *Annali Del Museo Civico Di Storia Naturale Di Ferrara*, **5**, 125–128.

Ficetola, G.F., Padoa-Schioppa, E., Monti, A., Massa, R., De Bernardi, F. & Bottoni, L. (2004) The importance of aquatic and terrestrial habitat for the European pond turtle (*Emys orbicularis*): implications for conservation planning and management. *Canadian Journal of Zoology*, **82**, 1704–1712.

Ficetola, G.F., Coïc, C., Detaint, M., Berroneau, M., Lorvelec, O. & Miaud, C. (2007a) Pattern of distribution of the American bullfrog *Rana catesbeiana* in Europe. *Biological Invasions*, **9**, 767–772.

Ficetola, G.F., Thuiller, W. & Miaud, C. (2007b) Prediction and validation of the potential global distribution of a problematic alien invasive species – the American bullfrog. *Diversity and Distributions*, **13**, 476–485.

Ficetola, G.F., Bonin, A. & Miaud, C. (2008) Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology*, **17**, 773–782.

Fowler, J.F. & Avery, J.L. (1994) Turtles. *Prevention and control of wildlife damage* (ed. by S.E. Hygnstrom and R. M. Timm), pp. F27–F31. University of Nebraska, Lincoln, Nebraska.

Gariboldi, A. & Zuffi, M.A.L. (1994) Notes on the population reinforcement project for *Emys orbicularis* (Linnaeus, 1758) in a natural park of northwestern Italy. *Herpetozoa*, **7**, 83–89.

Gibbons, J.W. & Semlitsch, R.D. (1982) Survivorship and longevity of a long-lived vertebrate species: how long do turtles live? *Journal of Animal Ecology*, **51**, 25.

Girondot, M., Fouillet, H. & Pieau, C. (1998) Feminizing turtle embryos as a conservation tool. *Conservation Biology*, **12**, 353–362.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773–785.

Hulme, P.E. (2006) Beyond control: wider implications for the management of biological invasions. *Journal of Applied Ecology*, **43**, 835–847.

Kearney, M. (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186–191.

Keller, R.P., Frang, K. & Lodge, D.M. (2008) Preventing the spread of invasive species: economic benefits of intervention guided by ecological predictions. *Conservation Biology*, **22**, 80–88.

Lapini, L., dall'Asta, A., Bressi, N., Dolce, S. & Pellarini, P. (1999) *Atlante corologico degli anfibi e dei rettili del Friuli-Venezia Giulia*. Museo Friulano di Storia Naturale, Udine, Italy.

Leprieur, F., Beauchard, O., Blanchet, S., Oberdorff, T. & Brosse, S. (2008) Fish invasions in the world's river systems: when natural processes are blurred by human activities. *PloS Biology*, **6**, 404–410.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **269**, 2407–2413.

Lever, C. (2003) *Naturalized amphibians and reptiles of the world*. Oxford University Press, New York.

Lindholm, A.K., Breden, F., Alexander, H.J., Chan, W.-K., Thakurta, S.G. & Brooks, R. (2005) Invasion success and genetic diversity of introduced populations of guppies *Poecilia reticulata* in Australia. *Molecular Ecology*, **14**, 3671–3682.

Luiselli, L., Capula, M., Capizzi, D., Filippi, E., Jesus, V.T. & Anibaldi, C. (1997) Problems for conservation of pond turtles (*Emys orbicularis*) in Central Italy: is the introduced Red-Eared Turtle (*Trachemys scripta*) a serious threat? *Chelonian Conservation and Biology*, **2**, 417–419.

Manel, S., Williams, H.C. & Ormerod, S.J. (2001) Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, **38**, 291–931.

Martínez-Freiría, F., Sillero, N., Lizana, M. & Brito, J.C. (2008) GIS-based niche models identify environmental correlates sustaining a contact zone between three species of European vipers. *Diversity and Distributions*, **14**, 452–461.

Morreale, S.J. & Gibbons, J.W. (1986) *Habitat suitability index models: slider turtle*. US Fish and Wildlife Service Biological Report 82. US Fish and Wildlife Service, Washington.

Nakicenovic, N. & Swart, R. (2000) *Emissions scenarios: a special report of working group III of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, UK.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*, **21**, 1–25.

Nielsen, C., Hartvig, P. & Kollmann, J. (2008) Predicting the distribution of the invasive alien *Heracleum mantegazzianum* at two different spatial scales. *Diversity and Distributions*, **14**, 307–317.

Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and comprehensive evaluation. *Ecography*, **31**, 161–175.

Pleguezuelos, J.M. (2002) Las especies introducidas de Amphibios y Reptiles. *Atlas y Libro Rojo de los Anfibios y Reptiles de España* (ed. by J.M. Pleguezuelos, R. Marquez and M. Lizana), pp. 501–532. Dirección General de Conservación de la Naturaleza-Asociación Herpetologica Española, Madrid, Spain.

Prévot-Julliard, A.C., Gousset, E., Archinard, C., Cadi, A. & Girondot, M. (2007) Pets and invasion risks: is the Slider turtle strictly carnivorous? *Amphibia-Reptilia*, **28**, 139–143.

Puth, L.M. & Post, D.M. (2005) Studying invasion: have we missed the boat? *Ecology Letters*, **8**, 715–721.

Raes, N. & ter Steege, H. (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727–736.

Ragni, B., Di Muro, G., Spilinga, C., Mandrici, A. & Ghetti, L. (2006) *Anfibi e Rettili dell'Umbria*. Petruzzi editore, Perugia, Umbria, Italy.

Ricciardi, A. (2007) Are modern biological invasions an unprecedented form of global change? *Conservation Biology*, **21**, 329–336.

Richardson, D.M. & Pysek, P. (2008) Fifty years of invasion ecology – the legacy of Charles Elton. *Diversity and Distributions*, **14**, 161–168.

Richardson, D.M., Pysek, P., Rejmanek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, **6**, 93–107.

Roura-Pascual, N., Suarez, A.V., Goomez, C., Pons, P., Touyama, Y., Wild, A.L. & Peterson, A.T. (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **271**, 2527–2534.

Ruffo, S. & Stoch, F. (2005) *Checklist e distribuzione della fauna Italiana*. Memorie del Museo Civico di Storia Naturale di Verona, 2a serie, sezione Scienze della Vita, Verona, Italy.

Sanderson, E.W., Jaiteh, M., Levy, M.A., Redford, K.H., Wannebo, A.V. & Woolmer, G. (2002) The human footprint and the last of the wild. *Bioscience*, **52**, 891–904.

Societas Herpetologica Italica (2008) Osservatorio Erpetologico Italiano. *Acta Herpetologica*, **3**, 83–84.

Sindaco, R., Doria, G., Razzetti, E. & Bernini, F. (2006) *Atlas of Italian amphibians and reptiles*. Polistampa, Firenze, Italy.

Spinks, P.Q., Pauly, G.B., Crayon, J.J. & Shaffer, H.B. (2003) Survival of the western pond turtle (*Emys marmorata*) in an urban California environment. *Biological Conservation*, **113**, 257–267.

Steiner, F.M., Schlick-Steiner, B.C., VanDerWal, J., Reuther, K.D., Christian, E., Stauffer, C., Suarez, A.V., Williams, S.E. & Crozier, R.H. (2008) Combined modelling of distribution and niche in invasion biology: a case study of two invasive *Tetramorium* ant species. *Diversity and Distributions*, **14**, 538–545.

Stephens, P.R. & Wiens, J.J. (2003) Ecological diversification and phylogeny of emydid turtles. *Biological Journal of the Linnean Society*, **79**, 577–610.

Strayer, D.L., Evinver, V.T., Jeschke, J.M. & Pace, M.L. (2006) Understanding the long-term effects of species invasions. *Trends in Ecology and Evolution*, **21**, 645–651.

Suarez, A.V., Holway, D.A. & Ward, P.S. (2005) The role of opportunity in the unintentional introduction of nonnative ants. *Proceedings of the National Academy of Sciences USA*, **102**, 17032–17035.

Teillac-Deschamps, P. & Prevot-Julliard, A.C. (2006) Impact of exotic slider turtles on freshwater communities: an experimental approach. *First European congress of conservation biology, book of abstracts*. pp. 162–163. Society for Conservation Biology, Heger, Hungary.

Thuiller, W., Richardson, D.M. & Midgley, G.F. (2007) Will climate change promote alien plant invasions? *Biological invasions* (ed. by W. Nentwig), pp. 197–211. Springer-Verlag, Berlin/Heidelberg, Germany.

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.

Tucker, J.K. & Paukstis, G.L. (2000) Hatching success of turtle eggs exposed to dry incubation environment. *Journal of Herpetology*, **34**, 529–534.

Vanni, S. & Nistri, A. (2006) *Atlante degli Anfibi e Rettili della Toscana*. Regione Toscana, Firenze, Italy.

Ward, D.F. (2007) Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions*, **9**, 723–735.

Williamson, M. & Fitter, A. (1996) The varying success of invaders. *Ecology*, **77**, 1661–1666.

Zuffi, M.A.L. (2000) Conservation biology of the European pond turtle *Emys orbicularis* (L.) in Italy. *Stapfia*, **69**, 219–228.

Editor: Mark Burgman

## SUPPORTING INFORMATION

The following Supporting Information is available for this article:

**Appendix S1** Projected suitability for reproducing populations of *Trachemys scripta* under future climatic conditions (2020, scenarios A1, B1, B2).

*Global Ecology and Biogeography, (Global Ecol. Biogeogr.)* (2009) **18**, 214–222

RESEARCH PAPER

# Human activities alter biogeographical patterns of reptiles on Mediterranean islands

Gentile Francesco Ficetola* and Emilio Padoa-Schioppa

*Department of Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy*

*Correspondence: G. F. Ficetola, Department of Environmental Sciences, University of Milan-Bicocca, Piazza della Scienza 1, 20128 Milan, Italy.
E-mail: francesco.ficetola@unimi.it

## ABSTRACT

**Aim** The theory of island biogeography predicts species richness based on geographical factors that influence the extinction–colonization balance, such as area and isolation. However, human influence is the major cause of present biotic changes, and may therefore modify biogeographical patterns by increasing extinctions and colonizations. Our aim was to evaluate the effect of human activities on the species richness of reptiles on islands.

**Location** Islands in the Mediterranean Sea and Macaronesia.

**Methods** Using a large data set ($n = 212$ islands) compiled from the literature, we built spatial regression models to compare the effect of geographical (area, isolation, topography) and human (population, airports) factors on native and alien species. We also used piecewise regression to evaluate whether human activities cause deviation of the species–area relationship from the linear (on log–log axes) pattern, and path analysis to reveal the relationships among multiple potential predictors.

**Results** The richness of both native and alien species was best explained by models combining geographical and human factors. The richness of native species was negatively related to human influence, while that of alien species was positively related, with the overall balance being negative. In models that did not take into account human factors, the relationship between island area and species richness was not linear. Large islands hosted fewer native species than expected from a linear (on log–log axes) species–area relationship, because they were more strongly affected by human influence than were small islands. Path analysis showed that island size has a direct positive effect on reptile richness. However, area also had a positive relationship with human impact, which in turn mediated a negative effect on richness.

**Main conclusion** Anthropogenic factors can strongly modify the biogeographical pattern of islands, probably because they are major drivers of present-day extinctions and colonizations and can displace island biodiversity from the equilibrium points expected by theory on the basis of geographical features.

**Keywords**

**Alien species, Anthropocene, elevation, human influence, island area, isolation, Macaronesia, Mediterranean Sea, species richness, topography.**

## INTRODUCTION

MacArthur and Wilson's theory of island biogeography (MacArthur & Wilson, 1967) is one of the major theoretical foundations of biogeography and ecology (Whittaker & Fernández-Palacios, 2007). The theory of island biogeography assumes that species richness on oceanic islands depends on the equilibrium between species extinction, species colonization and speciation. Simple geographical features, such as island size and isolation, are primary drivers of the extinction–colonization balance, thus determining the species richness of islands. Even though more complex models have been proposed, including factors such as energy availability (Wright, 1983; Kalmar & Currie, 2006), habitat heterogeneity (Diamond, 1969; Kadmon & Allouche, 2007) and disturbance (Whittaker, 1995), predictions from island biogeography models have repeatedly been confirmed

DOI: 10.1111/j.1466-8238.2008.00433.x

by both experimental and observational studies all over the world (Whittaker & Fernández-Palacios, 2007): area and isolation are still considered the key drivers of insular richness (e.g. Kreft *et al.*, 2008).

Habitat alteration and biological invasions are major causes of the current biodiversity crisis: human activities are thus directly responsible for most present-day extinctions, changes in community assemblies and modifications in the geographical ranges of species (Sala *et al.*, 2000; Wilson, 2002). Islands can be particularly vulnerable to human influences (Steadman, 1993; Whittaker & Fernández-Palacios, 2007). It is therefore likely that human activities are major drivers of the current extinction–colonization processes in oceanic islands. Several studies have documented biotic changes caused by humans, showing strong relationships between human activities, extinctions and the presence of alien species (Diamond, 1969; Chown *et al.*, 1998; Blackburn *et al.*, 2004; Blackburn *et al.*, 2008). However, large-scale studies are scarce, and we are far from a clear understanding of the effects of human factors on overall biogeographical patterns.

Here, we use an extensive database on reptiles in the Mediterranean Sea and Macaronesia (Fig. 1) to investigate whether anthropogenic factors can modify the biogeographical pattern of islands, and can cause a deviation of the species–area relationship from the expected linear (on log–log axes) pattern. The study area is a climatically homogeneous region, and some authors consider Macaronesia to be part of the Mediterranean area (Daget, 1977; Cox *et al.*, 2006). The islands of the Mediterranean Sea share a large proportion of reptiles with mainland Europe and Africa, but there are also several endemic species, particularly of lizards (e.g. *Algyroides fitzingeri*, *Archaeolacerta bedriagae* and *Podarcis pityusensis*). The native reptiles of Macaronesia comprise endemic species of the genera *Tarentola* and *Chalcides*, and by the endemic genus *Galliota*. Mediterranean islands have been subject to strong human influence for millennia (Blondel & Aronson, 1999), and constitute an ideal laboratory for evaluating the long-term effects of human activities on biodiversity patterns.

We tested three hypotheses describing the effect of human impact on the diversity of reptiles in Mediterranean islands. First, we evaluated whether human impact affects the biogeographical relationship of native species, by determining a lower species richness than predicted on the basis of classical species–area relationships. Second, we evaluated whether human impact is the major determinant of the richness of alien species. Finally, we tested whether the loss of species due to humans is compensated by the gain of alien species caused by human impact (Sax *et al.*, 2002).

## METHODS

### Data

We used faunistic atlases and the scientific literature (Lanza & Poggesi, 1986; Castanet & Guyetant, 1989; Delaguerre & Cheylan, 1992; Pleguezuelos *et al.*, 2002; Sultana & Falzon, 2002; Anonymous, 2003; Malkmus, 2004; Sindaco *et al.*, 2006) to obtain data on the reptile species composition on 212 islands of France, Italy, Malta, Portugal, Spain and Tunisia (Fig. 1). Following the literature (Pleguezuelos *et al.*, 2002; Lever, 2003; Sindaco *et al.*, 2006), species were classified either as native or as introduced by humans. For each island we obtained the area and the maximum elevation (a surrogate of habitat heterogeneity; e.g. Kreft *et al.*, 2008) using a variety of sources, including the original references, atlases, topographical maps and geographical publications, and we measured isolation as the distance from the continent or (if appropriate) from the nearest 'large island' (Corsica, Sardinia or Sicily). For each island, we also obtained human population size (years 2003–07) and the presence of airports from national statistical databases, atlases and geographical publications. These two variables were used to measure the human impact on the island. We considered airport presence as a surrogate measure of the intensity of commercial and human exchanges. Sea shipping may be more important than shipping by aircraft on islands. However, since nearly all the inhabited islands have a port, its presence could not be used as an independent variable. In our analysis, we included only islands with a surface ≥ 1 ha; this approximately corresponds to the size of the smallest islands inhabited by humans; islands with this area can host large population sizes of some reptiles (e.g. Buckley & Jetz, 2007; Pérez-Mellado *et al.*, 2008).

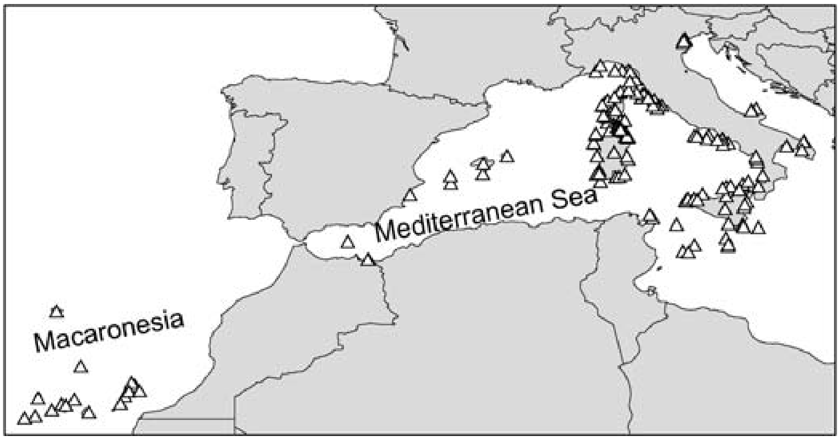


Figure 1 Study area and distribution of 212 analysed islands. Because of geographical proximity, several points are superimposed.

### Statistical analyses

The variation of environmental features and that of species richness (see Fig. 2) are spatially autocorrelated, and the presence of spatial autocorrelation may bias the results of classical statistical analyses (Dormann, 2007). In preliminary analyses, the residuals of ordinary least squares regression were significantly autocorrelated, particularly at small spatial scales ($P < 0.05$ at scales up to 300 km: Fig. 2). For this reason, we analysed the data using spatial eigenvector mapping (SEVM). This is a recently developed method allowing the translation of a spatial arrangement of data points into explanatory variables capturing the spatial effects (Dormann *et al.*, 2007). In SEVM, eigenvectors reducing the spatial autocorrelation of residuals are computed and then included as spatial predictors into generalized linear models. A recent comparison among statistical methods accounting for spatial autocorrelation showed that SEVM is one of the most flexible methods and is very efficient at removing residual spatial autocorrelation (Dormann *et al.*, 2007). We performed eigenvector selection following Griffith & Peres-Neto (2006). In our analyses, we carried out model selection based on the Akaike information criterion (AIC); therefore the pre-specified models (see below) were part of the eigenvector selection process. SEVM successfully removed spatial autocorrelation from the residuals of our models (Fig. 2).

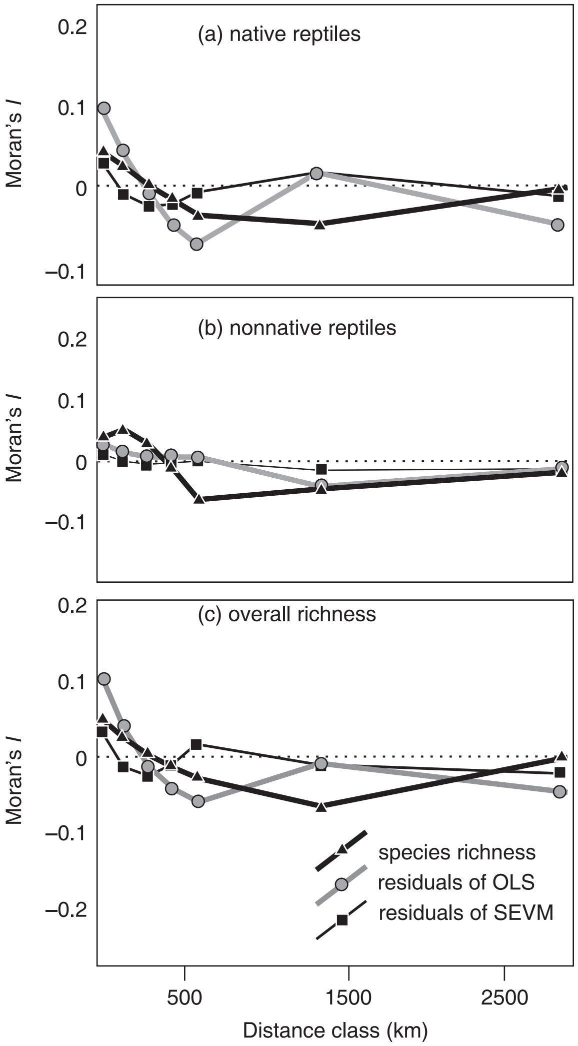


Figure 2 Moran's *I* correlograms for (a) richness of native reptiles; (b) richness of non-native reptiles and (c) overall richness of reptiles, measured. Bold black lines, number of species; grey lines, residuals of ordinary least squares regression (OLS); thin black lines, residuals of spatial eigenvector mapping (SEVM).

We then built a series of generalized linear models including richness of native and/or introduced species as dependent variables, five explanatory variables describing island geography and human presence as independent variables and the SEVM eigenvector(s). Species richness, island area, elevation, isolation and human populations were log-transformed to improve the normality of residuals. We transformed the richness of alien species using $\log(x + 0.5)$, since in several islands there were no alien species. Species richness is a count variable, therefore Poisson models can also be appropriate. However, the results of Poisson and normally distributed models were qualitatively very similar; we therefore report the models with normal errors only.

We used a model selection approach to evaluate the support of three different models, corresponding to three potential hypotheses explaining the pattern of reptile richness. Our first model (geographical-only model) hypothesized that geographical features (area, isolation and elevation) are the major drivers of species richness; this corresponds to the classical model of island biogeography. Our second model (human-only model) hypothesized a major influence of anthropogenic factors (total human population and the presence of airports). Finally, the geographical + human model hypothesized that species richness is best explained by a combination of both geographical and anthropogenic factors. We then used the AIC to rank models according to the strength of support from the data (Burnham & Anderson, 2002). We used Akaike weights to assess the likelihood of the alternative models, and we calculated evidence ratios to compare the relative likelihood of the different hypotheses (Burnham & Anderson, 2002).

We used piecewise regression (Muggeo, 2003; Toms & Lesperance, 2003) to evaluate the presence of threshold effects in the relationship between island area and species richness. Piecewise regression allows the simultaneous evaluation of the location and the standard error of break points in generalized linear models, and an evaluation of whether the addition of such break points improves the fit of the models.

Our independent environmental variables were moderately collinear, since larger islands had higher elevations and often hosted larger human populations and airports. The strongest correlations were between area and human population ($r = 0.84$), between human population and the presence of airports ($r = 0.72$) and between area and altitude ($r = 0.71$). Hence, we calculated the variance inflation factor for all regression models. For all models and variables, the variance inflation factor was always less than 5, indicating that collinearity does not pose major problems for our models (Bowerman & O'Connell, 1990).

Due to correlation among island area, human population and the presence of airports, it is difficult to tease apart their relative effect on species richness, and therefore to evaluate whether human activities modify the relationship between species richness and area. We therefore used path analysis (Bryman & Cramer, 1990; Bollen & Long, 1993) to compare three alternative models describing the relationship between island area and native species: (m1) island size affects species richness directly and through its relationship with human impact (i.e. human impact modifies the species–area relationship); (m2) both island size and human impact affect species richness, but there is no co-action among them; (m3) island size directly affects species richness, while human impact does not have an effect. Path analysis is an extension of regression models; it allows us to provide quantitative estimates of the causal connections among variables, in relation to explicitly formulated causal models, represented in the form of path diagrams (Bryman & Cramer, 1990). We performed the analyses with R 2.5 (R Development Core Team, 2007) using the packages *spdep*, *car*, *segmented* and *sem*. We built correlograms using SAM 3.0 (Rangel *et al.*, 2006).

## RESULTS

### Richness of native species

The number of native species per island ranged from 1 to 15. The 'geographical-only' model explained a substantial proportion of variation ($R^2 = 0.51$) of the richness of native reptiles (Table 1). Species richness related positively to island area and elevation, and negatively to isolation. The 'human-only' model explained a considerably lower proportion of variation ($R^2 = 0.24$). However, the model with the highest fit was the one assuming that both geographical and human factors drove the richness of native reptiles ($R^2 = 0.59$). This geographical + human model showed an AIC value substantially lower than that for either the geographical-only or human-only model, and had an AIC weight > 0.999 (Table 1). The AIC-evidence ratio between the geographical + human and the geographical-only model was about 11 million, meaning that the geographical + human model was about 11 million times more likely to be the best model. The evidence ratio between the geographical + human and the human-only model was even larger.

According to the geographical + human model, the richness of native reptiles increases with island area and elevation, and decreases with isolation. However, when taking into account these geographical variables, species richness is lower on islands with large human populations and with airports. The slope of the species–area relationship ($z$) on log–log scale for the geographical-only model was 0.08 [95% confidence interval (CI) = 0.06–0.1], and was much lower than the $z$-value obtained when using the model that takes into account human effects (geographical + human model: $z = 0.14$, CI = 0.12–0.17).

Human and geographical variables might also show interactive effects. When we tested for two-way interactions, we found a significant interaction between distance from the continent and human population ($F_{1,202} = 12.8$, $P = 0.0004$), indicating that the negative effect of human population was particularly strong on the most isolated islands.

**Table 1** Regression models relating richness of native reptiles to geographical and human factors.

| | *b* | *F* | d.f. | *P* |
|---|---|---|---|---|
| Model: geographical-only | | | | |
| ΔAIC = 32.47; $w < 0.0001$; $R^2 = 0.512$ | | | | |
| Area* | 0.08 | 61.1 | 1 | < 0.001 |
| Isolation* | –0.04 | 4.0 | 1 | 0.048 |
| Elevation* | 0.06 | 8.0 | 1 | 0.005 |
| SEVM eigenvectors | | | 3 | |
| Residuals | | | 205 | |
| Model: human-only | | | | |
| ΔAIC = 119.65, $w < 0.0001$, $R^2 = 0.239$ | | | | |
| Human population* | 0.05 | 21.7 | 1 | < 0.001 |
| Airport | –0.02 | 0.2 | 1 | 0.717 |
| SEVM eigenvectors | | | 3 | |
| Residuals | | | 206 | |
| Model: geographical + human | | | | |
| ΔAIC = 0, $w > 0.999$, $R^2 = 0.589$ | | | | |
| Area* | 0.14 | 105.8 | 1 | < 0.001 |
| Isolation* | –0.04 | 5.7 | 1 | 0.018 |
| Elevation* | 0.06 | 9.5 | 1 | 0.002 |
| Human population* | –0.05 | 17.9 | 1 | < 0.001 |
| Airport | –0.11 | 4.5 | 1 | 0.036 |
| SEVM eigenvectors | | | 3 | |
| Residuals | | | 203 | |

*b*, regression coefficients; ΔAIC, difference between a candidate model and the model with the lowest Akaike information criterion (AIC); *w*, AIC weight of the model; SEVM, spatial eigenvector mapping.
*This variable was log-transformed for the analysis.

If anthropogenic factors were not included in the regression model, the relationship between log(area) and log(species richness) was not linear. A piecewise regression with a break point (Fig. 3a) fitted the data significantly better than did a linear model ($F_{2,205} = 8.5$, $P = 0.0003$). The estimated position of the break point was area = 1.5 km$^2$. That is, the slope of the species–area relationship changes on islands larger than 1.5 km$^2$. Based on the AIC-evidence ratio, the model including a break point was 623 times more likely than the linear model. In the model including the break point, the $z$-value for islands with an area < 1.5 km$^2$ was 0.15 (CI = 0.11–0.18), while it was significantly smaller for larger islands ($z = 0.04$, CI = 0.01–0.08). The break point remained significant in a model that also included isolation and elevation ($F_{2,203} = 10.5$, $P < 0.0001$).

However, if anthropogenic factors were added to the model, the piecewise regression did not converge to any break point. This strongly suggests that large islands have reduced $z$-values because of the negative effects of human influence. It should be noted that the break point of 1.5 km$^2$ corresponds to an island size above which human impact rises abruptly. In our study, only 3% of islands smaller than 1.5 km$^2$ had permanent inhabitants, and only one had an airport. Conversely, 80% of islands larger than 1.5 km$^2$ were inhabited (Fig. 3b), and 26% had an airport.

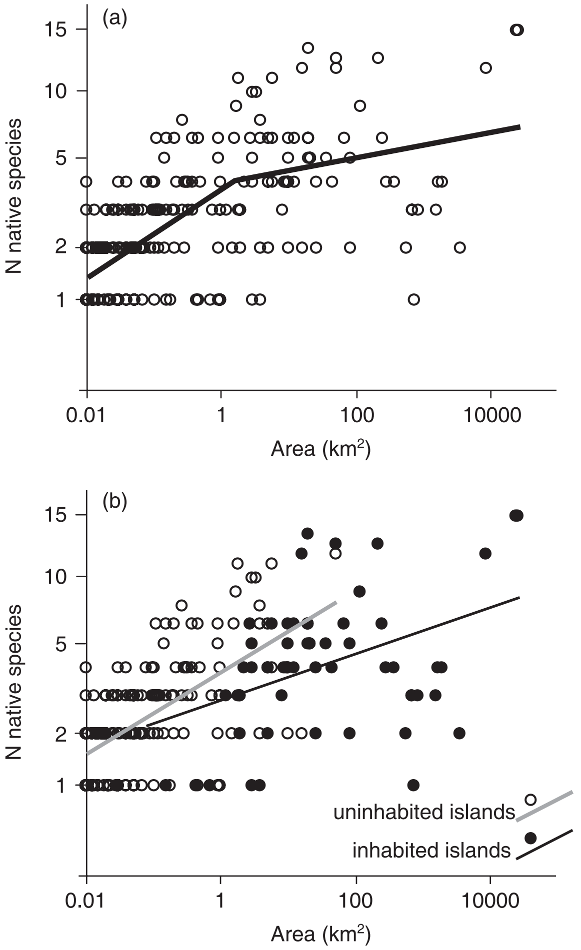


Figure 3 Relationship between island area and richness of native reptiles. (a) Piecewise linear regression, analysing all islands. (b) Linear regressions obtained analysing separately islands uninhabited (grey line, empty dots) and inhabited by humans (black line, filled dots). For uninhabited islands, $z = 0.14$; for inhabited islands, $z = 0.08$.

Table 2 Regression models relating richness of non-native reptiles to geographical and human factors.

| | $b$ | $F$ | d.f. | $P$ |
|---|---|---|---|---|
| Model: geographical-only | | | | |
| $\Delta$AIC = 14.66; $w = 0.0007$; $R^2 = 0.449$ | | | | |
| Area* | 0.07 | 46.6 | 1 | < 0.001 |
| Isolation* | 0.03 | 2.1 | 1 | 0.132 |
| Elevation* | 0.01 | 0.2 | 1 | 0.673 |
| SEVM eigenvector | | | 1 | |
| Residuals | | | 207 | |
| Model: human-only | | | | |
| $\Delta$AIC = 27.53, $w < 0.0001$, $R^2 = 0.401$ | | | | |
| Human population* | 0.04 | 21.5 | 1 | < 0.001 |
| Airport | 0.25 | 18.6 | 1 | < 0.001 |
| SEVM eigenvector | | | 1 | |
| Residuals | | | 208 | |
| Model: geographical + human | | | | |
| $\Delta$AIC = 0, $w = 0.999$, $R^2 = 0.496$ | | | | |
| Area* | 0.04 | 6.1 | 1 | 0.014 |
| Isolation* | 0.02 | 0.8 | 1 | 0.354 |
| Elevation* | 0.01 | 0.1 | 1 | 0.739 |
| Human population* | 0.02 | 1.4 | 1 | 0.223 |
| Airport | 0.19 | 11.5 | 1 | < 0.001 |
| SEVM eigenvector | | | 1 | |
| Residuals | | | 205 | |

*b*, regression coefficients; $\Delta$AIC, difference between a candidate model and the model with the lowest Akaike information criterion (AIC); *w*, AIC weight of the model; SEVM, spatial eigenvector mapping.
*This variable was log-transformed for the analysis.

Inhabited islands were larger than uninhabited (*t*-tests: $t_{210} = 5.7$, $P < 0.0001$); similarly those with airports were larger than those without ($t_{210} = 8.7$, $P < 0.0001$). Regression lines obtained analysing separately inhabited and uninhabited islands are shown in Fig. 3(b).

Path analyses showed that the model m1 (i.e. island area affects species richness directly *and* through its relationship with human impact) was the best path model (Fig. 4a). Island size had a direct positive effect on reptile richness. However, area also had a positive significant relationship with human population and the presence of airports, which in turn had a negative effect on richness. Therefore, area had a negative indirect effect on richness, mediated by its relationship with human impact (Fig. 4a). Model m1 had a significantly better fit than the less complex models m2 (Fig. 4b) and m3 (Fig. 4c) (m1–m2: likelihood ratio test: $\chi^2_1 = 5.489, P = 0.019$; m1–m3: $\chi^2_5 = 457.87, P < 0.0001$). Model m2 had a significantly better fit than model m3 ($\chi^2_4 = 452.38, P < 0.0001$). Path coefficients were consistent with m1 in more complex models as well, taking into account the effect of elevation and isolation.

## Alien species

The number of alien species per island ranged from 0 to 7. Most of the alien species were introduced in historical times from mainland Europe, northern Africa and other areas of the Mediterranean basin. Only a few alien species, such as the North American slider turtle, *Trachemys scripta*, originated from different continents.

The geographical + human model was the best, according to AIC values, and explained a substantial proportion of variance ($R^2 = 0.496$). In this model, the number of alien species is highest on larger islands and on islands with airports (Table 2). The evidence ratio indicates that support for the geographical-only and of the human-only models is very low, since they were more than 1500 times less likely than the geographical + human model. We did not observe significant interactions between human and geographical factors.

## Total number of species

Finally, to evaluate whether the increase in species richness caused by alien species compensates for the loss caused by

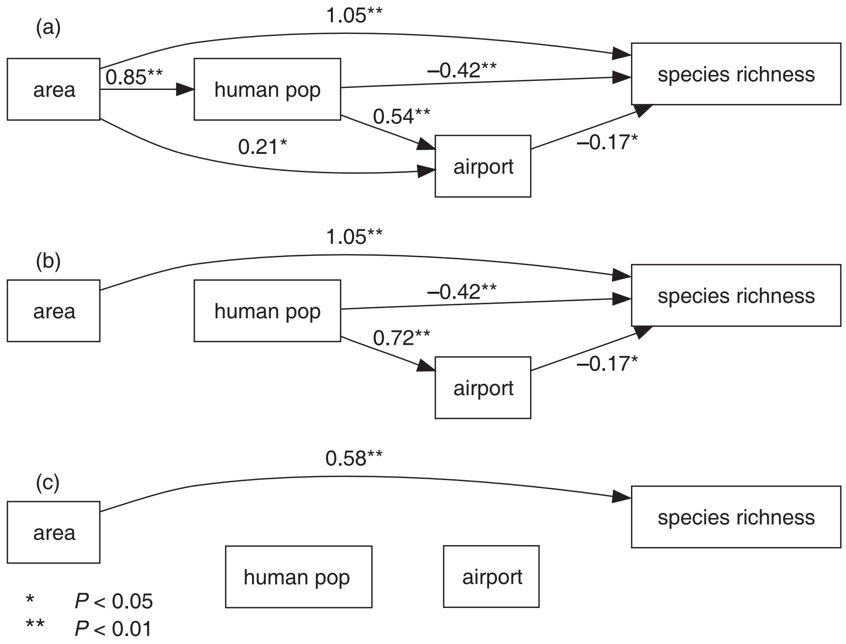


**Figure 4** Path diagrams, representing three models of the relationships between island area, human population, presence of airports and richness of native reptiles. (a) Full model, assuming that island size affects species richness directly and through its relationship with human impact (model m1 in the Methods). (b) Both island size and human activities affect species richness, but there is no co-action among them (model m2). (c) Island size directly affects species richness, while human activities do not have an effect (m3). The model in (a) performed significantly better than the alternative models (see Results). The numerical values are the path coefficients.

human presence (Sax *et al.*, 2002), we built a model analysing total species richness, including both native and alien species. The total number of species per island ranged between 1 and 18. Once again, the best model included both human and geographical factors (Table 3). The total number of species was higher on large, high-elevation islands, and was lower on islands with a dense human population. This strongly suggests that an increase in alien species does not compensate for a decrease in species richness caused by human influence. Isolation and airport presence did not have a significant effect in this model, probably because they influenced native and alien species in opposite directions (see Tables 1 & 2). When we tested for interactions, we found a significant interaction between distance from the continent and human population ($F_{1,202} = 4.8$, $P = 0.03$), similar to the one observed for native species.

Similar to the model analysing native species, if anthropogenic factors were not included in the model, the relationship between log(area) and log(species richness) was not linear. A relationship with a break point fitted the data significantly better than a linear model ($F_{2,206} = 4.7$, $P = 0.01$), and its shape was similar to that estimated for native species (small islands, $z = 0.21$, CI = 0.16–0.27; large islands, $z = 0.10$, CI = 0.06–0.15). The estimated position of the break point was very close to that estimated for the data set including only native species (1.03 km$^2$), but the break point was not present if anthropogenic factors were included in the model. This further confirms that alien species do not compensate for the decrease in species richness related to human influence.

**Table 3** Regression models relating the overall richness of reptiles (natives + non-natives) to geographical and human factors.

| | *b* | *F* | d.f. | *P* |
|---|---|---|---|---|
| Model: geographical-only<br>ΔAIC = 23.4; $w < 0.0001$; $R^2 = 0.580$ | | | | |
| Area* | 0.13 | 87.2 | 1 | < 0.001 |
| Isolation* | –0.03 | 1.2 | 1 | 0.282 |
| Elevation* | 0.08 | 9.0 | 1 | 0.003 |
| SEVM eigenvectors | | | 2 | |
| Residuals | | | 206 | |
| Model: human-only<br>ΔAIC = 102.4, $w < 0.0001$, $R^2 = 0.319$ | | | | |
| Human population* | 0.08 | 24.6 | 1 | < 0.001 |
| Airport | 0.08 | 0.8 | 1 | 0.372 |
| SEVM eigenvectors | | | 3 | |
| Residuals | | | 206 | |
| Model: geographical + human<br>ΔAIC = 0, $w > 0.999$, $R^2 = 0.589$ | | | | |
| Area* | 0.19 | 95.2 | 1 | < 0.001 |
| Isolation* | –0.04 | 2.8 | 1 | 0.095 |
| Elevation* | 0.11 | 15.2 | 1 | < 0.001 |
| Human population* | –0.06 | 12.1 | 1 | < 0.001 |
| Airport | –0.12 | 2.9 | 1 | 0.090 |
| SEVM eigenvectors | | | 3 | |
| Residuals | | | 203 | |

*b*, regression coefficients; ΔAIC, difference between a candidate model and the model with the lowest Akaike information criterion (AIC); *w*, AIC weight of the model; SEVM, spatial eigenvector mapping.
*This variable was log-transformed for the analysis.

## DISCUSSION

Our results show that human impact can strongly influence the pattern of island biogeography, and that it plays an important role in explaining the distribution of both native and alien species. Islands with high human impact host fewer native species than would be expected from a classical, linear biogeographical relationship, while they have an increased number of alien species. In other words, human impact seems to modify the biogeographical pattern, increasing the number of recent colonizations (Blackburn *et al.*, 2008) and altering the pattern of extinction; thus it can displace island biodiversity from the equilibrium points expected on the basis of geographical features alone (MacArthur & Wilson, 1967). Path analysis provides insights into the complex mechanisms driving the richness of native species. On the one hand, island area has a positive, direct effect on reptile richness. On the other hand, large islands suffer the strongest human impact, which in turn negatively affects species richness (Fig. 4a). Therefore, human impact mediates an indirect negative relationship between island area and the richness of native reptiles.

Other studies have used smaller data sets to show that islands hosting large human populations have more alien species, most probably because on these islands aliens have higher risk of introduction and/or a lower probability of extinction after establishment (Chown *et al.*, 1998; Sax *et al.*, 2002; Blackburn *et al.*, 2008). Our analysis confirms that human population and exchanges are major factors explaining the distribution of alien species. However, an important role of natural processes in the establishment of introduced species cannot be ruled out (Table 2). After accounting for human influence, the largest islands still have more alien species, suggesting that island area increases the possibility of establishment of propagules (Chown *et al.*, 1998; Sax *et al.*, 2002; but see also Blackburn *et al.*, 2008).

However, human influence was not limited to an increase in the number of alien species, since the biogeographical relationship of native species seems to have been modified: inhabited islands had fewer native species than expected, assuming the usual species–area relationship. This was particularly evident on large islands (Fig. 3), which host the largest human populations. Other studies have found that the establishment of introduced species compensated for recent extinctions on islands, with no net change or even an increase of the overall species richness per island (Sax *et al.*, 2002). Conversely, the richness of Mediterranean reptiles was negatively related to human influence, even when native and alien species were summed (Table 3). This may indicate that insular reptiles may be particularly prone to extinction and/or that a smaller number of species has been introduced compared with other groups.

Recent extinctions of reptiles on Mediterranean islands are poorly documented due to the lack of fossil remains and the paucity of studies (e.g. Capula *et al.*, 2002), and the pattern of extinction can often be estimated only from indirect evidence (e.g. Foufopoulos & Ives, 1999). Nevertheless, on most islands, the currently ongoing decline of several species is clearly related to human influence. For instance, 67% of European terrestrial reptiles that are critically endangered by extinction are island endemics (IUCN, 2007). Several species are nearly extinct on the largest, human-inhabited islands because of human-induced factors such as alien species, predatory pets and pests (such as cats and rats) and habitat loss, and survive only in the smallest, uninhabited satellite islands (Capula *et al.*, 2002; Cox *et al.*, 2006; Pérez-Mellado *et al.*, 2008). Indeed, in temperate regions humans tend to settle in areas with the highest biodiversity (Stohlgren *et al.*, 2006); therefore, it is very unlikely that the reduced richness of reptiles in human-inhabited islands arises because these islands naturally host a small number of native species. Despite the fact that we do not have direct data on historical extinctions, this reduced richness supports the idea that the smaller number of species on human-inhabited islands is probably due to a severe rise in the extinction rate.

An intense debate is ongoing in biogeography regarding the shape of species–area relationships. It has been proposed that natural ecological and geographical factors may affect the linearity of the relationships, yet evidence of deviation from a linear pattern (on log–log axes, i.e. a power law) is controversial (Lomolino, 2000, 2002; Williamson *et al.*, 2001). Some data sets suggest that the slope of the species–area relationship may decrease on large islands, and might even approach an asymptote (Lomolino, 2002; Kreft *et al.*, 2008). The strong relationship between human factors and species richness is probably widespread at the global scale, and we expect it to occur in other insular systems. Human influence may thus be an explanation of deviations from a simple linear pattern (e.g. Fig. 3). This hypothesis requires investigation across different groups and geographical areas, using data sets large enough to allow for testing of multiple hypotheses with enough statistical power.

Crutzen (2002) suggests that we are in a new geological era, the 'Anthropocene', in which humans are major drivers of global-scale processes, from atmospheric changes to biotic extinctions. Humans now rival natural geological and ecological processes in determining biodiversity at the global scale, and the present-day species distribution cannot be exhaustively analysed without considering human impact (Nogués-Bravo *et al.*, 2008). As macroecology becomes a vital tool in the study of global change (Kerr *et al.*, 2007), human influence should be integrated fully into biogeographical analyses for a more complete understanding of large-scale patterns and processes.

## ACKNOWLEDGEMENTS

We thank L. Bottoni, D. Fontaneto, S. Slavik, J. Lennon, C. Dormann and H. Kreft for comments on early versions of the manuscript. P. Perez-Neto helped with the development of SEVM models. G.F.F. was funded by a grant of EVK2-CNR on biodiversity conservation.

## REFERENCES

Anonymous (2003) *Etude des plans de gestion des sites MEDWET-COAST – site de Zembra*. Coastal Protection and Planning Agency, Ministry of Environment, Republic of Tunisia.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004) Avian extinction and mammalian introductions on oceanic islands. *Science*, **305**, 1955–1958.

Blackburn, T.M., Cassey, P. & Lockwood, J.L. (2008) The island biogeography of exotic bird species. *Global Ecology and Biogeography*, **17**, 246–251.

Blondel, J. & Aronson, J. (1999) *Biology and wildlife of the Mediterranean region*. Oxford University Press, New York.

Bollen, K.A. & Long, J.S. (1993) *Testing structural equation models*. Sage, Newbury Park, CA.

Bowerman, B.L. & O'Connell, R.T. (1990) *Linear statistical models*. PWS-Kent, Boston.

Bryman, A. & Cramer, D. (1990) *Quantitative data analysis for social scientists*. Routledge, London.

Buckley, L.B. & Jetz, W. (2007) Insularity and the determinants of lizard population density. *Ecology Letters*, **10**, 481–489.

Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag, New York.

Capula, M., Luiselli, L., Bologna, M.A. & Ceccarelli, A. (2002) The decline of the Aeolian wall lizard, *Podarcis raffonei*: causes and conservation proposals. *Oryx*, **36**, 66–72.

Castanet, J. & Guyetant, R. (1989) *Atlas de répartition des amphibiens et reptiles de France*. Société Herpétologique de France, Paris.

Chown, S.L., Gremmen, N.J.M. & Gaston, K.J. (1998) Ecological biogeography of southern ocean islands: species-area relationships, human impacts, and conservation. *The American Naturalist*, **152**, 562–575.

Cox, N., Chanson, J.S. & Stuart, S.N. (2006) *The status and distribution of reptiles and amphibians of the Mediterranean basin*. IUCN, Gland, Switzerland.

Crutzen, P.J. (2002) Geology of mankind. *Nature*, **415**, 23.

Daget, P. (1977) Le bioclimat méditerranéen: caractères généraux, modes de caractérisation. *Vegetatio*, **34**, 1–20.

Delaguerre, M. & Cheylan, M. (1992) *Atlas de répartition des batraciens et reptiles de Corse*. Parc Naturel Régional de Corse-Ecole Pratique de Hautes Etudes, Pampelune.

Diamond, J.M. (1969) Avifaunal equilibria and species turnover rates on the Channel Islands of California. *Proceedings of the National Academy of Sciences USA*, **64**, 57–63.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, **16**, 129–138.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Foufopoulos, J. & Ives, A.R. (1999) Reptile extinctions on land-bridge islands: life-history attributes and vulnerability to extinction. *The American Naturalist*, **153**, 1–25.

Griffith, D.A. & Peres-Neto, P.R. (2006) Spatial modelling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.

IUCN (2007) *2007 IUCN red list of threatened species*. http://www.iucnredlist.org/ (accessed 15 February 2008).

Kadmon, R. & Allouche, O. (2007) Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. *The American Naturalist*, **170**, 443–454.

Kalmar, A. & Currie, D.J. (2006) A global model for island biogeography. *Global Ecology and Biogeography*, **15**, 72–81.

Kerr, J.T., Kharouba, H.M. & Currie, D.J. (2007) The macroecological contribution to global change solutions. *Science*, **316**, 1581–1584.

Kreft, H., Jetz, W., Mutke, J., Kier, G. & Barthlott, W. (2008) Global diversity of island floras from a macroecological perspective. *Ecology Letters*, **11**, 116–127.

Lanza, B. & Poggesi, M. (1986) *Storia naturale delle isole satelliti della Corsica*. Istituto Geografico Militare, Firenze.

Lever, C. (2003) *Naturalized amphibians and reptiles of the world*. Oxford University Press, New York.

Lomolino, M.V. (2000) Ecology's most general, yet protean pattern: the species-area relationship. *Journal of Biogeography*, **27**, 17–26.

Lomolino, M.V. (2002) '... there are areas too small, and areas too large, to show clear diversity patterns ...' R.H. MacArthur (1972: 191). *Journal of Biogeography*, **29**, 555–557.

MacArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

Malkmus, R. (2004) *The amphibians and reptiles of Portugal, Madeira and the Azores archipelago*. Koeltz Scientific Books, Koenigstein.

Muggeo, V.M.R. (2003) Estimating regression models with unknown break-points. *Statistics in Medicine*, **22**, 3055–3071.

Nogués-Bravo, D., Araújo, M.B., Romdal, T. & Rahbek, C. (2008) Scale effects and human impact on the elevational species richness gradients. *Nature*, **453**, 216–219.

Pérez-Mellado, V., Hernández-Estévez, J.A., García-Díez, T., Terrassa, B., Ramón, M.M., Castro, J., Picornell, A., Martín-Vallejo, J. & Brown, R. (2008) Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain). *Amphibia-Reptilia*, **29**, 49–60.

Pleguezuelos, J.M., Márquez, R. & Lizana, M. (2002) *Atlas y libro rojo de los Anfibios y Reptiles de España*. Dirección General de Conservación de la Naturaleza-Asociación Herpetologica Española, Madrid.

R Development Core Team (2007) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna (http://www.r-project.org/).

Rangel, T.F.L.V.B., Diniz-Filho, J.A.F. & Bini, L.M. (2006) Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography*, **15**, 321–327.

Sala, O.E., Chapin, 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. & Wall, D.H. (2000) Biodiversity – global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.

Sax, D.F., Gaines, S.D. & Brown, J.H. (2002) Species invasions exceed extinctions on islands worldwide: a comparative study of plants and birds. *The American Naturalist*, **160**, 766–783.

Sindaco, R., Doria, G., Razzetti, E. & Bernini, F. (2006) *Atlas of Italian amphibians and reptiles*. Polistampa, Firenze.

Steadman, D.W. (1993) Biogeography of Tongan birds before and after human impact. *Proceedings of the National Academy of Sciences USA*, **90**, 818–822.

Stohlgren, T.J., Barnett, D.T., Flather, C., Fuller, P., Peterjohn, B., Kartesz, J.T. & Master, L.L. (2006) Species richness and pattern of invasion in plants, birds, and fishes in the United States. *Biological Invasions*, **8**, 427–447.

Sultana, J. & Falzon, V. (2002) *Wildlife of the Maltese islands*. BirdLife Malta, Malta.

Toms, J.D. & Lesperance, M.L. (2003) Piecewise regression: a tool for identifying ecological thresholds. *Ecology*, **84**, 2034–2041.

Whittaker, R.J. (1995) Disturbed island ecology. *Trends in Ecology and Evolution*, **10**, 421–425.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography*. Oxford University Press, Oxford.

Williamson, M., Gaston, K.J. & Lonsdale, W.M. (2001) The species–area relationship does not have an asymptote! *Journal of Biogeography*, **28**, 827–830.

Wilson, E.O. (2002) *The future of life*. Little Brown, London.

Wright, D.H. (1983) Species–energy theory: an extension of species–area theory. *Oikos*, **41**, 496–506.

## BIOSKETCHES

**G. Francesco Ficetola** is a post-doc. specializing in conservation and ecology of amphibians and reptiles. He combines multiple approaches (landscape ecology, macroecology, evolutionary ecology and population genetics) to obtain information that can be used to improve conservation actions. He is currently focusing on the issues of biological invasions.

**Emilio Padoa-Schioppa** is a researcher specializing in landscape ecology. He works on the management and conservation of landscapes and habitats in Europe, with a special interest in the Mediterranean herpetofauna.

Editor: Jack Lennon

*Journal of Biogeography* (*J. Biogeogr.*) (2013) **40**, 1202–1211

ORIGINAL ARTICLE

# Estimating patterns of reptile biodiversity in remote regions

Gentile Francesco Ficetola[1]*, Anna Bonardi[1], Roberto Sindaco[2] and Emilio Padoa-Schioppa[1]

*[1]Dipartimento di Scienze dell'Ambiente e del Territorio, Università degli Studi di Milano-Bicocca, Piazza della Scienza 1, 20126, Milano, Italy, [2]Museo Civico di Storia Naturale, I-10022, Carmagnola, Italy*

*Correspondence: Gentile Francesco Ficetola, Dipartimento di Scienze dell'Ambiente e del Territorio, Università degli Studi di Milano-Bicocca, Piazza della Scienza 1, 20126 Milano, Italy.
E-mail: francesco.ficetola@unimib.it

## ABSTRACT

**Aim** The incompleteness of information on biodiversity distribution is a major issue for ecology and conservation. Researchers have made many attempts to quantify the amount of biodiversity that still remains unknown. We evaluated whether models that integrate ecogeographical variables with measures of the effectiveness of sampling can be used to estimate biodiversity patterns (species richness) of reptiles in remote areas that have received limited surveys.

**Location** The Western Palaearctic (Europe, Northern Africa, the Middle East and Central Asia).

**Methods** We gathered data on the distribution of turtles, amphisbaenians and lizards. We used regression models integrating spatial autocorrelation (spatial eigenvector mapping and Bayesian autoregressive models) to analyse species richness, and identified relationships between species richness, ecogeographical features and large-scale measures of accessibility.

**Results** The two regression techniques were in agreement. Known species richness was dependent on ecogeographical factors, peaking in areas with high temperature and annual actual evapotraspiration, and intermediate cover of natural vegetation. However, richness declined sharply in the least accessible areas. Our models revealed regions where reptile richness is likely to be higher than currently known, particularly in the biodiversity hotspots in the south of the Arabian Peninsula, the Irano-Anatolian region, and the Central Asian mountains. An independent validation data set, with distribution data collected recently throughout the study region, confirmed that combining accessibility measures with ecogeographical variables allows a good estimate of reptile richness, even in remote areas that have received limited monitoring so far. Some remote regions that support very rich communities are covered very little by protected areas.

**Main conclusions** Integrating accessibility measures into species distribution models allows biologists to identify areas where current knowledge underestimates the actual richness of reptiles. Our study identifies regions requiring future biodiversity research, proposes a novel approach to biodiversity prediction in poorly studied areas, and identifies potential regions for conservation.

**Keywords**

**Amphisbaenians, biodiversity hotspots, conservation biogeography, ecogeography, lizards, predictive models, spatial autocorrelation, species richness, survey bias, turtles.**

## INTRODUCTION

Our knowledge of biodiversity distribution is far from complete (Mora *et al.*, 2011). First, many species on Earth are yet to be described (the Linnean shortfall), and second, we have a limited knowledge of the true distributions, even for the best-studied taxa (the Wallacean shortfall) (Lomolino, 2004; Cardoso *et al.*, 2011; Vale & Jenkins, 2012). This paucity of

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12060

information is a major challenge within conservation biogeography, limiting our ability to analyse biodiversity patterns and set conservation priorities (Lomolino, 2004; Cardoso *et al.*, 2011; Mora *et al.*, 2011). Unfortunately, quantifying the amount of biodiversity that remains unknown in a given area is challenging (Raxworthy *et al.*, 2003; Mora *et al.*, 2011).

Species distribution models based on relationships between biodiversity distribution and environmental features help to elucidate the factors determining biodiversity and to identify conservation targets even when information is incomplete (Raxworthy *et al.*, 2003; Qian & Ricklefs, 2008; Thuiller *et al.*, 2011). These models, being based on the available species distribution data, may be affected by poor sampling. Integrating data on the effectiveness of surveys may greatly improve the performance of models of species richness and species distribution (Kadmon *et al.*, 2004; Bini *et al.*, 2006; Phillips *et al.*, 2009; Barbosa *et al.*, 2010). Unfortunately, measures of the effectiveness of surveys are often lacking, particularly for large-scale data sets, where data are obtained from a variety of heterogeneous sources. Accessibility can be a major source of sampling bias. Monitoring in remote areas is often limited, and so biodiversity can be significantly underestimated (Nelson *et al.*, 1990; Margules & Pressey, 2000; Bini *et al.*, 2006; Sánchez-Fernández *et al.*, 2008; Barbosa *et al.*, 2010; Kent & Carmel, 2011). The accessibility of a region can thus be a useful proxy of the effectiveness of sampling in biogeographical studies. Recently, Nelson developed a global-scale measure of accessibility, quantified as the travel time from the nearest city using land-based or water-based transport (Nelson, 2008; Uchida & Nelson, 2010). However, we are not aware of studies integrating large-scale measures of accessibility into species distribution models.

The aim of this study was to analyse the richness of reptile species (considering turtles, amphisbaenians and lizards) in the Western Palaearctic, and to evaluate the importance of accessibility in determining known species richness at the biogeographical scale. The Western Palaearctic is a large biogeographical region, encompassing Europe, North Africa, the Middle East and Central Asia (Fig. 1); it contains hotspots of reptile biodiversity and endemism (Mittermeier *et al.*, 2004; Cox *et al.*, 2006; Sindaco & Jeremčenko, 2008), but is heavily threatened by growing human pressure (Brooks *et al.*, 2006; Cox *et al.*, 2006; Ficetola & Padoa-Schioppa, 2009). The area includes easily accessible, human-dominated regions (e.g. Western Europe), in which extensive inventories of reptile biodiversity have been performed, but also remote regions (Fig. 1a) where biodiversity data remain sparse (Sindaco & Jeremčenko, 2008). Furthermore, we evaluated whether integrating accessibility into species distribution models may help biologists to improve estimates of species richness, even in remote areas that have received limited surveys so far. In addition, reptiles are among the terrestrial vertebrates for which distribution data are more limited, and the only ones for which no comprehensive assessment of conservation status has been completed (Hoffmann *et al.*, 2010). The results of our study may improve knowledge of reptile biodiversity in the study area, and provide large-scale information that can be useful for conservation planning.

## MATERIALS AND METHODS

### Study area and data sets

We considered the 'Western Palaearctic' according to Sindaco & Jeremčenko (2008). This region includes the western portion of the Palaearctic realm [Europe, the Arabian Peninsula, western Asia (west of the Indus Valley), the former Soviet Republics of Central Asia and Russia (west of the Ural Mountains)] and several countries of northern Africa (including the Sahara Desert north of approximately 22° N) (Fig. 1). Our data set included more than 67,000 distributional records collected before 2008, showing the known distribution of the 480 native species of terrestrial turtles, amphisbaenians and lizards occurring in the area (Sindaco & Jeremčenko, 2008). The number of records per cell cannot be considered an index of the sampling effort, because in the best-surveyed areas (e.g. Europe), data were obtained from comprehensive sources such as atlases, summarizing a much larger number of observations. Presence records of each species were mapped on a grid of 3530 cells with a resolution of 1° × 1° (the resolution of data with less accurate coordinates). Taxonomy followed the checklist of Sindaco & Jeremčenko (2008), which was based on the critical review of about 880 papers. The list includes species recognized on the basis of biological, evolutionary and morphological species concepts (Uetz, 2010). Distribution maps of each of the 480 species used for analyses are available in Sindaco & Jeremčenko (2008). Reptile richness in each cell was then calculated by overlapping the distribution maps of all the 480 species. We also considered coastal cells and islands, as they harbour multiple endemic species and can constitute important biodiversity areas, when taking into account their limited surface area (Cox *et al.*, 2006; Ficetola & Padoa-Schioppa, 2009).

As environmental variables, we considered two geographical variables – cell surface occupied by non-aquatic environments (calculated on the basis of the GlobCover land cover; Bicheron *et al.*, 2008) and elevation range (calculated on the basis of a 30-arc-second digital elevation model; Hijmans *et al.*, 2005); and three climatic variables – annual mean temperature, annual actual evapotranspiration (annual AET hereafter), and annual summed precipitation (New *et al.*, 2002), which are considered to be major drivers of reptile biodiversity (Qian & Ricklefs, 2008; Powney *et al.*, 2010). Potential evapotranspiration can also have an important role (Rodríguez *et al.*, 2005), but was not considered because it is highly collinear to the other climatic variables. Furthermore, we considered the average accessibility of each cell (Nelson, 2008) and the percentage of each cell occupied by natural vegetation, calculated on the basis of GlobCover (Bicheron *et al.*, 2008). We also calculated the percentage of each cell covered by protected areas, on the basis of the World Database on Protected Areas (http://protectedplanet.net/) (Fig. 1b). All variables were upscaled

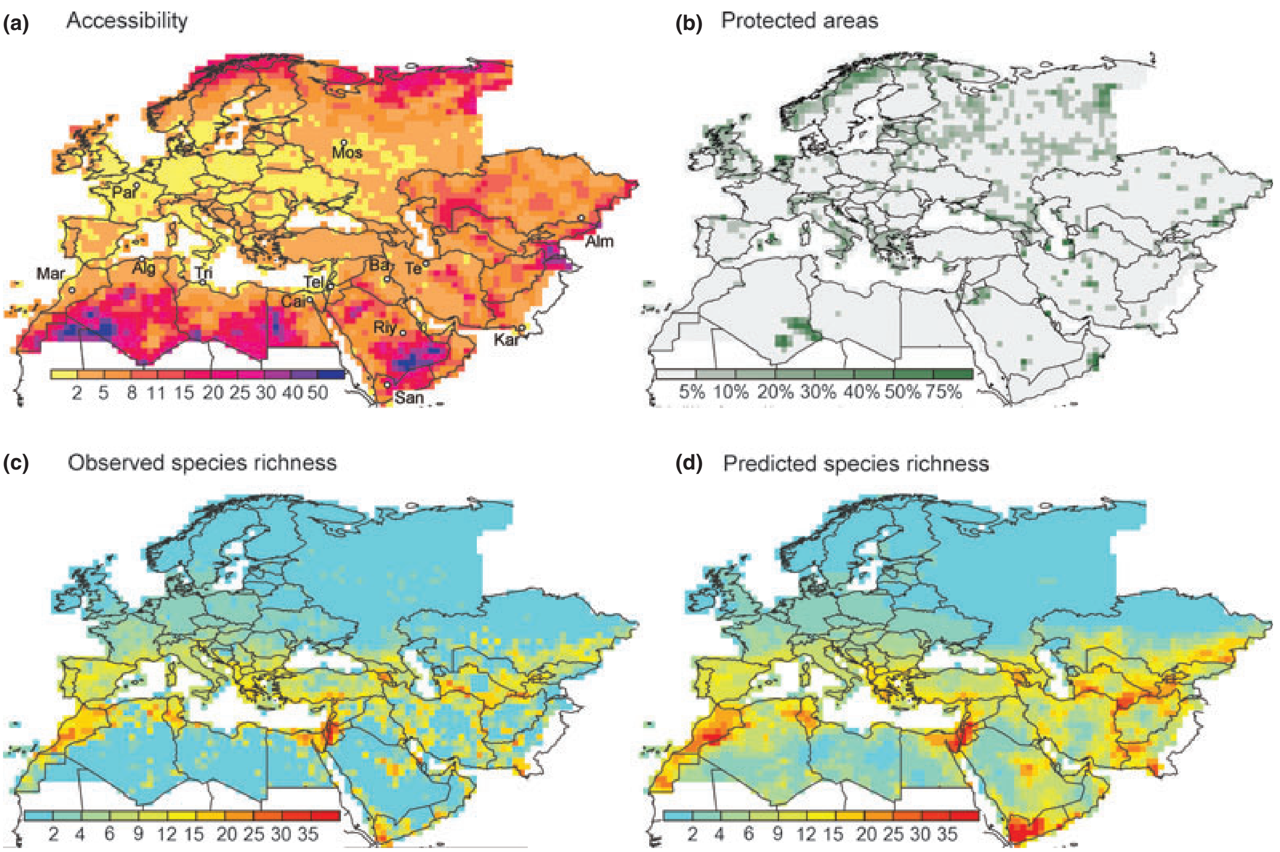


**Figure 1** Accessibility, protected areas and reptile richness within the Western Palaearctic. (a) Mean accessibility of 1° × 1° cells within the study area (calculated following Nelson, 2008). Accessibility is measured as the travel time (in hours) from the nearest city using land-based or water-based transport. (b) Percentage of each cell occupied by protected areas (World Database on Protected Areas; http://protectedplanet.net/). (c) Observed species richness of reptiles across the Western Palaearctic. (d) Reptile species richness predicted by models assuming that all cells are accessible in 1 h. The picture shows the average predictions of spatial eigenvector mapping and Bayesian conditional autoregressive model. Predictions of the two models are shown in Fig. S2 in Appendix S2. The locations of major cities are also shown: Alg, Algiers; Alm, Almaty; Ba, Baghdad; Cai, Cairo; Kar, Karachi; Mar, Marrakech; Mos, Moscow; Par, Paris; Riy, Riyadh; San, Sana'a; Te, Tehran; Tel, Tel Aviv; Tri, Tripoli.

(via averaging) from the original sources to match the 1° × 1° resolution. Accessibility represents the travel time to the nearest city (population $\geq$ 50,000 people) using land-based or water-based transport on a regular raster grid. It is based on a cost–distance algorithm computing the 'cost' of travelling between two locations, assuming that travelling time is affected by parameters such as road/railway network, presence of paths, navigable water bodies, slope and political boundaries. For each 1° × 1° cell, we obtained the average accessibility (in hours), from the 30″-resolution accessibility grid layer (Nelson, 2008; Fig. 1a). A few cells (3%) had average accessibility $\leq$ 1 h. To avoid these cells acting as outliers, we assumed that all areas within 1 h of the nearest city may be surveyed with similarly high sampling effort, and were assigned an accessibility of 1 (i.e. 1 h is the minimum accessibility across the whole study area). The appropriateness of this approach is confirmed by the observation that known richness follows a nonlinear pattern, with maximum richness at accessibility $\approx$ 1 h (Fig. 2). The study area includes several remote regions in the Arabian Peninsula, the Sahara Desert, in several areas of the Middle East and Central Asia, and in northern Russia (the latter almost devoid of reptile fauna). Although there are remote areas both in cold and warm regions (Fig. 1a), the most accessible areas tended to have higher annual AET [Dutilleul's (1993) correlation index: $r = -0.69$, $n = 3530$, modified d.f. = 22.4, $P < 0.001$] and more precipitation ($r = -0.59$, $n = 3530$, modified d.f. = 24.3, $P = 0.003$). The correlation with the other environmental variables was not significant [Dutilleul's (1993) correlation for all pairs: $n = 3530$, $P > 0.06$].

## Species richness models

We used regression models to evaluate the relationships between environmental features and species richness. The residuals of ordinary least squares regression were affected by significant spatial autocorrelation (Moran's $I = 0.40$, $P < 0.001$, $n = 207{,}214$), and so we integrated autocorrelation into our models. Because results may differ across statistical techniques, we used two different approaches that are

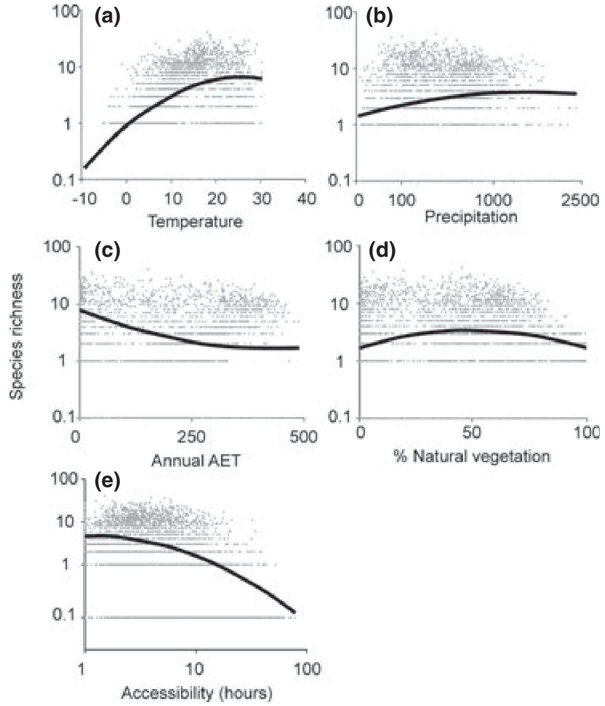


**Figure 2** Relationships between environmental variables and reptile richness for the Western Palaearctic, as estimated by spatial eigenvector mapping. Predictions are shown for (a) mean annual temperature; (b) summed annual precipitation; (c) annual actual evapotranspiration (annual AET); (d) proportion of each cell occupied by natural vegetation (%); (e) mean accessibility (travel time in hours from the nearest city using land-based or water-based transport). The plots include variables for which the model takes quadratic terms into account. Grey dots represent observed species richness in cells. The predictions for each variable are made while holding the other variables fixed at their global median. Fig. S2 in Appendix S2 reports the results from the Bayesian autoregressive model and the global medians of variables.

among the best-performing techniques when using non-normal data: spatial eigenvector mapping (SEVM) and Bayesian intrinsic conditional autoregressive models (BCA) (Dormann *et al.*, 2007; Beale *et al.*, 2010). Our dependent variable was a count, and so we assumed a Poisson error distribution rather than more standard techniques such as generalized least squares, simultaneous or conditional autoregressive models, which are suitable for normally distributed data (Dormann *et al.*, 2007; Beale *et al.*, 2010). Species richness may have a maximum when certain environmental variables reach certain values; furthermore, previous studies suggested that relationships between species richness and ecogeographical variables are often nonlinear (Whittaker *et al.*, 2007; Qian, 2010). We therefore also included quadratic terms for climatic variables, vegetation and accessibility in all models, to take into account potential nonlinear relationships.

SEVM allows the translation of the spatial arrangement of data points into explanatory variables (eigenvectors) that capture spatial effects (Dormann *et al.*, 2007). Neighbouring cells were linked using a king's connection scheme and we generated eigenvectors using Moran's eigenvector maps (Dray *et al.*, 2006; Borcard *et al.*, 2011). We then selected the eigenvectors that best reduced the spatial autocorrelation of residuals, and included them as spatial predictors in generalized linear models (GLM) (Griffith & Peres-Neto, 2006; Dormann *et al.*, 2007). Comparisons among statistical methods showed that this implementation of SEVM is flexible and efficient, even when analysing non-normal data (Dormann *et al.*, 2007). To take overdispersion into account, we used a quasi-Poisson family and tested significance using an *F* test (Crawley, 2007). We used the procedure proposed by Phillips and co-workers (Phillips, 2010; Elith *et al.*, 2011) to evaluate the relative importance of predictors in SEVM. First, we built the models including each environmental variable in isolation. Subsequently, each variable (including both linear and quadratic terms) was removed in turn, and the model run with the remaining variables. The difference in explained variation between the full model and the one excluding the variable of interest provides an estimate of the independent contribution of such variable, after taking into account other parameters and spatial autocorrelation (Phillips, 2010; Elith *et al.*, 2011).

BCA is a hierarchical Bayesian technique that incorporates a spatial random effect, describing the neighbourhood of each site, into Bayesian regression models. Simulations have shown that BCA is one of the spatial techniques with the best performance under a variety of conditions (Latimer *et al.*, 2006; Beale *et al.*, 2010). We built BCA assuming a Poisson error distribution. A model assuming negative binomial error had a higher deviance information criterion, indicating that overdispersion did not affect the BCA model. To ensure convergence, we ran three different MCMC chains. For each chain, we performed a burn-in of 15,000 iterations, followed by 45,000 iterations thinned every 10 iterations. For all variables, Brooks–Gelman–Rubin diagnostics were approximately 1, indicating convergence (Kéry, 2010). We then calculated the estimates of regression coefficients and their 95% highest posterior density credible intervals (HPDI).

The results of ecogeographical studies may be affected by the scale at which analyses are performed (Hawkins, 2012). Furthermore, the large number of cells might increase the chance of finding significant results. To confirm our results, we therefore repeated our analyses using a grid with a resolution of 2° × 2° (975 cells).

To reduce skewness and improve normality where needed, variables were transformed using logarithms (cell surface, elevational range and accessibility), square-root (summed annual precipitation) or square-root-arcsine (percentage of natural vegetation). We calculated the proportion of explained deviance ($R^2_D$) as a measure of the variation explained by species richness models. For SEVM, values of $R^2_D$ were adjusted to take into account the number of predictors in the model (Borcard *et al.*, 2011). When more than one modelling technique is used, differences among algorithms can determine differences in performance and lead to non-identical results (model-based uncertainty) (Marmion *et al.*, 2009). To reduce the model-based uncertainty for model predictions, we kept the average consensus prediction of the two species richness models (Marmion *et al.*, 2009). Using the individual results of either model yielded nearly identical results (see Results). Subsequently, we projected models using a new data set, with identical environmental variables and spatial structure, but assuming accessibility is 1 h for all cells. This allowed us to evaluate the potential species richness of each cell, if it received a surveying effort similar to the most accessible ones ('projected richness' hereafter).

### Validation of model predictions

To test whether the projected richness is an accurate estimate of the actual species richness, we obtained recent data on reptile distributions (from recent field research or publications) in 80 grid cells that have received intense sampling (Fig. 3, and see Appendix S1 in Supporting Information). These recent data were not included in the 2008 data set used to build the species richness models. Independent validation data were obtained mostly from recent publications on reptile distributions (73 cells). In six cells, recent data were obtained both from extensive field surveys performed by the authors and from the literature. Sampling was performed from 2008 to 2010 through diurnal and nocturnal visual encounter surveys, trying to cover the widest range of environmental features (Heyer *et al.*, 1994; Ficetola *et al.*, 2010). In one additional cell, data were obtained through field surveys only (total: 80 cells with recent data) (Appendix S1). As these cells recently received an intense and focused surveying effort, they represent areas with reliable information for the validation of model predictions. We thus assumed that reptile richness recorded from recent, intensive surveys ('validation richness' hereafter) represents an improved estimate of the actual species richness in these cells. Validation cells were in multiple regions of the study area, with a wide range of accessibility (Fig. 3, Appendix S1).

In order to assess the ability of the projected richness model to correctly predict the actual species richness in the validation cells, we built three Poisson GLMs considering validation richness as the dependent variable in all cases. As potential predictors, we included (1) the projected richness (i.e. the richness predicted by the species richness model assuming accessibility is 1 h in all cells), and (2) the number of species recorded in the 2008 data set (i.e. the original data set of Sindaco & Jeremčenko, 2008). In both these models, we assumed $B = 1$ and intercept = 0, i.e. a perfect overlap between best available value of the species richness and either (1) our projected richness or (2) the richness recorded until 2008. We considered the null model as a third GLM (3), assuming no relationships between the validation and projected species richness. We then used the Akaike information criterion (AIC) to compare the performance of the three GLMs.

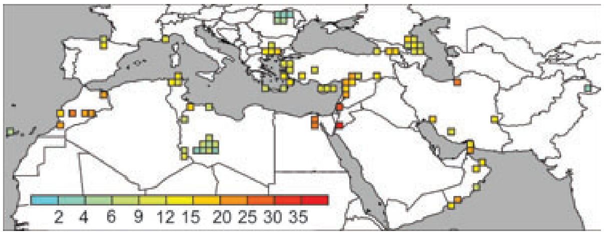


**Figure 3** Distribution of the 80 validation grid cells, and observed reptile richness in the validation cells.

Due to sampling constraints, remote cells were less frequent in our validation data. Furthermore, several validation cells were clustered, and this may affect the results of regression-based validation. As an additional metric of prediction performance, we compared the projected richness to the validation richness. For each cell, we calculated the absolute value of [(projection − validation)/validation] × 100%. We then subtracted that value from 100%, to obtain the 'percentage accuracy' for each cell. Subsequently, we repeated the same procedure with the richness recorded until 2008.

We used Spearman's correlation to evaluate the relationship between projected reptile richness in each cell, and the cell percentage covered by protected areas. Because of spatial autocorrelation, significance was calculated using Dutilleul's (1993) degrees of freedom. We performed statistical analyses in R 2.12 (R Development Core Team, 2010) using the package SPDEP (Bivand *et al.*, 2010); we ran BCA in WinBUGS 1.4 (Spiegelhalter *et al.*, 2008).

## RESULTS

Average reptile richness in 1° × 1° cells (± SD) was 4.3 ± 5.6 species (range: 0–41). The areas with the highest known richness were in the eastern Mediterranean and western Maghreb. Richness was irregularly distributed in wide areas of the Middle East, with species-rich cells near cells with very few known species (Fig. 1c). This pattern was most apparent in the area of Israel–Jordan–Sinai Peninsula, southwestern Yemen, and around Riyadh, Saudi Arabia (Fig. 1c). SEVM and BCA generally gave the same outcome, and their results were in strong agreement ($R^2 = 0.82$; Fig. S1 in Appendix S2). Both models included 12 predictors (seven variables with five quadratic terms); SEVM also included 91 eigenvectors representing spatial autocorrelation (Table 1). Both models explained a large amount of the variation in reptile species richness (SEVM: adjusted $R^2_{\mathrm{ADJ}} = 0.69$; BCA: $R^2_{\mathrm{D}} = 0.63$). When controlling for the effect of cell surface, species richness increased with elevational range, temperature and precipitation, and decreased with annual AET. Furthermore, richness was highest in cells with intermediate values of natural vegetation cover (Table 1, Fig. 2; see also Fig. S2 in Appendix S2). Temperature, natural vegetation, elevational range and accessibility explained 20–32% of variation, when used in isolation (Table 1). After taking into account spatial autocorrelation, accessibility was the variable with the highest independent contribution to the model. The independent contribution of vegetation cover and temperature was 2–4% while, when taking the other variables into account, the independent contribution of surface, elevational range and precipitation was limited (Table 1).

**Table 1** Results of multiple regression models using spatial eigenvector mapping (SEVM) and Bayesian conditional autoregressive models (BCA), relating reptile species richness to multiple environmental variables for the Western Palaearctic.

| | SEVM | | | | BCA | | | % Importance | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | *B* | *F* | d.f. | *P* | Median *B* | 95% HPDI | | with only | excluding |
| Surface | **0.128** | 42.9 | 1, 3426 | < 0.001 | **0.176** | 0.116 | 0.233 | 1.5% | 0.4% |
| Elevational range | **0.133** | 65.0 | 1, 3426 | < 0.001 | **0.154** | 0.101 | 0.205 | 19.6% | 0.6% |
| Temperature | **0.157** | 210.7 | 1, 3426 | < 0.001 | **0.096** | 0.058 | 0.138 | 31.9% | 4.1% |
| $(\text{Temperature})^2$ | **−0.003** | 67.2 | 1, 3426 | < 0.001 | **−0.002** | −0.003 | −0.001 | | |
| Precipitation | **0.051** | 14.6 | 1, 3426 | < 0.001 | **0.052** | 0.010 | 0.100 | 1.6% | 0.2% |
| $(\text{Precipitation})^2$ | **−0.001** | 6.8 | 1, 3426 | 0.009 | −0.001 | −0.002 | 0.0002 | | |
| Annual AET | **−0.007** | 84.9 | 1, 3426 | < 0.001 | **−0.004** | −0.006 | −0.002 | 4.5% | 1.1% |
| $(\text{AET})^2$ | **0.00001** | 42.8 | 1, 3426 | < 0.001 | 0.000003 | −0.000001 | 0.000007 | | |
| Natural vegetation | **1.767** | 107.0 | 1, 3426 | < 0.001 | **1.516** | 1.034 | 1.947 | 20.7% | 2.1% |
| $(\text{Nat. vegetation})^2$ | **−1.140** | 121.4 | 1, 3426 | < 0.001 | **−0.952** | −1.219 | −0.661 | | |
| Accessibility | 0.083 | 1.3 | 1, 3426 | 0.254 | 0.185 | −0.004 | 0.368 | 14.7% | 5.3% |
| $(\text{Accessibility})^2$ | **−0.236** | 108.4 | 1, 3426 | < 0.001 | **−0.280** | −0.338 | −0.220 | | |
| SEVM eigenv. | | 20.3 | 91, 3426 | < 0.001 | | | | 37.3% | 17.9% |

Significant coefficients (SEVM) and coefficients with 95% highest posterior density credible intervals (HPDI) not overlapping zero (BCA) are in bold. %Importance: importance of each variable estimated for SEVM. In the 'with only' column, the importance (adjusted $R^2$) is assessed by building models including each variable in isolation; in the 'excluding' column, importance refers to the difference between the $R^2_{\text{ADJ}}$ of the full model and the $R^2_{\text{ADJ}}$ of the partial model excluding the variable of interest. Variables are: cell surface occupied by non-aquatic environments, elevational range, annual mean temperature, annual summed precipitation, annual actual evapotranspiration (annual AET), cell percentage occupied by natural vegetation, average accessibility.

When taking the other variables into account, the highest richness was observed in easily accessible cells, and gradually decreased in more remote areas, following a nonlinear pattern. For instance, if environmental features were held constant, reptile richness was 4–5 species in cells with an average accessibility of 1 h, compared with 1–2 species in cells with an accessibility of 10 h (Fig. 2e, Fig. S2e in Appendix S2). Results were extremely similar according to SEVM and BCA, although the relationship between two climatic features (annual AET and annual precipitation) were significantly nonlinear in SEVM but not in BCA (Table 1). The analysis performed at 2° × 2° resolution yielded very similar results (Appendix S3). For accessibility, regression coefficients were slightly different from the model at 1° × 1° resolution, but the negative, nonlinear relationship was confirmed (Appendix S3).

We then projected species richness models, assuming that all cells are accessible in 1 h, while keeping the original values of environmental variables. This allowed us to estimate the potential richness patterns, if remote areas received the same surveying effort as the accessible ones. In several regions, the projected consensus model predicted a higher species richness than the currently known values, particularly in the western Maghreb, in the south of the Arabian peninsula, in the Irano-Anatolian region, and in the Central Asian mountains (Fig. 1d, Fig. S3 in Appendix S2). Cells with high projected richness tended to be those with less cover of protected areas (Fig. 1; correlation between projected richness and cover of protected areas: $r_s = -0.25$, Dutilleul's modified d.f. = 96.9, $P = 0.015$). Furthermore, in the Western Palaearctic, the cover of protected areas was highest in easily accessible cells ($r_s = -0.24$, d.f. = 145.3, $P = 0.005$).

### Validation of biodiversity predictions

The recent distribution data updated the available information, and usually increased known richness in validation cells. The average number of newly recorded species per cell was 2.5 (range: 0–12; Appendix S1). The projected model predicted reptile richness in the recently surveyed cells well (GLM: $\chi^2_1 = 195.6$, $P < 0.001$; $R^2_D = 0.69$; Fig. 4). It allowed us to estimate the validation species richness better than using the number of species reported in the 2008 data set ($\Delta$AIC = 73.0) or using the null model ($\Delta$AIC = 193.6) as predictors. This indicates very strong support for the predictions of the projected model. Results remain the same if SEVM or BCA are used rather than the consensus model (not shown). Nevertheless, statistical tests should be considered with caution, as validation cells were not independent and only a few of them were in remote regions, owing to the limited recent surveys in those areas.

The average accuracy of predictions in validation cells ($n = 80$) was 79%, and was higher than the accuracy calculated using 2008 data (77%). The good performance of predictions was particularly evident in remote cells (cells with accessibility > 3 h: $n = 39$ accuracy of prediction = 74%, accuracy of 2008 data = 70%; cells with accessibility > 6 h: $n = 7$, accuracy of prediction = 80%, accuracy of 2008 data = 55%).

## DISCUSSION

In remote areas, we only know part of the species richness (Nelson *et al.*, 1990), but quantifying the magnitude of richness underestimation is extremely difficult. Our analyses suggest that, after taking key ecogeographical variables into

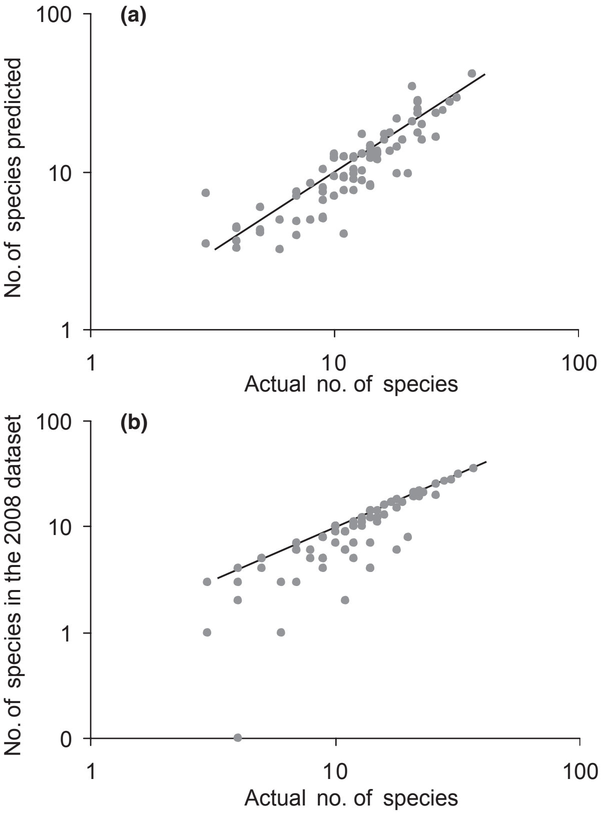


**Figure 4** Relationship between 'validation richness' (species richness of reptiles of the Western Palaearctic, obtained from recently available information) and: (a) the number of species predicted in the validation cells, assuming accessibility from the nearest city by land and water transport in 1 h; (b) the number of species in the 2008 data set. Lines have an intercept of 0 and a slope of 1.

account, accessibility influences the known reptile biodiversity at the biogeographical scale. Thus, integrating accessibility into models helps to estimate reptile richness in remote, poorly studied areas (Fig. 4a). Obviously, accessibility is not a perfect proxy of sampling effort because, particularly at fine scales, other factors may influence the completeness of inventories, such as perceived attractiveness for researchers and the presence of institutions specializing in certain taxa or specific geographical areas (Nelson *et al.*, 1990; Sánchez-Fernández *et al.*, 2008; Boakes *et al.*, 2010; Vale & Jenkins, 2012). Nonetheless, broad-scale measures of accessibility can provide information on the bias of field surveys, which may be used to improve the performance of distribution models and allows first assessments of the biodiversity in remote areas where observations are sparse (Bini *et al.*, 2006).

Our analysis suggests that reptile richness in remote areas such as the Iranian Plateau (including Afghanistan and south-western Pakistan), as well as parts of Central Asia, may be higher than currently known (Fig. 1, Fig. S3 in Appendix S2). Although this region includes two of the world's recognized biodiversity hotspots (the mountains of Central Asia and the Irano-Anatolian region), known reptile richness is still relatively low compared with other hotspots (Mittermeier *et al.*, 2004). The remoteness of the region, the roughness of landscape and the lasting political instability may have limited systematic surveys, and the actual biodiversity is likely to be higher than currently known. Two major mechanisms may cause such underestimation of species richness (Lomolino, 2004; Mokany & Ferrier, 2011): (1) species distribution is poorly known (Wallacean shortfall), and (2) some species are yet to be described (Linnean shortfall). Recent surveys provide a measure of causes of richness underestimation: on average, surveys increased the known richness by 47% and, in 13 cells, the increase in richness was ≥ 100% (Appendix S1). Limited knowledge of species distribution was the major cause of richness underestimation, as many species are recorded for only one or very few localities, and most new records in the validation cells expanded the range of known species. The discovery of new species is much less frequent than distribution updates. Nevertheless, most of the recently described reptile species in the Western Palaearctic (see e.g. Masroor, 2008; Rösler *et al.*, 2008; Nazarov *et al.*, 2009; Torki, 2010; Busais & Joger, 2011; Torki *et al.*, 2011) have been found in areas where our model suggests a higher richness than currently reported, such as the south of the Arabian Peninsula and the Irano-Anatolian region. Finally, in recent years, molecular tools are improving our definition of species identity, with the recognition of cryptic taxa and the raising of geographically isolated 'subspecies' to the rank of species. These studies tend to have a geopolitical bias (Harris & Froufe, 2005), but molecular tools and integrated taxonomy are nevertheless starting to expand our knowledge of biodiversity, even for remote or understudied regions (e.g. Barata *et al.*, 2012). Overall, we conclude that both Wallacean and Linnean shortfalls affect known reptile richness and our models help to identify regions where these shortfalls are particularly strong.

Reptiles are the only terrestrial vertebrates for which, so far, no comprehensive maps of species distribution range are available at the global scale, thereby limiting large-scale analyses of biodiversity patterns. For instance, this is the first analysis of reptile richness covering the whole Western Palaearctic on the basis of a regular grid (see Rodríguez *et al.*, 2005, for an analysis of Europe). Previous large-scale studies of reptile richness highlighted that the availability of energy is the major determinant of reptile richness (Rodríguez *et al.*, 2005; Qian *et al.*, 2007; Whittaker *et al.*, 2007; Qian & Ricklefs, 2008; Powney *et al.*, 2010; Qian, 2010). Our results were in agreement with these findings, as reptile richness was highest in areas characterized by warm climate and low annual AET (Fig. 2). Integrating proxies of sampling effort, such as accessibility, may improve the performance of biodiversity models, allowing better estimates of the relationships between species richness and environmental features (Phillips *et al.*, 2009; Bonardi *et al.*, 2011; Gómez-Rodríguez *et al.*, 2012), and helping to improve our understanding of the impact of environmental changes (e.g. climate or land-use change) on species richness. This can be particularly valuable for regions or taxa where insufficient data are available for modelling each species individually (Mokany & Ferrier, 2011). More accurate estimates of alpha diversity can also be integrated into a metacommunity framework, to estimate community composition in poorly known areas (Mokany *et al.*, 2011). Nevertheless, it should be noted that no single environmental variable explained richness well, and that spatial autocorrelation accounted for a substantial amount of the variation (Table 1), suggesting an important role of historical or other unmeasured factors.

Systematic conservation planning requires more complete knowledge of biodiversity distribution and the identification of relationships between biodiversity and environmental features (Margules & Pressey, 2000; Rondinini *et al.*, 2006; Mokany & Ferrier, 2011). Efforts to fill this gap are a priority for conservation biogeography, and must be performed in all regions and environments. Unfortunately, large areas of the globe have received limited surveys so far. Our approach allows spatially explicit estimates of unknown species richness, and can therefore guide the allocation of funding for biodiversity surveying towards specific regions, in which actual biodiversity values may be much higher than currently known.

Conservation actions are usually based on known biodiversity, with protected areas targeted to regions with high species richness, endemism and threat (Hughes *et al.*, 2002; Brooks *et al.*, 2006; Rondinini *et al.*, 2006). This may bias the allocation of conservation efforts towards the best-known areas, while regions with rich biodiversity, but that are poorly studied, may remain under-represented (Bini *et al.*, 2006). In the Western Palaearctic, remote regions can host high biodiversity, but tend to receive limited protection (Fig. 1). Our study identifies areas where the true biodiversity is probably much higher than the current known estimates. An exhaustive assessment of the network of protected areas for the conservation of reptiles is beyond the aim of this study, and should be performed at a finer spatial scale, taking into account not only species richness, but also other parameters such as the degree of endemism, the irreplaceability of areas, the presence of endangered taxa, and threatening processes (Margules & Pressey, 2000; Bini *et al.*, 2006; Brooks *et al.*, 2006). Our improved estimates of species richness can be integrated with the available information on endemism and threat, to identify regions where new protected areas may be required, despite biodiversity data remaining scarce (Hughes *et al.*, 2002; Bini *et al.*, 2006; Rondinini *et al.*, 2006). Furthermore, integrating our broad-scale analysis with high-resolution data on habitat availability and changes in land cover may allow finer-scale information to be developed (Rondinini *et al.*, 2011), which could help to refine the identification of conservation priorities. Some remote regions with high reptile richness had a very limited cover of protected areas, or no protected areas at all. There is a notable scarcity of protected areas in the south-west of the Arabian Peninsula (particularly in Yemen), in the east of the Iranian Plateau and in the western Maghreb (Fig. 1b), even though these regions are recognized as global biodiversity hotspots, and have a very high proportion of endemism for both reptiles and other taxa, including several micro-endemic reptiles recorded only in one or very few localities (Mittermeier *et al.*, 2004; Sindaco & Jeremčenko, 2008).

## ACKNOWLEDGEMENTS

We thank R. Ambrosini for statistical suggestions. J. Veech, T. Garner, D. Fontaneto, A. Aswad, S. Meiri, A. Feldman, L. Signorile and three referees provided comments on early versions of the manuscript. G.F.F. was funded by a scholarship of Università degli Studi di Milano-Bicocca.

## REFERENCES

Barata, M., Perera, A., Martínez-Freiría, F. & Harris, D.J. (2012) Cryptic diversity within the Moroccan endemic day geckos *Quedenfeldtia* (Squamata: Gekkonidae): a multidisciplinary approach using genetic, morphological and ecological data. *Biological Journal of the Linnean Society*, **106**, 828–850.

Barbosa, A.M., Fontaneto, D., Marini, L. & Pautasso, M. (2010) Is the human population a large-scale indicator of the species richness of ground beetles? *Animal Conservation*, **13**, 432–441.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Achard, F., Herold, M., Ranera, F. & Olivier, A. (2008) *GLOBCOVER*. Products description and validation report. Medias France, Toulouse, France.

Bini, L.M., Diniz, J.A.F., Rangel, T.F.L.V.B., Bastos, R.P. & Pinto, M.P. (2006) Challenging Wallacean and Linnean shortfalls: knowledge gradients and conservation planning in a biodiversity hotspot. *Diversity and Distributions*, **12**, 475–482.

Bivand, R., Altman, M., Anselin, L. *et al.* (2010) *spdep: spatial dependence: weighting schemes, statistics and models*. R package version 0.5-24. Available at: http://cran.r-project.org/package=spdep.

Boakes, E.H., McGowan, P.J.K., Fuller, R.A., Ding, C.-Q., Clark, N.E., O'Connor, K. & Mace, G.M. (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology*, **8**, e1000385.

Bonardi, A., Manenti, R., Corbetta, A., Ferri, V., Fiacchini, D., Giovine, G., Macchi, S., Romanazzi, E., Soccini, C., Bottoni, L., Padoa-Schioppa, E. & Ficetola, G.F. (2011) Usefulness of volunteer data to measure the large scale decline of "common" toad populations. *Biological Conservation*, **144**, 2328–2334.

Borcard, D., Gillet, F. & Legendre, P. (2011) *Numerical ecology with R*. Springer, New York, NY.

Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A.B., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S.L. (2006) Global biodiversity conservation priorities. *Science*, **313**, 58–61.

Busais, S.M. & Joger, U. (2011) Three new species and one new subspecies of *Hemidactylus* Oken, 1817 from Yemen (Squamata, Gekkonidae). *Vertebrate Zoology*, **61**, 267–280.

Cardoso, P., Erwin, T.L., Borges, P.A.V. & New, T.R. (2011) The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, **144**, 2647–2655.

Cox, N., Chanson, J.S. & Stuart, S.N. (2006) *The status and distribution of reptiles and amphibians of the Mediterranean Basin*. IUCN, Gland, Switzerland and Cambridge, UK.

Crawley, M.J. (2007) *The R book*. Wiley, Chichester, UK.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dray, S., Legendre, L. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Dutilleul, P. (1993) Modifying the *t* test for assessing the correlation between two spatial processes. *Biometrics*, **49**, 305–314.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Ficetola, G.F. & Padoa-Schioppa, E. (2009) Human activities alter biogeographical patterns of reptiles on Mediterranean islands. *Global Ecology and Biogeography*, **18**, 214–222.

Ficetola, G.F., Crottini, A., Casiraghi, M. & Padoa-Schioppa, E. (2010) New data on amphibians and reptiles of the Northern Areas of Pakistan: distribution, genetic variability and conservation issues. *North-Western Journal of Zoology*, **6**, 1–12.

Gómez-Rodríguez, C., Bustamante, J., Díaz-Paniagua, C. & Guisan, A. (2012) Integrating detection probabilities in species distribution models of amphibians breeding in Mediterranean temporary ponds. *Diversity and Distributions*, **18**, 260–272.

Griffith, D.A. & Peres-Neto, P.R. (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.

Harris, D.J. & Froufe, E. (2005) Taxonomic inflation: species concept or historical geopolitical bias? *Trends in Ecology and Evolution*, **20**, 6–7.

Hawkins, B.A. (2012) Eight (and a half) deadly sins of spatial analysis. *Journal of Biogeography*, **39**, 1–9.

Heyer, W.R., Donnelly, M.A., McDiarmid, R.W., Hayek, L.-A.C. & Foster, M.S. (1994) *Measuring and monitoring biological diversity: standard methods for amphibians*. Smithsonian Institution Press, Washington, DC.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hoffmann, M., Hilton-Taylor, C., Angulo, A. *et al.* (2010) The impact of conservation on the status of the world's vertebrates. *Science*, **330**, 1503–1509.

Hughes, T.P., Bellwood, D.R. & Connolly, S.R. (2002) Biodiversity hotspots, centres of endemicity, and the conservation of coral reefs. *Ecology Letters*, **5**, 775–784.

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401–413.

Kent, R. & Carmel, Y. (2011) Presence-only versus presence–absence data in species composition determinant analyses. *Diversity and Distributions*, **17**, 474–479.

Kéry, M. (2010) *Introduction to WinBUGS for ecologists: a Bayesian approach to regression, ANOVA, mixed models, and related analyses*. Academic Press, Burlington, MA.

Latimer, A.M., Wu, S., Gelfand, A.E. & Silander, J.A. (2006) Building statistical models to analyze species distributions. *Ecological Applications*, **16**, 33–50.

Lomolino, M.V. (2004). Conservation biogeography. *Frontiers of biogeography: new directions in the geography of nature* (ed. by M.V. Lomolino and L.R. Heaney), pp. 293–296. Sinauer, Sunderland, MA.

Margules, C.R. & Pressey, R.L. (2000) Systematic conservation planning. *Nature*, **405**, 243–253.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15**, 59–69.

Masroor, R. (2008) A new species of *Cyrtopodion* (Sauria: Gekkonidae) from the northern areas of Pakistan. *Zootaxa*, **1857**, 33–43.

Mittermeier, R.A., Gil, P.R., Hoffman, M., Pilgrim, J., Brooks, T., Mittermeier, C.G., Lamoreux, J. & da Fonseca, G.A.B. (2004) *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions*. CEMEX, Mexico City.

Mokany, K. & Ferrier, S. (2011) Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity and Distributions*, **17**, 374–380.

Mokany, K., Harwood, T.D., Overton, J.M., Barker, G.M. & Ferrier, S. (2011) Combining $\alpha$- and $\beta$-diversity models to fill gaps in our knowledge of biodiversity. *Ecology Letters*, **14**, 1043–1051.

Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B. & Worm, B. (2011) How many species are there on Earth and in the ocean? *PLoS Biology*, **9**, e1001127.

Nazarov, R., Ananjeva, N. & Radjabizadeh, M. (2009) Two new species of angular-toed geckoes (Squamata: Gekkonidae) from South Iran. *Russian Journal of Herpetology*, **16**, 311–324.

Nelson, A. (2008) *Travel time to major cities: a global map of accessibility*. Office for Official Publications of the European Communities, Luxembourg.

Nelson, B.W., Ferreira, C.A.C., da Silva, M.F. & Kawasaki, M.L. (1990) Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, **345**, 714–716.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*, **21**, 1–25.

Phillips, S.J. (2010) A brief tutorial on Maxent. *Lessons in Conservation*, **3**, 107–135.

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

Powney, G.D., Grenyer, R., Orme, C.D.L., Owens, I.P.F. & Meiri, S. (2010) Hot, dry and different: Australian lizard richness is unlike that of mammals, amphibians and birds. *Global Ecology and Biogeography*, **19**, 386–396.

Qian, H. (2010) Environment–richness relationships for mammals, birds, reptiles, and amphibians at global and regional scales. *Ecological Research*, **25**, 629–637.

Qian, H. & Ricklefs, R.E. (2008) Global concordance in diversity patterns of vascular plants and terrestrial vertebrates. *Ecology Letters*, **11**, 547–553.

Qian, H., Wang, X., Wang, S. & Li, Y. (2007) Environmental determinants of amphibian and reptile species richness in China. *Ecography*, **30**, 471–482.

R Development Core Team (2010) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A. & Peterson, A.T. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, **426**, 837–841.

Rodríguez, M.Á., Belmontes, J.A. & Hawkins, B.A. (2005) Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica*, **28**, 65–70.

Rondinini, C., Wilson, K.A., Boitani, L., Grantham, H. & Possingham, H.P. (2006) Tradeoffs of different types of species occurrence data for use in systematic conservation planning. *Ecology Letters*, **9**, 1136–1145.

Rondinini, C., Di Marco, M., Chiozza, F., Santulli, G., Baisero, D., Visconti, P., Hoffman, M., Schipper, J., Stuart, S.N., Tognelli, M.F., Amori, G., Falcucci, A., Maiorano, L. & Boitani, L. (2011) Global habitat suitability models of terrestrial mammals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2633–2641.

Rösler, H., Köhler, J. & Böhme, W. (2008) A new species of the diurnal gekkonid genus *Pristurus* Rüppell, 1835 from the Red Sea island Hanish al-Kabir, Yemen. *Amphibia-Reptilia*, **29**, 217–227.

Sánchez-Fernández, D., Lobo, J.M., Abellán, P., Ribera, I. & Millán, A. (2008) Bias in freshwater biodiversity sampling: the case of Iberian water beetles. *Diversity and Distributions*, **14**, 754–762.

Sindaco, R. & Jeremčenko, V.K. (2008) *The reptiles of the Western Palearctic*. Edizioni Belvedere, Latina, Italy.

Spiegelhalter, D., Thomas, A., Best, N. & Lunn, D. (2008) *WinBUGS 1.4.3*. Imperial College, MRC, UK.

Thuiller, W., Lavergne, S., Roquet, C., Boulangeat, I., Lafourcade, B. & Araujo, M.B. (2011) Consequences of climate change on the tree of life in Europe. *Nature*, **470**, 531–534.

Torki, F. (2010) Beschreibung eines neuen *Asaccus* (Reptilia: Phyllodactylidae) aus der Provinz Lorestan, Iran. *Sauria*, **32**, 3–16.

Torki, F., Ahmadzadeh, F., Ilgaz, C., Avcı, A. & Kumlutas, Y. (2011) Description of four new *Asaccus* Dixon and Anderson, 1973 (Reptilia: Phyllodactylidae) from Iran and Turkey. *Amphibia-Reptilia*, **32**, 185–202.

Uchida, H. & Nelson, A. (2010) Agglomeration index: towards a new measure of urban concentration. *Urbanization and development: multidisciplinary perspectives* (ed. by J. Beall, B. Guha-Khasnobis and R. Kanbur), pp. 41–60. Oxford University Press, Oxford, UK.

Uetz, P. (2010) The original descriptions of reptiles. *Zootaxa*, **2334**, 59–68.

Vale, M.M. & Jenkins, C.N. (2012) Across-taxa incongruence in patterns of collecting bias. *Journal of Biogeography*, **39**, 1744–1748.

Whittaker, R.J., Nogués-Bravo, D. & Araújo, M.B. (2007) Geographical gradients of species richness: a test of the water-energy conjecture of Hawkins *et al.* (2003) using European data for five taxa. *Global Ecology and Biogeography*, **16**, 76–89.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Validation grid cells.

**Appendix S2** Model predictions (Figs S1–S2) and difference between projected and known richness (Fig. S3).

**Appendix S3** Results of regression models performed at the resolution of 2° × 2°.

## BIOSKETCHES

**G. Francesco Ficetola** is a post-doc specializing in conservation and ecology of amphibians and reptiles. He combines multiple approaches (landscape ecology, macroecology, evolutionary ecology and population genetics) to obtain information that can be used to improve conservation actions.

Author contributions: All authors jointly designed the study; R.S. gathered most of the data; G.F.F. and A.B. performed the analyses; G.F.F. wrote the first version of the paper; and all authors discussed the results and commented on the draft manuscript.

---

Editor: Joseph Veech

*Global Ecology and Biogeography, (Global Ecol. Biogeogr.)* (2014) **23**, 1303–1313

Global Ecology and Biogeography
A Journal of Macroecology

RESEARCH PAPER

# Sampling bias inverts ecogeographical relationships in island reptiles

Gentile Francesco Ficetola[1,2]*, Massimo Cagnetta[1], Emilio Padoa-Schioppa[1], Anita Quas[3], Edoardo Razzetti[4], Roberto Sindaco[5] and Anna Bonardi[1]

[1]*Dipartimento di Scienze dell'Ambiente e del Territorio, e di Scienze della Terra, Università degli Studi di Milano-Bicocca, Piazza della Scienza 1, 20126 Milano, Italy,* [2]*Laboratoire d'Ecologie Alpine (LECA), CNRS UMR 5553, Université Grenoble Alpes, BP 53, 38041 Grenoble Cedex 9, France,* [3]*Department of Economics, Finance and Control, EMLYON Business School, 23 Avenue Guy de Collongue, 69134 ECULLY Cedex, France,* [4]*Museo di Storia Naturale, Università degli Studi di Pavia, Piazza Botta 9, 27100 Pavia, Italy,* [5]*Museo Civico di Storia Naturale, 10022 Carmagnola, Italy*

*Correspondence: Gentile Francesco Ficetola, Laboratoire d'Ecologie Alpine (LECA), CNRS UMR 5553, Université Grenoble Alpes, BP 53, 38041 Grenoble Cedex 9, France.
E-mail: francesco.ficetola@gmail.com

## ABSTRACT

**Aim** Species richness is one of the commonest measures of biodiversity, and is a basis for analyses at multiple scales. Data quality may affect estimations of species richness, but most broad-scale studies do not take sampling biases into account. We analysed reptile richness on islands that have received different sampling efforts, and assessed how inventory completeness affects the results of ecogeographical analyses. We also used simulations to evaluate under what circumstances insufficient sampling can bias the outcome of biodiversity analyses.

**Location** Mediterranean islands.

**Methods** We gathered data on reptile richness from 974 islands, assuming better sampling in islands with specific inventories. We used Moran's eigenvector mapping to analyse the factors that determine whether an island has been surveyed, and to identify the relationships between reptile richness, geographical parameters and anthropic parameters. We simulated islands, mimicking patterns of true data, and sampled them with varying effort. Simulated richness was analysed using the same approach used for real-world data.

**Results** The probability that islands were sampled for reptiles was higher in large, human-populated islands. The relationship between human impact and reptile richness was negative in well-surveyed islands, but positive in islands that had not been systematically surveyed, because densely populated and accessible islands receive better sampling. In simulations, analyses successfully retrieved the relationships between species richness and human presence only if the average species detection probability was ≥75%. Poorer sampling resulted in biased regression results.

**Main conclusions** Human activities may strongly affect biodiversity, but human presence and accessibility improve sampling effort and thus the quality of biodiversity information. Therefore, regressing known species richness on parameters representing human presence may result in apparent positive relationships. These two facets of human presence (positive on biodiversity knowledge, negative on actual biodiversity) represent a major challenge for ecogeographical studies, as not taking them into account would bias analyses and underestimate human impact.

**Keywords**
**Accessibility, biodiversity surveys, detection probability, herpetological atlas, island biogeography, virtual ecologist, Wallacean shortfall, zero-inflated models.**

## INTRODUCTION

The analysis of species richness over broad spatial scales is a central topic in ecology and biogeography. These analyses can help to understand the environmental factors that drive biodiversity, to evaluate the ongoing effect of human activities, and to forecast the potential impact of future environmental changes (Kerr *et al.*, 2007). Analyses of species richness over broad

DOI: 10.1111/geb.12201
http://wileyonlinelibrary.com/journal/geb

geographical scales are increasingly popular due to technical advancements (e.g. geographical information systems, statistical tools suited for the analysis of spatially explicit data), but also due to the growing availability of environmental layers and distribution data, particularly in faunistic/floristic atlases and online databases (Yang *et al.*, 2013; Ficetola *et al.*, 2014). Our knowledge of species distribution remains, however, largely incomplete, even for the best-studied taxa (Wallacean shortfall) (Lomolino, 2004). Even if we assume that all the species detected in a given area are correctly identified and reported, the known species richness may be a severe underestimation of the true richness, because many species remain undetected. Analyses on plants and reptiles suggest that, in broad-scale databases, the underestimation of the true richness in a given area can exceed 30%, and underestimation is likely to be stronger in tropical regions and in less-studied, species-rich taxa (Ficetola *et al.*, 2013; Yang *et al.*, 2013).

Incomplete sampling can influence the outcome of analyses in multiple ways. In some cases, some species remain undetected at certain sites, but imperfect detection is randomly distributed across environmental patches. Under these circumstances, imperfect detection would moderately bias the results of analyses, for instance by reducing the absolute values of regression coefficients (Gu & Swihart, 2004), but random misdetection is not considered to have a major impact on the interpretation of results (Van Buskirk, 2005; but see also Foody, 2011). Unfortunately, it is unlikely that imperfect sampling would be random over broad spatial scales. If sampling is systematically biased toward areas with specific environmental features, the impact on ecogeographical analyses can be severe. For instance, regions with higher richness (e.g. strong elevation gradients) often attract more research interest (Sánchez-Fernández *et al.*, 2008; Vale & Jenkins, 2012; Yang *et al.*, 2013). More intensive sampling in mountainous regions is expected to overestimate the importance of topographic gradients on species richness, while neglecting other, relevant variables (Yang *et al.*, 2013).

Human presence can have contrasting effects on known biodiversity. On the one hand, human impact often has negative effects on many native species, reducing the richness of species that avoid human habitats. On the other hand, areas with high human presence and accessibility are most likely to receive accurate sampling and can therefore show more known species, regardless of the true species richness (Sánchez-Fernández *et al.*, 2008; Ficetola *et al.*, 2013). This sampling artefact may create misleading patterns of known biodiversity, determining underestimation of the human impact on biodiversity. In recent years, we have increasingly come to appreciate the importance of sampling accuracy on the analysis of species distribution and richness. A growing body of literature explored the effects of incomplete sampling, to identify new approaches for the analysis of biodiversity data (MacKenzie *et al.*, 2006; Miller *et al.*, 2011; Gómez-Rodríguez *et al.*, 2012). Most of these studies, however, were performed at rather fine spatial scales (e.g. regional or landscape), while less attention has been devoted to broad-scale analyses of data quality, and to the impacts on macroecological analyses (but see Sánchez-Fernández *et al.*, 2011; Foody, 2011; Kéry, 2011; Comte & Grenouillet, 2013; Yang *et al.*, 2013; Lahoz-Monfort *et al.*, 2014).

Islands are an excellent study system in which to assess the impact of data quality on species-richness analyses. Islands have attracted the attention of ecologists and biogeographers for more than 100 years, and many studies have identified the ecological, geographical and human factors that determine species richness (McArthur & Wilson, 1967; Whittaker & Fernández-Palacios, 2007). The availability of faunistic and floristic checklists for many islands has boosted biogeographical analyses, even over broad spatial scales (e.g. Kalmar & Currie, 2006; Kreft *et al.*, 2008; Triantis *et al.*, 2012). These sources often rely on a variety of surveys (Triantis *et al.*, 2012), but limited information is available on their completeness, and on the potential consequences of their quality for analyses.

The aim of this study was to evaluate the contrasting effects of human impact on species-richness data for islands. We focused on Mediterranean reptiles; reptiles are abundant vertebrates in small islands, and many data are available for this group in faunistic atlases and other publications, although the data quality is uneven. First, we identified the geographical and human factors related to the availability and quality of biodiversity data. Second, we evaluated whether different levels of sampling quality may affect the observed pattern of species richness. Specifically, we predicted a negative relationship between human impact and the true reptile richness on islands, but non-negative or positive relationships between richness and human impact if sampling is not evenly distributed, because islands that are easily accessible or with greater human presence are better sampled. Finally, we used the 'virtual ecologist' approach (Zurell *et al.*, 2010) to simulate islands suffering different human impacts. We evaluated how human impact may interact with sampling quality, and identified the conditions under which poor sampling may bias the estimation of the relationship between species richness and human impact.

## MATERIALS AND METHODS

### Study area and datasets

We analysed the islands of the Mediterranean basin and the adjacent Macaronesia (Fig. 1). We considered three geographical features that are known to affect reptile richness on islands – area, isolation and maximum elevation (McArthur & Wilson, 1967; Ficetola & Padoa-Schioppa, 2009) – and three variables that represent human impact – total human population, the presence of an airport, and naval connections. Airport presence is strongly related to the accessibility of islands for people, and is also a good indicator of economic development, trading exchanges and tourism (Green, 2007). Isolation was measured as the distance from the continent or from the nearest large island (Corsica, Sardinia, Sicily or Crete). Area, elevation, human population (in the period 2003–2007), naval connections and airport presence were mostly obtained from Arnold's (2008) database of Mediterranean islands, and from Ficetola & Padoa-Schioppa (2009). Additional data were obtained from

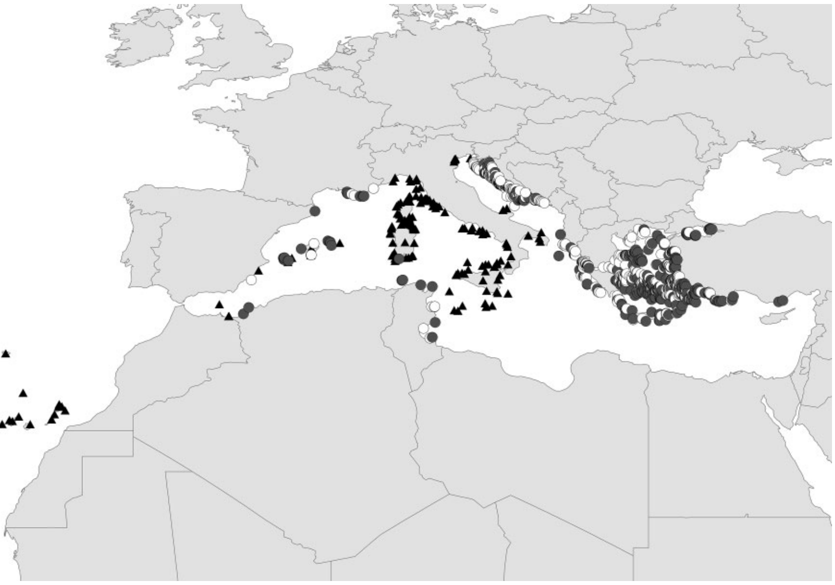


Figure 1 Study area and distribution of the 974 analysed Mediterranean and Macaronesian islands. Black triangles, islands in the 'atlas' dataset; open circles, islands in the 'no-atlas' dataset for which we did not find reptile records; grey circles, islands of the 'no-atlas' dataset with at least one reptile record.

national statistical databases, atlases and geographical publications. Naval connections were coded as follows: 0, no regularly scheduled connections with the mainland or with other islands; 1, ≤ 7 connections; 2, > 7 connections per week (Arnold, 2008). We only considered islands with surface area greater than 1 ha; islands with this area can host large populations of reptiles (e.g. Pérez-Mellado *et al.*, 2008).

Land use and climate are additional and potentially important parameters. We calculated the percentage cover of non-natural habitats (urban and cropland) as an additional measure of human impact, on the basis of the 300-m resolution GlobCover land-cover map (Bicheron *et al.*, 2008). Potential evapotranspiration (PET) is a climatic parameter that represents the amount of atmospheric energy, and is strongly related to reptile species richness (Rodríguez *et al.*, 2005). We extracted PET from Trabucco & Zomer (2009). Land-use and PET were only available for a subset of islands with relatively large area, because of their relatively coarse resolution.

We considered two different sources of data for reptile distributions, resulting in two distinct datasets. First, herpetological atlases have been published for several countries/regions (Appendix S1 in Supporting Information). These sources include specific sections on island biodiversity, representing the summary of multiple, intensive surveys, and provide detailed accounts of reptile communities on a number of islands. We thus assume that the islands described in detail in these atlases ('atlas' dataset) represent islands for which high-quality data are available.

For the remaining islands ('no-atlas' islands), we searched for information in a variety of sources, including published papers, the grey literature, publications summarizing the grey literature, and direct surveys by the authors (Sindaco & Jeremčenko, 2008; Sindaco *et al.*, 2013, and references in Appendix S1). Although a herpetological guide of Greece is available (Valakos *et al.*, 2008), it does not detail the species composition on the islands, and Greek islands were therefore included in the 'no-atlas' dataset. The islands in this dataset probably received variable survey effort and, for many of them (about 50%; see Results), we found no information on reptile communities. For each island, we then calculated the number of known native reptile species (see Appendix S1).

### Analysis of reptile data

In both datasets, environmental features (area, isolation, elevation, human population, naval connections and PET) and species richness showed spatial autocorrelation. Autocorrelation may be related to different biogeographical processes occurring among archipelagos and may affect regression analyses. In preliminary analyses, the residuals of ordinary least-squares regressions were significantly autocorrelated ($P < 0.05$; Fig. S1 in Appendix S2), so we integrated autocorrelation into the models, and analysed the data using Moran's eigenvector maps (MEM). MEM allows the spatial arrangement of data points to be translated into explanatory variables (eigenvectors) that capture spatial effects (Dray *et al.*, 2006; Thayn & Simanis, 2013). We used a permutation test to evaluate the autocorrelation of residuals for all models. We generated MEMs on the basis of geographical distances, using the minimum distance that keeps all sites connected (3°; Borcard & Legendre, 2002). We then selected the eigenvectors that best reduce the spatial autocorrelation of residuals, and included them as spatial predictors in generalized linear models, until residual autocorrelation was not significant any more ($P > 0.05$ after a permutation test) (Griffith & Peres-Neto, 2006; Dormann *et al.*, 2007). Comparisons among statistical methods showed that this implementation of MEM is flexible and efficient, even when analysing non-normal data, and provides results that are congruent with other spatially explicit techniques (Dormann *et al.*, 2007; Siesa *et al.*, 2011; Ficetola *et al.*, 2013; Thayn & Simanis, 2013). We performed the MEM analysis using the 'ME' function in the R package SPDEP (Griffith & Peres-Neto, 2006; Bivand *et al.*, 2013).

If surveyed, even the most isolated islands of the Mediterranean, and all the islands with an area greater than 1 ha revealed the presence of some reptile species (e.g. Pérez-Mellado *et al.*, 2008). We assumed that an island received some sampling if it has at least one observation of either native or non-native species. We used MEM (binomial error distribution) to identify the factors that are most likely to determine whether an island has been surveyed. For the 'no-atlas' dataset, the presence or absence of reptile records was considered as the dependent variable, whereas the six environmental variables (area, isolation, elevation, human population, airport presence and naval connections) were the independent variables. We used Akaike's information criterion (AIC) to compare candidate models, including all the possible combinations of independent variables. AIC trades off explanatory power against the number of predictors; parsimonious models explaining more variation have low AIC values and are considered to be the 'best AIC' models (Symonds & Moussalli, 2011). AIC may select overly complex models; we therefore considered a complex model only if it showed a lower AIC than all its simpler nested models (Richards *et al.*, 2011). For each candidate model, we also calculated the AIC weight, which measures the relative support of models (Symonds & Moussalli, 2011).

We then used MEM to relate the observed richness of native reptiles to the geographical and human variables. To compare the outcome of analyses performed on high-quality data, with analyses performed on data with varying quality, this analysis was repeated four times: (1) on all the islands that received some sampling, i.e. for which at least one species (either native or non-native) was recorded; (2) to test whether the relationship between human impact and observed richness was different between the 'atlas' and 'no-atlas' datasets, for which we repeated the analysis of the full dataset, including the interaction between data quality ('atlas'/'no-atlas') and variables representing human impact; (3) on the 'atlas' dataset only; (4) on the islands of the 'no-atlas' dataset which had received some sampling.

Human population was strongly correlated to naval connections (Pearson's $r = 0.84$). As highly correlated variables may bias regression results, we first ran analyses using human population and excluding naval connections, and then repeated them including naval connections and excluding human population. Finally, to evaluate the potential role of climate and land use, we repeated analyses for the subset of islands for which these data were available. Models were built with Poisson-distributed errors or, if the residual deviance was strongly different from residual degrees of freedom, with quasi-Poisson errors.

When necessary, variables were transformed using logarithm (island area, isolation, elevation and human population), square-root (PET) or arcsine-square-root (% land-use) transformations. We tested significance using likelihood-ratio tests. We detected no collinearity issues (variance-inflation factor always $< 5$). Models were run in R 3.0; partial regression plots were built using the package VISREG (R Core Team, 2013).

## Analysis of simulated data

The analysis of virtual data that simulate ecological processes and sampling is a powerful framework, allowing the effect of sampling and analytical methods to be assessed (Zurell *et al.*, 2010). We simulated data on species richness in islands, assuming that the true richness was influenced by biogeographical variables and negatively affected by human impact. Species richness was then virtually sampled with varying effort, assuming a better sampling in human-dominated islands. Simulated data of observed richness were then analysed using the same approach used for real-world data. This allowed us to evaluate the consequences of poor sampling: under what circumstances does better sampling in human-dominated islands outweigh the negative effects of human impact? How good should sampling be in order to obtain reliable estimates of the true relationships between species richness and environmental variables?

In each simulation run, we considered a set of 300 islands. For each island, we generated variables representing environmental features, true species richness and the number of species detected. With regard to environmental features, we focused on area, isolation and presence/absence of airport, as these were the most important independent variables in explaining reptile distribution in the analysis of real-world data. Airport presence was modelled as a function of island area, as larger real islands are more likely to have an airport. The simulated number of species effectively present in each island, $S\_Ns$, was generated from a Poisson distribution with parameter equal to the lambda value predicted by the regression model for the islands in the 'atlas' dataset (Table 1c).

We then generated the number of species detected in each islands, $S\_Nd$. To simulate different sampling intensities across islands, we computed for each island the variable $Q$, representing the probability of species detection, and therefore sampling quality. $Q$ is related to environmental variables, as shown in eqn. 1, built from real-world results:

$$Q = \frac{1}{e^{-[K+\beta_1 \times (S\_airport)+\beta_2 \times (S\_log.area)]}+1} \quad (1)$$

$\beta_1$ and $\beta_2$ are the coefficients of the regression model that relate the presence of reptile observations to environmental variables (Table 2a). Sampling quality, $Q$, varies between 0 and 1, and is better in large islands and in islands hosting an airport (see Results). $K$ is a constant which differs across simulation cycles, allowing variation in the average sampling effectiveness across all the islands to be represented; higher $K$ determines better sampling. In real data, the presence of a herpetological atlas leads to higher $K$. In an island hosting $S\_Ns$ species, each species has a different probability of being detected. The detection probability of each species ($p$) in an island was drawn from a beta distribution with mean of $Q$, and variance of 0.01. $S\_Nd$ was the number of species detected in an island hosting $S\_Ns$ species, and was calculated as the sum of $S\_Ns$ Bernoulli trials, each with probability of success $p$.

Table 1 Poisson regression models relating species richness of native reptiles to geographical and human variables of Mediterranean and Macaronesian islands. (a) Model built for all the islands for which at least one species was recorded. (b) Model built for the same islands used in (a), also including the interaction between data quality ('atlas'/'no-atlas') and the variables representing human impact (human population and airport). (c) Model built for the islands of the 'atlas' dataset. (d) Model built for the islands of the 'no-atlas' dataset for which at least one species was recorded. Correlograms of residual spatial autocorrelation are shown in Appendix S2, Fig. S1.

| Independent | *B* | $\chi^2$ | d.f. | *P* |
|---|---|---|---|---|
| (a) all islands with data | | | | |
| $R^2_{\mathrm{D}} = 0.61$; residual autocorrelation: $P = 0.12$ | | | | |
| Area | 0.29 | 41.82 | 1 | **< 0.001** |
| Isolation | –0.21 | 51.71 | 1 | **< 0.001** |
| Elevation | 0.18 | 8.15 | 1 | **0.004** |
| Human population | 0.09 | 13.74 | 1 | **< 0.001** |
| Airport | 0.02 | 0.05 | 1 | 0.829 |
| MEM eigenvectors | | 42.97 | 2 | **< 0.001** |
| (b) all islands; interactions between data quality and human impact | | | | |
| $R^2_{\mathrm{D}} = 0.68$; residual autocorrelation: $P = 0.07$ | | | | |
| Area | 0.36 | 72.41 | 1 | **< 0.001** |
| Isolation | –0.12 | 18.47 | 1 | **< 0.001** |
| Elevation | 0.09 | 2.57 | 1 | 0.109 |
| Human population | 0.13 | 12.29 | 1 | **< 0.001** |
| Airport | 0.08 | 0.02 | 1 | 0.889 |
| Human population × data quality | –0.16 | 30.21 | 1 | **< 0.001** |
| Airport × data quality | –0.53 | 9.45 | 1 | **0.002** |
| Data quality* | 0.27 | 16.79 | 1 | **< 0.001** |
| MEM eigenvectors | | 21.95 | 1 | **< 0.001** |
| (c) Islands on atlas† | | | | |
| $R^2_{\mathrm{D}} = 0.60$; residual autocorrelation: $P = 0.17$ | | | | |
| Area | 0.43 | 80.74 | 1 | **< 0.001** |
| Isolation | –0.11 | 5.03 | 1 | **0.025** |
| Elevation | 0.12 | 3.00 | 1 | 0.083 |
| Human population | –0.08 | 5.80 | 1 | **0.016** |
| Airport | –0.48 | 11.83 | 1 | **< 0.001** |
| MEM eigenvectors | | 58.18 | 3 | **< 0.001** |
| (d) Islands not on atlas | | | | |
| $R^2_{\mathrm{D}} = 0.71$; residual autocorrelation: $P = 0.20$ | | | | |
| Area | 0.26 | 16.31 | 1 | **< 0.001** |
| Isolation | –0.13 | 12.95 | 1 | **< 0.001** |
| Elevation | 0.11 | 1.56 | 1 | 0.212 |
| Human population | 0.18 | 31.14 | 1 | **< 0.001** |
| Airport | 0.11 | 2.33 | 1 | 0.127 |
| MEM eigenvectors | | 8.46 | 1 | **0.003** |

$R^2_{\mathrm{D}}$: proportion of null-deviance explained by the model.
*Coded as follows: 0, 'no-atlas' dataset; 1, 'atlas' dataset.
†Quasi-Poisson error distribution.

We subsequently evaluated whether regression can successfully estimate the relationships between species richness and the three environmental predictors (area, isolation and airport presence), by relating the detected species richness *S_Nd* to the predictors. For each simulation, we built the regression models, and recorded the model coefficients as well as the average values of *Q* and *S_Nd*/*S_Ns*, which represents the proportion of species successfully detected. We ran simulations with five values of *K* (−2, −1, 0, 1 and 2), and 300 replicates per each value of *K*. A complete description of the simulation methods is reported in Appendix S2; the R code of simulations is available in Appendix S3.

## RESULTS

### Mediterranean island data

We obtained data from 974 islands in 11 countries (Fig. 1). The 'atlas' dataset included 217 islands, and the 'no-atlas' dataset included 757 islands. The 'no-atlas' islands were mostly concentrated in the Eastern Mediterranean basin (72%), although several small islands of the Western Mediterranean were not detailed in herpetological atlases (Fig. 1). Atlas islands tended to be less isolated, smaller, with lower maximum elevation and larger human populations than no-atlas islands (*t*-test, all $P \leq 0.01$), whereas the frequency of ferry connections was similar between the two datasets (Mann–Whitney test, $P = 0.35$). There was nonetheless a wide overlap in geographical features and human impact between the two datasets, and the atlas dataset included the largest and most isolated islands (Appendix S2, Fig. S2).

#### *Which islands are more likely to receive reptile surveys?*

Out of the 757 islands of the 'no-atlas' dataset, we found at least one species record in 372 islands (see Appendix S4 for a list of islands for which we found no reptile records). The best AIC model indicates that the probability of having at least one recorded species was positively related to island area, airport presence and human population (Table 2). Three eigenvectors were also included into the model. They represent significant spatial autocorrelation (Table 2), and reduced residual autocorrelation to non-significant values (Moran's $I = 0.005$, $P = 0.76$). Some models with lower AIC weights did not include human population, and included a negative relationship with elevation (Table 2b). Area and airport presence were included in all candidate models with high support.

#### *Analysis of observed species richness*

When all the islands with at least one record were analysed (589 islands), species richness clearly increased with island area and elevation, and was negatively related to isolation. The observed richness was not related to airport presence, but was positively related to human population. Two eigenvectors were also included into the model, to take spatial autocorrelation into account (Table 1a).

We detected significant interactions between human population, airport presence and dataset of origin ('atlas' vs. 'no-atlas'): the relationship between reptile richness and the variables representing human impact was significantly more negative in the 'atlas' dataset than in the 'no-atlas' dataset (significant

**Table 2** Binomial spatial eigenvector models relating presence/absence of reptile records on Mediterranean and Macaronesian islands to human and geographical variables. (a) Best model, selected on the basis of Akaike's information criterion (AIC); (b) environmental variables included in the other candidate models. Only the 12 models with lowest AIC are reported here; all these models also include MEM eigenvectors. *K*, number of parameters in the model; *w*, AIC weight; $R^2_D$, proportion of null deviance explained by the model.

| Independent variables | *B* | $\chi^2$ | d.f. | *P* | *K* | AIC | *w* | $R^2_D$ |
|---|---|---|---|---|---|---|---|---|
| (a) best-AIC model | | | | | 7 | 811.5 | 0.63 | 0.24 |
| Area | 0.55* | 12.20 | 1 | **< 0.001** | | | | |
| Human population | 0.27 | 5.48 | 1 | **0.020** | | | | |
| Airport | 15.37† | 7.65 | 1 | **0.006** | | | | |
| MEM eigenvectors | | 147.64 | 3 | **< 0.001** | | | | |
| (b) Alternative models | | | | | | | | |
| Area (+), airport presence (+), elevation (−) | | | | | 7 | 814.1 | 0.17 | 0.24 |
| Area (+), airport presence (+) | | | | | 6 | 815.0 | 0.11 | 0.23 |
| Area (+), human population (+), elevation (−) | | | | | 7 | 817.1 | 0.04 | 0.23 |
| Area (+), human population (+) | | | | | 6 | 817.1 | 0.04 | 0.23 |
| Airport presence (+), isolation (+), human population (+) | | | | | 7 | 821.6 | 0.004 | 0.23 |
| Airport presence (+),human population (+) | | | | | 6 | 821.7 | 0.004 | 0.23 |
| Area (+), elevation (−) | | | | | 6 | 822.3 | 0.003 | 0.23 |
| Area (+) | | | | | 5 | 823.6 | 0.001 | 0.22 |
| Isolation (+), human population (+) | | | | | 6 | 828.4 | < 0.001 | 0.22 |
| Human population (+) | | | | | 5 | 828.6 | < 0.001 | 0.22 |
| Airport presence (+), elevation (+), naval connections (+) | | | | | 7 | 840.0 | < 0.001 | 0.21 |

*Used as $\beta_2$ in eqn. 1.

†Used as $\beta_1$ in eqn. 1.

interactions in Table 1b). When only the 'atlas' islands were analysed, richness was negatively related to human population and airport presence. Parameters representing human impact therefore showed opposite regression coefficients, compared to the results of the full dataset (Table 1c, Fig. 2). Conversely, when we analysed only the 'no-atlas' islands with at least one presence record (372 islands), the relationships between species richness and parameters representing human impact were similar to the analysis of the full dataset. The relationship between reptile richness and human population was positive, whereas the relationship with airport presence was not significant (Table 1d, Fig. 2). Spatial eigenvectors were included in all the models, and successfully reduced residual spatial autocorrelation to non-significant values (Table 1; Appendix S2, Fig. S1).

We obtained analogous results for models using naval connections instead of human population (Table S1 in Appendix S2). The relationships between richness and variables representing human impact (naval connections and airport presence) were generally positive both in the analysis of the full dataset, and in the analysis of the 'no-atlas' islands, but the interaction between naval connections and dataset was strongly significant (Table S1b), because the relationship between reptile richness and naval connections was weakly negative for the islands of the 'atlas' dataset (Table S1).

We obtained data on land use and climate (PET) for 307 islands ('atlas', 78 islands; 'no-atlas', 229 islands). After taking into account the other variables, reptile richness was not significantly related to either land use or PET (Appendix S2, Table S2). It should nonetheless be remarked that data on land use and climate were only available for the largest islands. For instance, in this dataset the median area was 9.11 km$^2$, whereas the median area of the global dataset was 0.65 km$^2$.

### Simulations

The five values of *K* considered allow potential variation in sampling quality (Fig. 3a) to be well represented. When $K = -2$, only 21% of present species were detected; the proportion of detected species rose to 54% with $K = 0$, and to 87% with $K = 2$ (Fig. 3a). The results of regression models were strongly affected by sampling quality. For both island area and isolation, coefficients were close to the actual values when $K \geq 1$ (i.e. high sampling quality and *c.* 75% of species or more detected), whereas coefficient estimates were less accurate with lower values of *K* (Fig. 3b,c), although the sign of the coefficient remained unchanged for both area and isolation. The coefficient of the relationship between airport presence and species richness was strongly affected by changes to *K*. The relationship between airport presence and richness was positive if $K = -2$, and approached the real, negative value only when $K \geq 1$ (Fig. 3d). Poisson regression coefficients were therefore strongly affected by sampling bias, and the relationship with variables representing human presence may be inverted when bias was strong.

## DISCUSSION

### Human presence, data quality and known richness

Species richness is one of the commonest measures of biodiversity, and is a basis of ecological analyses at multiple scales. However, the outcome of species-richness analyses may be severely affected by data quality. Human presence, accessibility and attractiveness of areas are strongly related to sampling effort and the quality of biodiversity information (Sánchez-Fernández

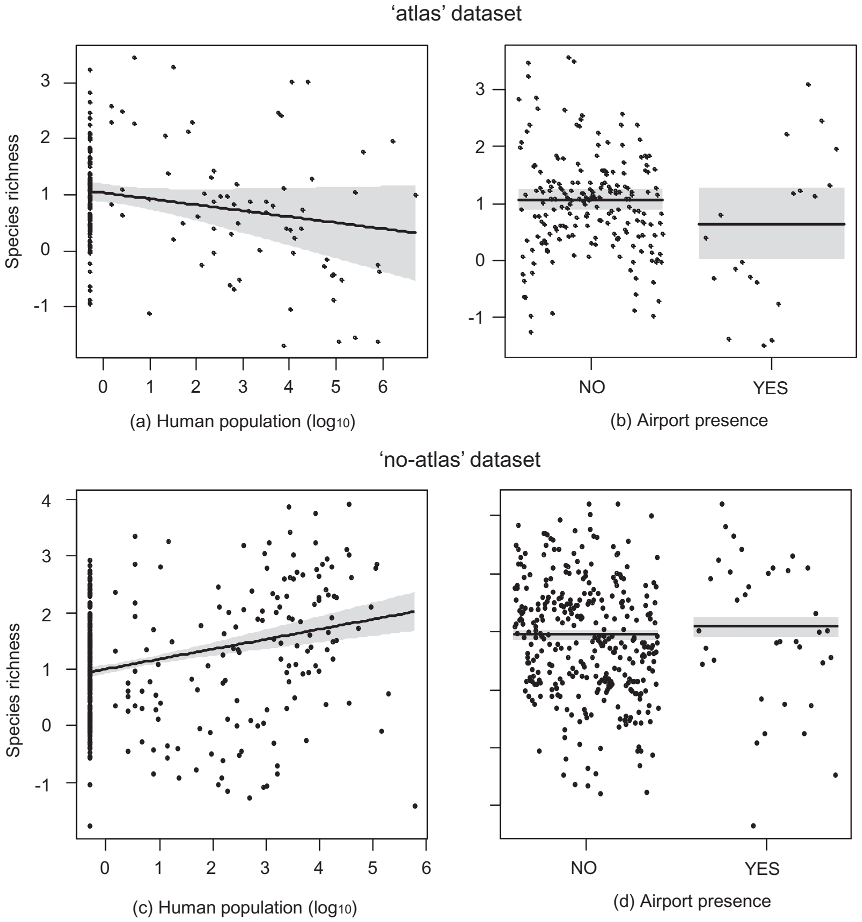


Figure 2 Partial residual plots showing the relationship between species richness of reptiles and variables representing human impact (total human population and airport presence) for Mediterranean and Macaronesian islands included ('atlas' dataset) and not included ('no-atlas' dataset) in faunistic/herpetological atlases. Shaded areas represent 95% confidence bands.

*et al.*, 2008; Kent & Carmel, 2011; Ficetola *et al.*, 2013; Yang *et al.*, 2013). Across the Mediterranean, sampling is clearly better in islands with good transport connections and housing large human populations. As a result, regressing species richness against parameters representing human presence (e.g. human population, airport presence and naval connections) may result in apparently positive relationships (Table 1, Fig. 2c–d), even though human impact is likely to have negative effects (Table 1c, Fig. 2a–b). These two facets of human impact (positive on biodiversity knowledge; likely to be negative on actual biodiversity) represent a major challenge for ecogeographical studies, as not taking them into account would result in biased analyses and misunderstanding of the true human impact.

This interpretation of our results is supported by simulation analyses. Low sampling quality determines less accurate estimates for all relationships (e.g. inflated error in the relationship between isolation and richness; Fig. 3c), but the bias is particularly relevant for parameters related to both richness and detection probability, such as airport presence (Fig. 3d). Actually, the negative effects of human impact in our simulations could only be successfully assessed if island communities are well-sampled, and at least 75% of species are detected (Fig. 3). We do not know the actual level of completeness of surveys in the Mediterranean islands of our datasets. Richness underestimations of more than 50% are not infrequent in poorly-surveyed localities (Ficetola *et al.*, 2013; Yang *et al.*, 2013), but survey quality was not evenly distributed across islands in our real-world dataset. Even for the 'no-atlas' dataset, some islands received repeated focused surveys, whereas others have only been occasionally sampled.

Some studies suggested that positive relationships between biodiversity and human presence may arise because certain climates (e.g. warmer, with more precipitation) favour both human activities and biodiversity (Stohlgren *et al.*, 2006). However, this effect is unlikely to influence our results, because islands in our set span a limited latitudinal range (Fig. 1), and share a similar Mediterranean climate. In fact, observed species richness was not related to evapotranspiration, and the interactions between sampling quality and human impact remained evident when taking climate into account (Table S2). A positive correlation between reptile richness and human presence might also arise because humans settle on islands where water is available and with more habitats, and these parameters can also positively affect species richness (Kadmon & Allouche, 2007). Our models nonetheless considered island area and elevation, which strongly correlate with habitat availability and diversity (Davidar *et al.*, 2002; Kreft *et al.*, 2008). Furthermore, the positive correlation was only observed for the 'no-atlas' islands, supporting the role of sampling quality.

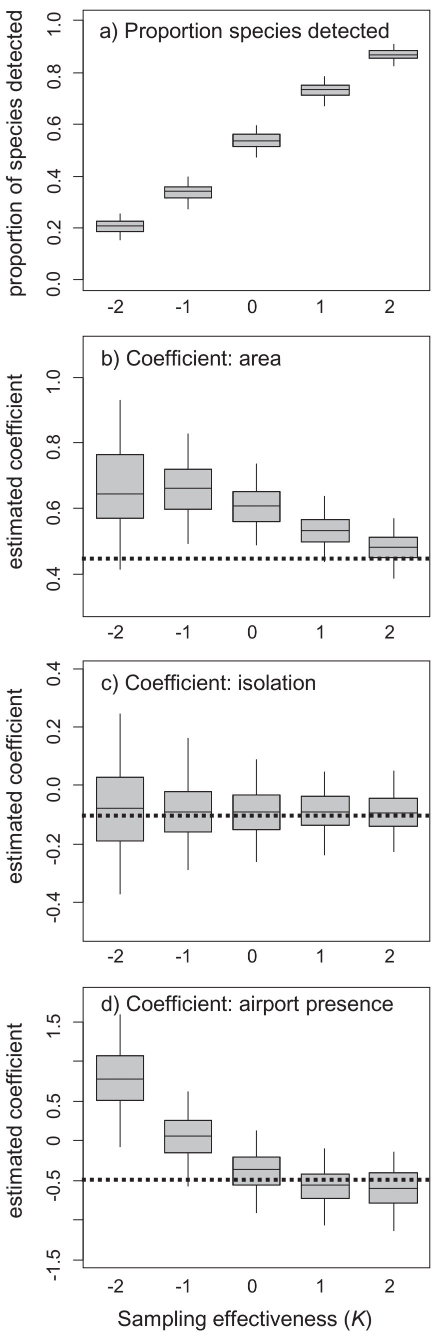


Figure 3 Results of simulations: effect of different values of sampling effectiveness (*K*) on (a) average proportion of reptile specie detected in Mediterranean and Macaronesian islands; coefficients of Poisson regression relating observed species richness to (b) log(area), (c) log(isolation) of the islands and (d) airport presence. Box include the second and the third quartile of simulation results; error bars are 95% CI, the bold dotted lines represent the true parameters used to generate the data.

## Human impact and reptile richness

For the 'atlas' islands, the relationships between species richness and both human population and airport presence were clearly negative, suggesting a strong impact on reptile communities (Table 2, Fig. 2a–b). Human activities in the Mediterranean basin date back to the Neolithic, and their effects extend to most of the islands. The introduction of alien species is likely to be the factor with the strongest negative effects on native reptiles (Corti *et al.*, 1999). Many alien species have been introduced to Mediterranean islands in historical times, but the rate of introductions is quickly rising due to increasing commercial exchanges (Pinya & Carretero, 2011). Exhaustive data on trade intensity were not available, so we used airport presence and naval connections as proxies of tourism and trade between the islands and the continents. Airport traffic is a major predictor of economic development (Green, 2007), and airport presence is one of the strongest predictors of the richness of alien reptiles on islands (Ficetola & Padoa-Schioppa, 2009). Other human impacts may have weaker effects on reptiles. Many Mediterranean reptiles are well-adapted to traditional agriculture, and may therefore tolerate the partial loss of natural habitats. Nevertheless, recent urbanization poses the greatest threat, and some endemic reptiles are actually endangered by habitat loss and degradation. Finally, persecution by humans may have locally important effects, particularly on snakes (Mullin & Seigel, 2009).

## How can we improve broad-scale ecogeographical analyses?

Given the heterogeneous quality of broad-scale distribution data, how can we improve their use for ecological analysis? One potential approach is to identify localities for which data quality is sufficiently high, and to focus analyses on well-surveyed sites (Yang *et al.*, 2013). This is the approach we used here, as we assumed an adequate quality for the 'atlas' dataset. Nevertheless, this approach remains subjective, and we did not have direct control of the actual quality for many of the islands in the 'atlas' dataset. Even if these islands generally received repeated and intensive sampling, information could be incomplete for some of them. Alternative and more objective approaches include using accumulation curves to estimate the completeness of data, but this would require access to the original data, which is rarely available for broad-scale databases. Furthermore, estimates obtained through accumulation curves may not be consistently reliable (Sánchez-Fernández *et al.*, 2011; Yang *et al.*, 2013).

A different approach would be to obtain estimates of sampling effort (e.g. accessibility to researchers, number of known surveys performed), and integrate them into models to improve the results of analyses (Sánchez-Fernández *et al.*, 2008; Gómez-Rodríguez *et al.*, 2012; Ficetola *et al.*, 2013). These measures were not available for all the islands in our dataset – for instance, accessibility maps have a resolution of 1 km$^2$ (Nelson, 2008), and many islands are smaller than this – but could be successfully integrated into studies that focus on larger spatial units. Finally, some analytical techniques have been proposed for taking the complex effects of variables on detection probability and richness into account, such as zero-inflated Poisson models (ZIP; Zuur *et al.*, 2009). Even with ZIP, however, we were unable to successfully capture the negative relationships between human presence and species richness. Indeed, the ZIP results were no better than those of standard regression models (Appendix S2, Table S3, Fig. S3).

Even if there is general awareness that species distribution data are rarely complete (Lomolino, 2004), the impact of this underestimation on biodiversity analysis remains underappreciated (Kéry, 2011). Uneven geographical sampling can, however, strongly affect patterns of species richness and may even invert ecogeographical relationships if the same factors is related to both sampling quality and actual species richness. Insufficient sampling can have consequences both for ecological understanding and practical conservation planning. For instance, misjudging the effects of human impact may result in a lack of management actions. In the last decade, a growing body of literature has investigated the effects of imperfect detection on local-scale studies, although this issue has attracted less attention in the macro-ecological literature. More research on this theme will allow a better understanding of broad-scale ecological patterns, with potential consequences on biodiversity conservation.

## ACKNOWLEDGEMENTS

We thank J. A. Diniz-Filho and two reviewers for comments on the manuscript. C. Arnold provided his database of Mediterranean Islands. G.F.F. was funded by a scholarship of University Milano-Bicocca.

## REFERENCES

Arnold, C. (2008) *Mediterranean islands*. Survival Books, London.

Bicheron, P., Huc, M., Henry, C., Bontemps, S. & GLOBCOVER partners (2008) *GLOBCOVER products description manual*. Medias France, European Space Agency, Toulouse.

Bivand, R., Altman, M., Anselin, L. *et al.* (2013) *spdep: spatial dependence: weighting schemes, statistics and models.* R package version 0.5-62. Available at: http://cran.r-project.org/web/packages/spdep/

Borcard, D. & Legendre, P. (2002) All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, **153**, 51–68.

Comte, L. & Grenouillet, G. (2013) Species distribution modelling an imperfect detection: comparing occupancy versus consensus methods. *Diversity and Distributions*, **19**, 996–1007.

Corti, C., Massetti, M., Delfino, M. & Pérez-Mellado, V. (1999) Man and herpetofauna of the mediterranean islands. *Revista Española de Herpetología*, **13**, 83–100.

Davidar, P., Yogananad, K., Ganesh, T. & Devy, S. (2002) Distribution of forest birds and butterflies in Andaman islands, Bay of Bengal: nested patterns and process. *Ecography*, **25**, 5–16.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dray, S., Legendre, L. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Ficetola, G.F. & Padoa-Schioppa, E. (2009) Human activities alter biogeographical patterns of reptiles on Mediterranean islands. *Global Ecology and Biogeography*, **18**, 214–222.

Ficetola, G.F., Bonardi, A., Sindaco, R. & Padoa-Schioppa, E. (2013) Estimating patterns of reptile biodiversity in remote regions. *Journal of Biogeography*, **40**, 1202–1211.

Ficetola, G.F., Rondinini, C., Bonardi, A., Katariya, V., Padoa-Schioppa, E. & Angulo, A. (2014) An evaluation of the robustness of global amphibian range maps. *Journal of Biogeography*, **41**, 211–221.

Foody, G.M. (2011) Impacts of imperfect reference data on the apparent accuracy of species presence–absence models and their predictions. *Global Ecology and Biogeography*, **20**, 498–508.

Gómez-Rodríguez, C., Bustamante, J., Díaz-Paniagua, C. & Guisan, A. (2012) Integrating detection probabilities in species distribution models of amphibians breeding in Mediterranean temporary ponds. *Diversity and Distributions*, **18**, 260–272.

Green, R.K. (2007) Airports and economic development. *Real Estate Economics*, **35**, 91–112.

Griffith, D.A. & Peres-Neto, P.R. (2006) Spatial modelling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.

Gu, W.-D. & Swihart, R.K. (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation*, **116**, 195–203.

Kadmon, R. & Allouche, O. (2007) Integrating the effects of area, isolation, and habitat heterogeneity on species diversity: a unification of island biogeography and niche theory. *The American Naturalist*, **170**, 443–454.

Kalmar, A. & Currie, D.J. (2006) A global model for island biogeography. *Global Ecology and Biogeography*, **15**, 72–81.

Kent, R. & Carmel, Y. (2011) Presence-only versus presence–absence data in species composition determinant analyses. *Diversity and Distributions*, **17**, 474–479.

Kerr, J.T., Kharouba, H.M. & Currie, D.J. (2007) The macroecological contribution to global change solutions. *Science*, **316**, 1581–1584.

Kéry, M. (2011) Towards the modelling of true species distributions. *Journal of Biogeography*, **38**, 617–618.

Kreft, H., Jetz, W., Mutke, J., Kier, G. & Barthlott, W. (2008) Global diversity of island floras from a macroecological perspective. *Ecology Letters*, **11**, 116–127.

Lahoz-Monfort, J.J., Guillera-Arroita, G. & Wintle, B.A. (2014) Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, **23**, 504–515.

Lomolino, M.V. (2004) Conservation biogeography. *Frontiers of biogeography: new directions in the geography of nature* (ed. by M.V. Lomolino and L.R. Heaney), pp. 293–296. Sinauer Associates, Sunderland, MA.

McArthur, R.H. & Wilson, E.O. (1967) *The theory of island biogeography*. Princeton University Press, Princeton, NJ.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.A. & Hines, J.E. (2006) *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Academic Press, Burlington, MA.

Miller, D.A., Nichols, J.D., McClintock, B.T., Campbell Grant, E.H., Bailey, L.L. & Weir, L.A. (2011) Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification. *Ecology*, **92**, 1422–1428.

Mullin, S.J. & Seigel, R.A. (eds) (2009) *Snakes: ecology and conservation*. Cornell University Press, Ithaca, NY.

Nelson, A. (2008) *Travel time to major cities: a global map of accessibility*. Office for Official Publications of the European Communities, Luxembourg.

Pérez-Mellado, V., Hernández-Estévez, J., García-Díez, T., Terrassa, B., Misericordia Ramón, M., Castro, J., Picornell, A., Martín-Vallejo, J. & Brown, R. (2008) Population density in *Podarcis lilfordi* (Squamata, Lacertidae), a lizard species endemic to small islets in the Balearic Islands (Spain). *Amphibia-Reptilia*, **29**, 49–60.

Pinya, S. & Carretero, M.A. (2011) The Balearic herpetofauna: a species update and a review on the evidence. *Acta Herpetologica*, **6**, 59–80.

R Core Team (2013) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Richards, S.A., Whittingham, M.J. & Stephens, P.A. (2011) Model selection and model averaging in behavioural ecology: the utility of the IT-AIC framework. *Behavioral Ecology and Sociobiology*, **65**, 77–89.

Rodríguez, M., Belmontes, J.A. & Hawkins, B.A. (2005) Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica*, **28**, 65–70.

Sánchez-Fernández, D., Lobo, J.M., Abellán, P., Ribera, I. & Millán, A. (2008) Bias in freshwater biodiversity sampling: the case of Iberian water beetles. *Diversity and Distributions*, **14**, 754–762.

Sánchez-Fernández, D., Lobo, J.M., Abellán, P. & Millán, A. (2011) How to identify future sampling areas when information is biased and scarce: an example using predictive models for species richness of Iberian water beetles. *Journal for Nature Conservation*, **19**, 54–59.

Siesa, M.E., Manenti, R., Padoa-Schioppa, E., De Bernardi, F. & Ficetola, G.F. (2011) Spatial autocorrelation and the analysis of invasion processes from distribution data: a study with the crayfish *Procambarus clarkii*. *Biological Invasions*, **13**, 2147–2160.

Sindaco, R. & Jeremčenko, V.K. (2008) *The reptiles of the Western Palearctic, vol. 1. Annotated checklist and distributional atlas of the turtles, crocodiles, amphisbaenians and lizards of Europe, North Africa, Middle East and Central Asia*. Belvedere, Latina, Italy.

Sindaco, R., Venchi, A. & Grieco, C. (2013) *The reptiles of the Western Palearctic, vol. 2. Annotated checklist and distributional atlas of the snakes of Europe, North Africa, Middle East and Central Asia*. Belvedere, Latina, Italy.

Stohlgren, T.J., Barnett, D.T., Flather, C., Fuller, P., Peterjohn, B., Kartesz, J. & Master, L.L. (2006) Species richness and pattern of invasion in plants, birds, and fishes in the United States. *Biological Invasions*, **8**, 427–447.

Symonds, M.R.E. & Moussalli, A. (2011) A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology*, **65**, 13–21.

Thayn, J.B. & Simanis, J.M. (2013) Accounting for spatial autocorrelation in linear regression models using spatial filtering with eigenvectors. *Annals of the Association of American Geographers*, **103**, 47–66.

Trabucco, A. & Zomer, R.J. (2009) *Global potential evapo-transpiration (Global-PET) and global aridity index (Global-Aridity) geo-database*. CGIAR Consortium for Spatial Information, Washington, DC. Available at: http://www.csi.cgiar.org/

Triantis, K.A., Guilhaumon, F. & Whittaker, R.J. (2012) The island species–area relationship: biology and statistics. *Journal of Biogeography*, **39**, 215–231.

Valakos, E.D., Pafilis, P., Sotiropoulos, K., Lymberakis, P., Maragou, P. & Foufopoulos, J. (2008) *The amphibians and reptiles of Greece*. Chimaira, Frankfurt am Main.

Vale, M.M. & Jenkins, C.N. (2012) Across-taxa incongruence in patterns of collecting bias. *Journal of Biogeography*, **39**, 1744–1748.

Van Buskirk, J.V. (2005) Local and landscape influence on amphibian occurrence and abundance. *Ecology*, **86**, 1936–1947.

Whittaker, R.J. & Fernández-Palacios, J.M. (2007) *Island biogeography*. Oxford University Press, Oxford.

Yang, W.-Y., Ma, K. & Kreft, H. (2013) Geographical sampling bias in a large distributional database and its effects on species richness–environment models. *Journal of Biogeography*, **40**, 1415–1426.

Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. & Grimm, V. (2010) The virtual ecologist approach: simulating data and observers. *Oikos*, **119**, 622–635.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009) *Mixed effects models and extensions in ecology with R*. Springer, New York.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

**Appendix S1** References used as source for species distribution in the 'atlas' and in the 'no-atlas' datasets, and for the identification of non-native species.

**Appendix S2** Additional analyses: correlograms of residual autocorrelation; complete description of methods used for simulations; results of regression models considering naval connections, land-use and climate; analyses using zero-inflated Poisson regression.

**Appendix S3** R code used to run simulations.

**Appendix S4** List of islands for which no reptile records were found.

## BIOSKETCH

**G. Francesco Ficetola** is a researcher combining multiple approaches (landscape ecology, macroecology, evolutionary ecology) to improve knowledge of the ecology and conservation of amphibians and reptiles.

Author contributions: G.F.F., A.B., E.P.-S and E.R. designed the study; A.B., M.C., E.R., R.S and G.F.F. gathered species data; G.F.F., A.B., M.C and A.Q. performed the analyses; G.F.F. wrote the first draft of the manuscript. All authors commented on and approved the final manuscript version.

Editor: José A. F. Diniz-Filho