



¿Es posible predecir la distribución geográfica de las especies basándonos en variables ambientales?

Jorge M. Lobo

Departamento de Biodiversidad y Biología Evolutiva
Museo Nacional de Ciencias Naturales
C/ José Gutiérrez Abascal, 2.
28006, Madrid
mcnj117@mncn.csic.es

Hacia un Proyecto CYTED para el Inventario y Estimación de la Diversidad Entomológica en Iberoamérica: PrIBES-2000.
Martín-Piera, F., J.J. Morrone & A. Melic (Eds.)
ISBN:
m3m : Monografías Tercer Milenio
vol. 1, SEA, Zaragoza, 2000
pp. 55 - 68.

PrIBES-2000:

Proyecto para Iberoamérica de Entomología Sistemática.
<http://entomologia.rediris.es/pribes2000>
Coordinador del proyecto :
Dr. Fermín Martín-Piera
Dpto. Biodiversidad y Biología Evolutiva
Museo Nacional Ciencias Naturales-CSIC
c/ José Gutiérrez Abascal, 2
28006 Madrid (ESPAÑA)
fermin@mncn.csic.es

Coeditores del volumen:

Sociedad Entomológica Aragonesa (SEA),
<http://entomologia.rediris.es/sea>
Avda. Radio Juventud, 6
50012 Zaragoza (ESPAÑA)
Director Publicaciones: Antonio Melic
amelic@retemail.es

CYTED—Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo.
Coordinador Internacional:
Dr. Gonzalo Halffter.
Instituto de Ecología
2,5 km antigua ctra. a Coatepec
Apdo. Correos, 63
Xapala 9100, Veracruz (MÉXICO).

Con la colaboración de
Instituto HUMBOLDT
COLOMBIA

¿ES POSIBLE PREDECIR LA DISTRIBUCIÓN GEOGRÁFICA DE LAS ESPECIES BASÁNDONOS EN VARIABLES AMBIENTALES?

Jorge M. Lobo

Resumen

Conocer el reparto geográfico del número total de especies de un taxón es una tarea tan colosal que, incluso en países con un conocimiento taxonómico y biogeográfico modélico, las áreas cuyos inventarios pueden considerarse bien establecidos son una modesta proporción del total. Ante esta situación, se propone que la elaboración de funciones predictivas basadas en variables ambientales, puede ser un método fiable y efectivo para conocer la biodiversidad de un taxón en un área concreta insuficientemente colectada. Mediante diversos ejemplos, se efectúa un repaso de los métodos e inconvenientes a considerar a la hora de realizar funciones predictivas de este tipo y se proponen algunas soluciones que pueden ser de utilidad para mejorar la efectividad de dichas funciones.

Palabras clave: Biodiversidad, distribución riqueza, modelos predictivos.

It is possible to predict the geographical distribution of species on the basis of environmental variables?

Abstract

Getting to know the geographical distribution of the various species of a taxon is a task so colossal that, even in countries with an important taxonomic and biogeographic body of knowledge, the well-surveyed areas are a rather modest proportion of the total. The development of predictive functions based on environmental variables is proposed as a way to tackle this problem. Predicting species richness can be an effective method to describe a taxon's biodiversity in a poorly collected area. The methods and problems associated with accomplishing such predictive functions are reviewed, with the help of various examples, and some solutions are proposed that may be useful to improve their efficiency.

Key words: Biodiversity, Species richness distribution, Predictive models.

1. – LA CUESTIÓN

La disminución y desaparición de poblaciones y especies debida a la perturbación ejercida sobre el medio por las actividades humanas, constituye uno de los conflictos ambientales más graves que debe enfrentar actualmente la humanidad (Wilson, 1988; May *et al.*, 1995). Ante la magnitud e importancia del problema, parece adecuado recomendar que cualquier estrategia de protección de la biodiversidad deba sustentarse en evidencias científicas (Murphy, 1990; Noss, 1991). Sin embargo, la comunidad científica sólo posee respuestas aproximadas sobre los procesos que generan y mantienen la diversidad biológica y es necesario, como paso preliminar ineludible, recopilar, compendiar y cartografiar la información biológica actualmente dispersa en la literatura y las colecciones (Soule, 1991; Lubchenco *et al.*, 1991; Reid *et al.*, 1992; Systematics Agenda 2000, 1994). Esta colosal tarea necesita de una aproximación nueva que fusione y enlace las fronteras de disciplinas, a menudo separadas, como la Sistemática, la Biogeografía y la Ecología que, ante la denominada crisis de la biodiversidad, tienen la obligación de compartir objetivos y amalgamarse.

Los patrones de variación de la diversidad y los procesos que los producen cambian con la escala espacial de análisis (Hengeveld, 1990), de modo que los fundamentos basados en el estudio de las comunidades locales pueden ser inaplicables en el caso de escalas espaciales mayores. Estudiar las relaciones entre distribución, diversidad y condiciones ambientales a grandes escalas espaciales es, no sólo un requisito imprescindible si queremos conocer la distribución geográfica de la biodiversidad, sino una posibilidad científica con nuevas expectativas. En esta aproximación al conocimiento de la distribución de la biodiversidad, debe destacarse el especial papel que juega la Biogeografía y la consideración que merece el espacio. Los hallazgos científicos de la Ecología Demográfica y de la Genética de Poblaciones, han producido una imagen de la naturaleza en la cual las especies se ajustan gradualmente al medio mediante la modificación de sus frecuencias génicas o demográficas, mientras que las condiciones abióticas se encuentran en una situación de equilibrio fluctuante alrededor de unas condiciones medias estables. En esta visión del mundo natural, las comunidades están formadas por asociaciones de especies plásticas que evolucionan conjuntamente y tienden al equilibrio con el medio que las alberga (Mayr, 1963; Avers, 1989). Sin embargo, la Paleontología, la Climatología del Cuaternario, la Genética y la Biogeografía, han aportado resultados que avalan una modificación de este esquema de pensamiento. En la evolución pueden producirse saltos evolutivos sustanciales debidos a modificaciones genéticas relativamente pequeñas que afectan a los genes reguladores. Estos complejos génicos pueden ser muy antiguos y mantenerse intactos desde los tiempos de origen de un taxón (Turner y Paterson, 1991; Ahlberg *et al.*, 1996; Rollo, 1994; Slack, 1994), de modo que las adaptaciones eco-fisiológicas que actualmente observamos en un grupo de especies, podrían haberse generado en tiempos y lugares muy alejados de los actuales (Beerling y Woodward, 1996; Farrell *et al.*, 1992; Latham y Ricklefs, 1993). Además, el estudio de las condiciones climáticas durante los dos últimos millones de años demuestra que el medio abiótico está en un permanente estado de no-equilibrio y que las condiciones ambientales siguen tendencias cambiantes de distinta duración. La respuesta de las especies ante esos cambios ha sido la dispersión sin apenas evidencias de modificaciones morfológicas importantes (Coope, 1979; Elias, 1994; Huntley y Birks, 1983; Webb *et al.*, 1993). De este modo, las especies exhiben una respuesta individual ante los cambios ambientales con velocidades de dispersión diferentes y sin existir una tendencia uniforme en la dirección del movimiento. De acuerdo con esta visión, las comunidades actuales serían asociaciones provisionales de especies que resultan de la yuxtaposición de taxones con distintas historias evolutivas, no existiendo unas condiciones de equilibrio ideales entre las distintas especies que las componen.

Si las especies son relativamente rígidas y sólo una pequeña parte del genoma puede adaptarse gradualmente al medio y si, además, el medio abiótico está en permanente estado de transformación, la reacción más eficaz de una especie ante una modificación ambiental consiste en variar espacialmente su distribución geográfica a fin de acoplar espacialmente los requerimientos eco-fisiológicos en aquellos enclaves que reproducen las condiciones ambientales para las que está adaptada (adaptación espacial; véase Hengeveld, 1997 y Lobo, 1999 para una revisión sobre el tema). Este modo de entender el medio natural concede gran importancia a la dimensión espacial y es de gran ayuda si queremos explicar la diversidad y la composición faunística de un determinado territorio. La Biogeografía se convierte así en

una disciplina de alto valor explicativo, que estudia el dinamismo de las áreas de distribución como parte del proceso evolutivo y que busca regularidades en la forma, tamaño, posición y otras características de las distribuciones. El estudio de la variación espacial de la biodiversidad es uno de los asuntos en los que está interesada la Biogeografía y que le concierne especialmente, ya que una localidad con gran número de especies es esencialmente aquella en la que convergen muchas áreas de distribución. Sin embargo, para superar una fase narrativa, la Biogeografía debe reforzarse metodológicamente mediante la utilización de la acumulación exhaustiva de información, la búsqueda de patrones subyacentes, la propuesta de interpretaciones para los mismos y, por último, la elaboración de modelos con valor predictivo que permitan rechazar estadísticamente hipótesis nulas formuladas con anterioridad (Legendre, 1990; Walter y Paterson, 1994; Blondel, 1995; Hengeveld, 1997).

Actualmente, el estudio de la biogeografía de la biodiversidad tiene como principales objetivos: (1) elaborar estrategias generales de inventario e identificación rápidas y eficaces, (2) obtener una información de base extensa mediante la compilación y cartografía del conocimiento disponible sobre la diversidad biológica, y (3) buscar modelos capaces de predecir el número de especies u otra característica relacionada con la biodiversidad en ausencia de datos exhaustivos.

Las potentes herramientas informáticas disponibles, en especial los programas de bases de datos y los sistemas de información geográfica, facilitan enormemente la recopilación y cartografía de la información biológica (Davis, 1994; Margules y Austin, 1994; Miller, 1994; ver <http://viceroy.eeb.uconn.edu/Biota>). El análisis ulterior de esta información debe permitir evaluar el esfuerzo taxonómico y faunístico realizado, describir el emplazamiento de las áreas y los grupos taxonómicos necesitados de prospección, estimando el trabajo pendiente necesario para completar los atlas de la biodiversidad (véase Morrone en este mismo volumen). Sin embargo, son pocos los casos en los que se está acometiendo esta labor de manera permanente (Beaufort y Maurin, 1988; Harding y Sheail, 1992; ver también <http://kaos.erin.gov.au> (Environmental Resources Information Network, Australia), <http://www.conabio.gob.mx> (Comisión Nacional para el Conocimiento y el Uso de la Biodiversidad, México), o <http://www.inbio.ac.cr> (Instituto Nacional de Biodiversidad, Costa Rica) y muchos los lamentos sobre el poco esfuerzo dedicado a la recolección, determinación y estudio sistemático animal y vegetal, sobre todo en aquellos países económicamente pobres, pero ricos en especies y endemismos.

2. - ¿CUANTO ESFUERZO ES SUFICIENTE?

Cuantificar el esfuerzo de muestreo es imprescindible para evaluar la calidad de la información biológica disponible, pero no es sencillo determinar cuan completos son nuestros inventarios. Rara vez puede demostrarse que el incremento en el esfuerzo de muestreo de un territorio permite descubrir nuevas especies de un grupo cualquiera (Gaston, 1996). Sin embargo, los datos sobre la distribución de una especie se limitan a constatar su presencia, sin existir medida alguna que permita diferenciar la ausencia de la falta de muestreo. En los pocos casos en los que se manejan datos sobre la distribución geográfica del número de especies de un grupo, se evita mencionar generalmente que dicha información procede de diversas fuentes, es decir, que ha sido recogida utilizando distintos métodos y esfuerzos. De este modo, resulta imposible discernir cuál es el efecto de la variación en la intensidad de muestreo sobre la distribución del número de especies y,

por tanto, cuál es el patrón real de distribución (Williams y Gaston, 1994). Algunas aproximaciones metodológicas permiten realizar comparaciones entre la riqueza de especies de distintas localidades, corrigiendo el número de especies de un área según el esfuerzo de muestreo realizado en ella (Hurlbert, 1971; Prendergast *et al.*, 1993). Estas técnicas pretenden obtener estimas relativas comparables de riqueza entre distintas localidades, pero no ofrecen información sobre el número absoluto de especies que contienen.

Pueden obtenerse medidas absolutas del número total de especies de un área, examinando las curvas de acumulación, es decir, la relación asintótica entre el incremento en el esfuerzo de muestreo y el número de especies colectadas (Colwell y Coddington, 1995; Colwell, 1997; Soberón y Llorente, 1993). Las curvas de acumulación permiten estimar, además, la cantidad de esfuerzo necesario para obtener un determinado incremento en el número de especies o el número de especies nuevas que encontraremos si aumentamos el esfuerzo colector en una tasa dada (Soberón y Llorente, 1993). Debido a estas posibilidades, el análisis de la relación entre el esfuerzo de muestreo y el número de especies colectadas, resulta especialmente útil para discernir las localidades “bien” y “mal” muestreadas. Conociendo las técnicas de colecta de un grupo taxonómico y el trabajo que es preciso realizar para su ejecución, puede establecerse un nivel de esfuerzo para el cual la tasa de incremento en la riqueza sea lo suficientemente baja, de modo que el inventario faunístico resultante pueda considerarse bien establecido o difícilmente mejorable. También es posible obtener estimas absolutas de la diversidad biológica, analizando la distribución de frecuencias lognormal del número de especies en clases de abundancias (Magurran, 1988; Krebs, 1989; Ludwig y Reynolds, 1988; Lobo y Deloya, 1993; Fagan y Kareiva, 1997) o utilizando diversos estimadores no paramétricos que se basan en la información acumulada en diversas unidades muestrales (Magurran, 1988; Krebs, 1989; Palmer, 1990; Baltanas, 1992; Chao y Lee, 1992; Colwell y Coddington, 1995; Chazdon *et al.*, 1998).

Sin embargo, para la aplicación de todas estas metodologías se necesitan siempre datos con un mínimo de calidad. Datos que ofrezcan información sobre la abundancia de las especies o que estén asociados con alguna medida de esfuerzo claramente definida. Cuando procede, éstas metodologías son muy valiosas pero, evidentemente, nada pueden hacer cuando el problema fundamental es la ausencia de información. Si el grado de conocimiento taxonómico y biogeográfico de un territorio es precario, el esfuerzo aún por realizar puede ser de tales proporciones, que resulta imposible obtener a medio plazo una aproximación fiable del reparto espacial del número de especies. Aunque resulte descorazonador, no hay que descartar que esta situación sea corriente incluso en aquellos grupos bien estudiados y en países con una prolongada tradición sistemática. Veamos un ejemplo.

Los coleópteros coprófagos de la subfamilia Scarabaeinae son uno de los grupos de insectos mejor estudiados en Europa Occidental (ver, por ejemplo, Hanski y Cambefort, 1991; Baraud, 1992 y referencias allí citadas). Tanto en la Península Ibérica como en Francia, existen bancos de datos exhaustivos que compendian toda la información museística y bibliográfica disponible (Lumaret, 1990; Lobo y Martín-Piera, 1991). Utilizando como medida de esfuerzo la cantidad de registros contenidos en dichas bases de datos para cada cuadrícula UTM de 50 km, se realizó un análisis de la relación asintótica entre el número de especies y el esfuerzo de muestreo expresado como el número de registros informáticos (Fig. 1). Dicho análisis demostró que únicamente entre un 25% y 32% del total de las cuadrículas pueden considerarse

relativamente bien muestreadas (Lobo *et al.*, 1997; Lobo y Martín-Piera, datos no publicados). Es decir, recopilar durante años toda la información disponible desde hace más de un siglo, sobre un grupo faunística y taxonómicamente bien estudiado (la última nueva especie fue descrita hace más de 20 años), sólo permite tener inventarios fiables en apenas poco más de la cuarta parte del territorio. Considerando, además, que se trata de una escala espacial grosera (cuadrículas de unos 2.500 km² aproximadamente) y de un país europeo, ¿cuál puede ser entonces la situación en otras áreas más ricas y con menor información almacenada? ¿Qué podemos hacer entonces?

3. – BUSCANDO LA FUNCIÓN PREDICTIVA

En este caso, la única alternativa es realizar pronósticos razonables sobre la distribución del número de especies. La relación entre el número de taxones de alto rango y el número de especies (ver Martín-Piera en este mismo volumen), la búsqueda de especies indicadoras y la utilización de información ambiental de diverso tipo para obtener modelos con valor predictivo, son las tres principales aproximaciones seguidas para obtener estimas de la diversidad biológica cuando falta la información biológica (Gaston, 1996). En adelante me centraré en la última de estas aproximaciones, describiendo sus requerimientos y algunos fundamentos básicos sobre el Análisis de Regresión, la técnica estadística generalmente aplicada en esta metodología. Después repasaré los principales inconvenientes a tener en cuenta, enumerando algunas soluciones que pueden ser de utilidad a la hora de encontrar funciones predictivas de mayor capacidad. Ejemplos de esta metodología e información más detallada pueden encontrarse en Austin *et al.* (1984 y 1996), Buckland y Elston (1993), Heikkinen y Neuvonen (1997), Kirkpatrick y Brown (1994), Nicholls (1989) o Margules *et al.* (1987).

Son necesarios tres requisitos para buscar una función predictiva: i) una serie de unidades espaciales cuyo inventario faunístico sea fiable y bien establecido, ii) datos ambientales veraces para cada una de esas unidades y iii) una metodología capaz de generar una ecuación con las variables explicativas que permita predecir la variable dependiente (número de especies o presencia-ausencia en el caso de una única especie).

3.1. Necesidades

Como se mencionó en el apartado anterior, cuando existe información sobre la abundancia relativa de cada especie o cuando se delimita una unidad de esfuerzo de muestreo, pueden determinarse aquellas localidades bien muestreadas en las que el inventario faunístico o la presencia-ausencia de cada especie pueda considerarse incuestionable (ver Prendergast *et al.*, 1993; Soberón y Llorente, 1993; Colwell y Coddington, 1995; Colwell, 1997 o Fagan y Kareiva, 1997).

Respecto al segundo requisito, para reunir información ambiental de calidad suele ser necesario un esfuerzo considerable debido, generalmente, a la falta de fuentes y a su dispersión. Los mapas temáticos constituyen una posibilidad adecuada siempre y cuando nuestra escala espacial de análisis no sea muy detallada y los datos de la variable cartografiada no estén obsoletos. Para superponer la cartografía ambiental y la topográfica, de modo que pueda calcularse el valor promedio de la variable en la unidad espacial que fundamenta la utilización de los Sistemas de Información Geográfica (SIGs). Un SIG es un tipo especializado de base de datos capaz de manejar datos geográficos como imágenes

deseamos, es o mapas de la misma porción del territorio, de modo que es posible superponer y analizar las características espaciales y temáticas de una zona concreta (Miller, 1994). Numerosos estudios han tratado de poner de manifiesto la relación existente entre el número de especies de un grupo determinado y distintos parámetros ambientales (ver Gaston, 1996 y referencias allí citadas). Estos trabajos demuestran que el número de variables ambientales que pueden utilizarse para buscar una función predictiva es muy grande, siendo difícil decidir cuales son las más convenientes en cada caso concreto. Como criterio preliminar, parece conveniente seleccionar aquellas variables relacionadas con la energía que entra en el sistema como productividad, evapotranspiración, pluviosidad, temperatura, etc. (ver Wright et al., 1993), pero la amplitud altitudinal o alguna medida de la heterogeneidad ambiental suelen dar también buenos resultados (Wohlgemuth, 1998). Finalmente, si sospechamos que la influencia de la alteración antrópica del territorio puede ser una variable importante, resulta útil considerar alguna medida del uso del suelo o del tipo de paisaje.

Existen diversos métodos, unos más sofisticados que otros, para intentar predecir el número de especies o la presencia de una sola de ellas a partir de variables ambientales. Por ejemplo, los algoritmos genéticos (Mitchell, 1991; ver http://kaos.erin.gov.au/general/biodiv_model/ERIN/GARP/home.html), el análisis de huecos (*Gap Analysis*, Scott et al., 1993; ver <http://www.sdvc.uwyo.edu/wbn/gap.html>), los autómatas celulares (Carey, 1996) o los métodos de clasificación y regresión (*Regression Tree Analysis*; Iverson y Prasad, 1998). Sin embargo, el método más socorrido y sencillo para explorar la relación entre los datos de especies y diversas variables es el Análisis de Regresión (Ter Braak y Looman, 1995). En un Análisis de Regresión simple existe una variable respuesta o dependiente (y) que puede ser el número de especies, la abundancia o la presencia-ausencia de una sola especie y una variable explicativa o independiente (x). El propósito es obtener una función sencilla de la variable explicativa, que sea capaz de describir lo más ajustadamente posible la variación de la variable dependiente. Como los valores observados de la variable dependiente difieren generalmente de los que predice la función, ésta posee un error. La función más eficaz es aquella que describe la variable dependiente con el menor error posible o, dicho en otras palabras, con la menor diferencia entre los valores observados y predichos. La diferencia entre los valores observados y predichos (el error de la función) se denomina variación residual o residuos. Para estimar los parámetros de la función se utiliza el ajuste por mínimos cuadrados. Es decir, se trata de encontrar la función en la cual la suma de los cuadrados de las diferencias entre los valores observados y esperados sea menor. Sin embargo, con este tipo de estrategia es necesario que los residuos o errores estén distribuidos normalmente y que varíen de modo similar a lo largo de todo el rango de valores de la variable dependiente. Estas suposiciones pueden comprobarse examinando la distribución de los residuos y su relación con la variable dependiente.

Cuando la variable dependiente es cuantitativa (por ejemplo, el número de especies) y la relación entre ambas variables sigue una línea recta, la función es del tipo $y = c + bx$, en donde c es el intercepto o valor del punto de corte de la línea de regresión con el eje de la variable dependiente (una medida del número de especies existente cuando la variable ambiental tiene su mínimo valor) y b es la pendiente o coeficiente de regresión (la tasa de incremento del número de especies con cada unidad de la variable ambiental considerada). Si la relación no es lineal pueden transformarse los

valores de una o ambas variables para intentar linealizarla. Si no es posible convertir la relación en lineal, puede comprobarse el grado de ajuste de una función polinomial más compleja. La función polinomial más sencilla es la cuadrática ($y = c + bx + bx^2$) que describe una parábola, pero puede usarse una función cúbica u otra de un orden aun mayor capaz de conseguir un ajuste casi perfecto a los datos. Por último, puede acudir a los métodos de regresión curvilineales para observar el grado de ajuste de la función no lineal que, a priori, consideremos que define la relación entre las variables. Cuando la variable dependiente se expresa en datos cualitativos (presencia-ausencia de una especie) es aconsejable utilizar las regresiones logísticas ($y = [\exp(c + bx)] / [1 + \exp(c + bx)]$). Buenos ejemplos del uso de regresiones logísticas para predecir la distribución de una especie pueden encontrarse en Walker (1990), Osborne y Tigar (1992) y Parker (1999).

Si se poseen datos de varias variables ambientales explicativas deben utilizarse los Análisis de Regresión Múltiples que, en esencia, siguen los mismos principios que los Análisis de Regresión simples. En estos análisis suele ser frecuente que algunas variables independientes estén correlacionadas entre ellas, por lo que hay que conducirse cuidadosamente para realizar la elección de las variables significativas. Puede hacerse un Análisis de Regresión Múltiple con todas las variables a la vez y comprobar cuales tienen una pendiente significativamente diferente de cero, pero también pueden eliminarse aquellas variables no significativas mediante un proceso de eliminación hacia delante (*forward*) o hacia atrás (*backward*). En el primero se realiza una regresión individual sobre cada una de las variables independientes y se elige aquella que explica un mayor porcentaje de la variación en la variable dependiente. Entonces se vuelven a realizar regresiones parciales incluyendo la variable elegida y, una a una, se introduce el resto de las variables hasta encontrar aquella que explica un mayor porcentaje de la varianza que queda aún por explicar. El proceso finaliza cuando no puede encontrarse ninguna otra variable que contribuya significativamente. En el procedimiento de selección *backward* se comienza por incluir todas las variables a la vez, eliminándose en cada paso la variable explicativa menos significativa, hasta que sólo queden aquellas que contribuyen significativamente a explicar la variación de la variable dependiente. Un método más fiable es el Análisis de Regresión por Pasos (*Stepwise Multiple Regression*) en el cual se alterna un proceso de selección *forward* de las variables con una eliminación *backward*, de modo que cada vez que se incluye una nueva variable en el modelo, se comprueba la significación de todas las previamente incluidas. Desgraciadamente puede ocurrir que los distintos métodos elijan diferentes conjuntos de variables.

Los Análisis de Regresión por Pasos son cada vez más criticados (ver, por ejemplo, <http://www.fmdc.calpoly.edu/libarts/jgill/360.stepwise.html>), debido a que, con relativa frecuencia, son incapaces de producir una función en la cual las variables integrantes sean las de mayor poder predictivo. En general, los mayores inconvenientes se producen cuando existe un alto número de variables explicativas y cuando algunas de éstas están correlacionadas entre sí. En esos casos, el grado de correlación entre las variables predictivas, suele afectar a la frecuencia con que aparecen aquellas que son verdaderamente explicativas, pudiendo entrar a formar parte del modelo variables que no poseen ninguna significación biológica o estén pobremente relacionadas con la variable dependiente. Si tenemos N variables independientes, existen 2^N posibles ecuaciones y varias de ellas pueden ser capaces de representar los datos igualmente bien (Daniel y Wood, 1980). ¿Cuál debemos elegir? La solución óptima consiste en

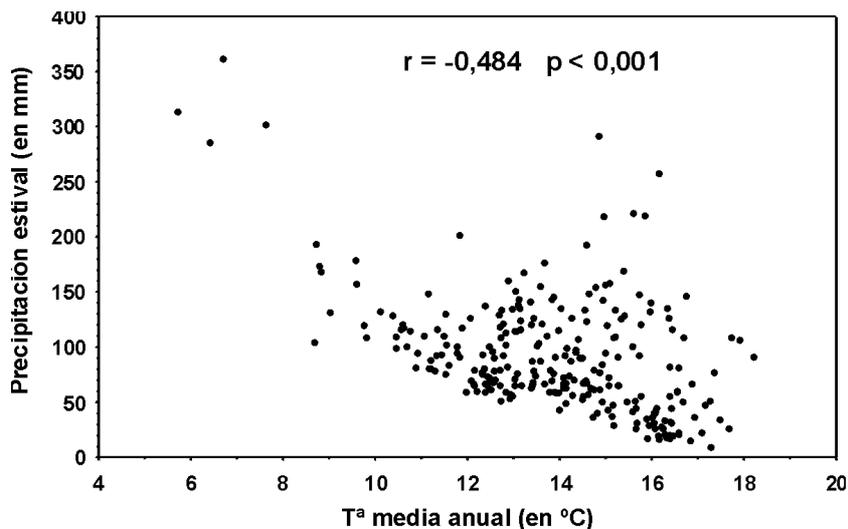


Fig. 2. Relación entre la temperatura media anual en las 255 cuadrículas UTM de 50 km de la Península Ibérica y la precipitación estival, un ejemplo de dos variables que no son independientes ya que una temperatura ambiental mayor suele asociarse con una sequía estival más pronunciada. En la figura se indica el coeficiente de correlación de Pearson entre las dos variables.

estudiar todas las ecuaciones posibles y elegir la que posea menores residuos. Esta aproximación, aunque infalible, puede ser inviable en la práctica con los métodos tradicionales: si poseemos 15 variables independientes, existen 32.728 ecuaciones diferentes. Cuando se trata únicamente de buscar una función que nos permita predecir no es trascendente que la función obtenida excluya variables significativas o incluya otras menos explicativas y difíciles de interpretar biológicamente. Podemos identificar los factores correlacionados con la diversidad de especies de un grupo cualquiera a una determinada escala espacial, pero resulta más difícil decidir si esa correlación es accidental o si puede identificarse una causa biológica subyacente (Shipley, 1999). La relación causal establecida con los análisis de regresión es hipotética y sus posibilidades de actuación deberán evaluarse posteriormente acumulando evidencias teóricas y empíricas. Examinando los argumentos teóricos a favor y en contra de su influencia, realizando manipulaciones experimentales específicas, explorando las correlaciones parciales entre cada variable explicativa y la variable dependiente (ver un ejemplo en Legendre y Legendre, 1983), o realizando diversos análisis estadísticos más específicos que requieren la utilización de pocas variables e hipótesis causales previas (*Path Analysis*; Legendre y Legendre, 1998).

3.2. Los inoportunos inconvenientes

La búsqueda de una relación sencilla que pueda ofrecernos una función de alto valor predictivo es, con toda seguridad, una pretensión exagerada e irreal. Estimar el número de especies (o la presencia de una sola de ellas) en un territorio concreto y a una escala espacial determinada, mediante la utilización de variables ambientales, es una aproximación más simple que coleccionar, identificar y contar directamente el número de especies, pero no está exenta de dificultades. Examinaré las que considero fundamentales.

Como se mencionó anteriormente, un problema frecuente es la correlación existente entre las distintas variables ambientales que se utilizan (Fig. 2), un atributo propio de los sistemas que, como la Biosfera, poseen una notable complejidad en las relaciones entre sus elementos. Muchas variables ambientales varían de forma similar a lo largo de un gradiente espacial, de modo que se pierde la necesaria independencia entre sus valores. Elegir una variable explicativa suele suponer desestimar otras relacionadas con ella, pero que son ligeramente menos efectivas a la hora de predecir la variable

dependiente. Por ello, pueden encontrarse distintas funciones con similar capacidad predictiva constituidas por conjuntos de variables ambientales diferentes. Cualquiera de esas funciones puede ser útil a la hora de realizar pronósticos, pero siempre hay que tener cuidado con las interpretaciones causales, es decir, con la consideración de esas variables como los factores que originan la modificación en la variable estudiada. Examinar previamente el grado de correlación existente entre las variables explicativas no soluciona completamente el problema, pero permite desechar algunas variables ambientales y conocer las razones de su elección o rechazo. Una alternativa consiste en utilizar algún método multivariante de ordenación, como el análisis de componentes principales (ver Legendre y Legendre, 1998). Con estos métodos puede reducirse el número de variables mediante la generación de otras nuevas que son composiciones de las anteriores y, lo que es más importante, que no están correlacionadas entre sí. El inconveniente es que, en no pocas ocasiones, es difícil interpretar estas nuevas variables.

Otro inconveniente habitual es que, a veces, el efecto de una variable depende del valor de otra y esa *interacción* posee un mayor valor predictivo que las propias variables por separado (Fig. 3; ver Margules *et al.*, 1987). Para comprobar si las interacciones son efectivas a la hora de predecir la variable dependiente, se utiliza el producto de dos variables ($x_3 = x_1 * x_2$) como una variable más, comprobándose su capacidad predictiva mediante su inclusión en el Análisis de Regresión. Un tipo de interacciones especialmente interesantes, son aquellas que existen entre una variable espacial (latitud, longitud o algún término cuadrático o cúbico de éstas) y una variable ambiental. En el caso hipotético de que una variable de interacción de este tipo posea valor predictivo, es probable que la variable ambiental en cuestión tenga un efecto significativo únicamente en determinadas porciones del territorio.

La relación entre las variables ambientales y la variable dependiente no tiene por qué ser lineal (Fig. 4; ver Austin *et al.*, 1996). Es más, hay numerosos ejemplos de relaciones que no son lineales. Tal es el caso de las que se dan entre productividad y diversidad (Rosenzweig y Abramsky, 1993) o entre diversidad y altitud (McCoy, 1990): parabólicas, exponenciales, potenciales, etc. Las relaciones deben de ser lineales para la mayoría de los métodos de regresión y estadísticos sencillos. Por ello, es conveniente visualizar una a una las relaciones entre las variables explicativas y la variable dependiente y en el caso de que no sean lineales intentar linealizarlas

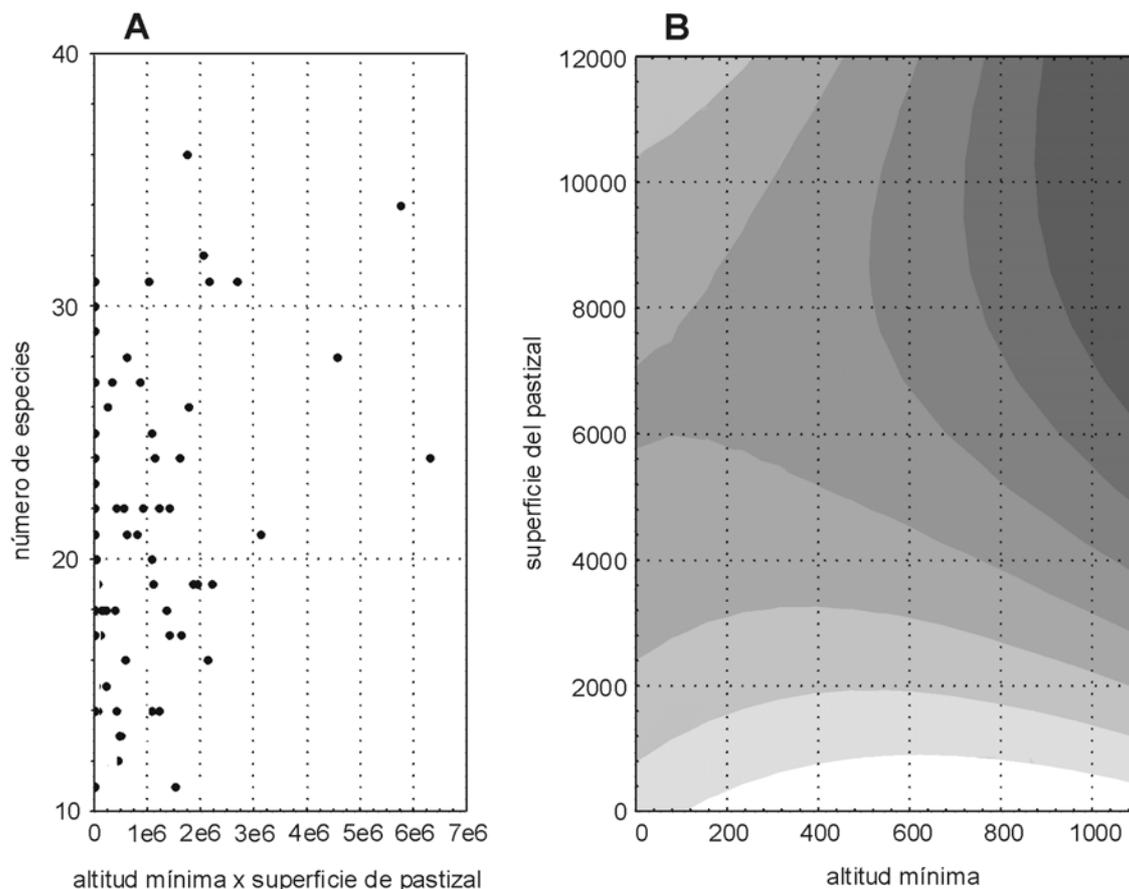


Fig. 3. Relación entre el número de especies de Escarabedios y la interacción entre la altitud mínima y la superficie de pastizal (A). Cuando se introducen ambas variables y su producto en un análisis de regresión por pasos, únicamente el término de interacción resulta estadísticamente significativo ($b=0,337$, $t_{(80)}=3,20$, $p=0,002$). Como puede observarse en el gráfico de contorno (B), las localidades con una altitud mínima muy baja o con poca superficie de pastizales albergan pocas especies (zonas más claras), como también ocurre en las localidades con una gran superficie de pastizales y altitudes bajas o en las localidades con altitudes mínimas elevadas pero poca superficie de pastizal. Sin embargo, cuando el valor de ambas variables es grande su efecto sinérgico produce un aumento notable de la riqueza de especies. Lo que en realidad demuestra esta interacción es que los pastizales de las zonas montañosas ibéricas son especialmente ricos en especies de Escarabedios.

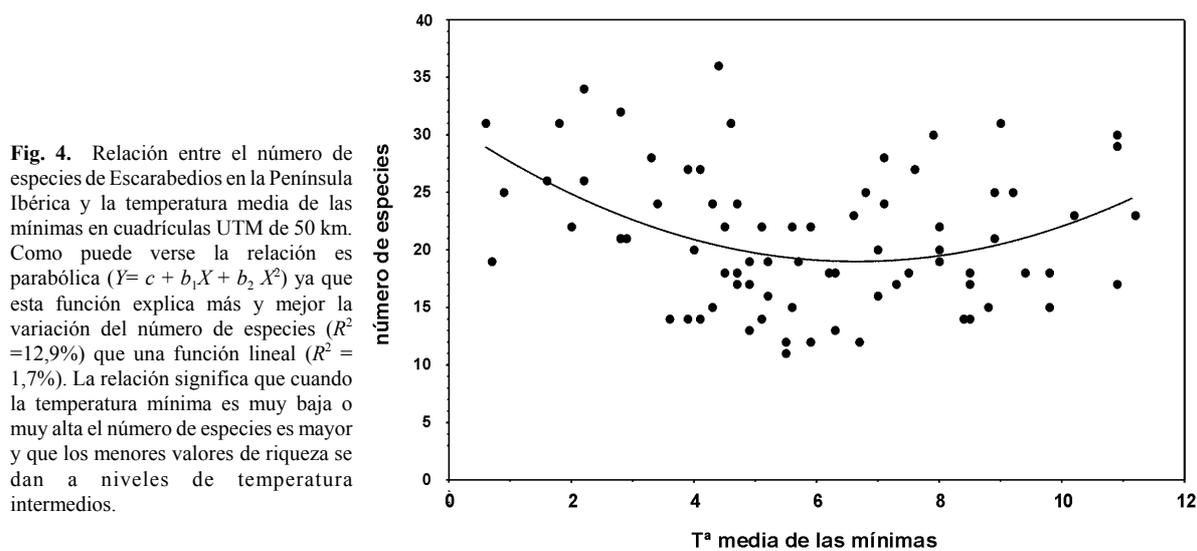


Fig. 4. Relación entre el número de especies de Escarabedios en la Península Ibérica y la temperatura media de las mínimas en cuadrículas UTM de 50 km. Como puede verse la relación es parabólica ($Y= c + b_1X + b_2 X^2$) ya que esta función explica más y mejor la variación del número de especies ($R^2 = 12,9\%$) que una función lineal ($R^2 = 1,7\%$). La relación significa que cuando la temperatura mínima es muy baja o muy alta el número de especies es mayor y que los menores valores de riqueza se dan a niveles de temperatura intermedios.

mediante algún tipo de transformación.

Tabla I

VARIABLES explicativas utilizadas para elaborar un modelo predictivo de la distribución del número de especies de Escarabeidos en la Península Ibérica. Los datos de cada variable se tomaron para cada cuadrícula UTM de 50 km utilizando un modelo digital del terreno, mapas de usos del suelo y mapas climáticos. Para obtener los datos promedios en cada cuadrícula, dicha cartografía se superpuso con un mapa digital de las cuadrículas mediante un Sistema de Información Geográfica. La diversidad de usos del suelo y la diversidad geológica se calcularon mediante el índice de diversidad de Shannon utilizando los valores de todas las categorías de usos del suelo (n=44) y geológicas (n=3). Mediante Modelos Lineales Generalizados se averiguó la significación de las funciones lineales ($y=c+bx$), cuadráticas ($y=c+bx+bx^2$) o cúbicas ($y=c+bx+bx^2+bx^3$) de cada variable a la hora de explicar el número de especies en las 82 cuadrículas UTM bien muestreadas, mostrándose la función elegida en cada caso.

Variables	Función seleccionada
Espaciales	
Latitud	Cubica
Longitud	Lineal
Geográficas	
Superficie terrestre en cada cuadrícula	Lineal
Distancia a los Pirineos	Lineal
Topográficas	
Altitud mínima	Lineal
Altitud máxima	Lineal
Altitud media	Lineal
Diferencia de altitudes	Lineal
Geológicas	
Superficie caliza	Lineal
Superficie silicea	Lineal
Diversidad geológica	Lineal
Climáticas	
Temperatura media de las mínimas	Cuadrática
Temperatura media de las máximas	Lineal
Temperatura media anual	Cuadrática
Variación anual de temperatura	Lineal
Precipitación anual	Cúbica
Precipitación estival	Lineal
Variación anual de precipitaciones	Lineal
Días con sol anuales	Lineal
Usos del suelo	
Superficie cultivada y urbana	Lineal
Superficie forestal	Lineal
Superficie con matorral	Lineal
Superficie con pastizal	Cúbica
Diversidad de usos del suelo	Cúbica

Tabla II

Resultados del análisis de regresión múltiple por pasos entre el número de especies de Escarabaeidos en las cuadrículas Ibéricas bien muestreadas (Fig. 1) y las variables ambientales y espaciales elegidas (Tabla I). Se muestran las variables significativas, sus coeficientes y el error en la estima de los mismos.

	coeficientes	error estandard
Altitud máxima	0,062	0,026
+ Superficie de pastizal ²	0,014	0,006
+Diversidad de usos del suelo	0,114	0,041
+Superficie forestal x Diversidad geológica	-0,079	0,025
+Altitud máxima x Area terrestre	0,093	0,026
+Superficie caliza x Diversidad geológica	0,054	0,029
+ Latitud ²	-0,021	0,009
+ Latitud ³	-0,009	0,002
<i>intercepto</i>	2,940	0,050

Como generalmente las variables biológicas no se distribuyen siguiendo la curva Normal, ni las relaciones entre ellas suelen lineales, se está imponiendo cada vez con mayor frecuencia el uso de los Modelos Lineales Generalizados (McCullagh y Nelder, 1983; Crawley, 1993). La utilización de esta técnica estadística constituye una herramienta fundamental para solventar los inconvenientes anteriores, ya que permite especificar el tipo de distribución para los errores y elaborar modelos aunque las relaciones entre las variables no sean lineales. Además, facilita el análisis de las interacciones, siendo utilizable, asimismo, en el caso de variables categóricas (ver algunos ejemplos en Austin *et al.*, 1996 o Heikkinen y Neuvonen, 1997).

Las variables ambientales están a menudo espacialmente autocorrelacionadas, o sea, el valor de la variable en una localidad depende del valor en las localidades contiguas, violándose así la independencia entre las observaciones requerida por la mayoría de las pruebas estadísticas (Cliff y Ord, 1981; Odland, 1988). La autocorrelación de una variable puede suprimirse con bastante efectividad mediante un *Trend Surface Analysis* (TSA; ver Legendre, 1993). Se trata de construir un polinomio de tercer grado con la latitud (LAT) y la longitud (LON) de cada observación ($LAT + LON + LAT^2 + LAT \times LON + LON^2 + LAT^3 + LAT^2 \times LON + LAT \times LON^2 + LON^3$) y someter sus nueve términos como variables independientes a un análisis de regresión por pasos, con la variable a la cual queremos quitar su estructura espacial como dependiente. De este modo, si aparece algún término significativo, la variable posee una estructura espacial y pueden elegirse los residuos de la regresión (la parte de la variable no explicada por el espacio), como una estima de la variable sin autocorrelación o estructura espacial. Como casi todas las variables ambientales varían siguiendo algún gradiente espacial, sobre todo a escala geográfica, eliminar el componente espacial de una variable supone casi siempre desechar una parte importante de su capacidad explicativa. De este modo, es necesario: i) comprobar que las variables que utilizamos están autocorrelacionadas, ii) incorporar sin problema esa autocorrelación en nuestros modelos predictivos y iii) examinar los residuos del modelo para ver si estos están autocorrelacionados (Cliff y Ord, 1981; Legendre y Legendre, 1998; Smith, 1994). Si los residuos (lo que no predice nuestro modelo) poseen una clara estructura espacial o están autocorrelacionados, es que probablemente no hemos considerado alguna variable con una clara variación espacial. Será entonces necesario buscar esa variable o variables (elaborar un mapa de esos residuos puede dar algunas pistas) e introducirla para lograr que la capacidad predictiva de nuestro modelo no sea significativamente mayor en unos lugares determinados que en otros. Una manera de asegurar que la variación espacial de nuestra variable dependiente esté recogida, disminuyendo de este modo las posibilidades de que aparezcan residuos autocorrelacionados, es introducir al final del modelo y como variables independientes, los nueve términos del polinomio de tercer grado de la latitud y la longitud. Para realizar los análisis de autocorrelación existen algunos programas informáticos de acceso gratuito (por ejemplo: R - Package, <http://www.fas.umontreal.ca/biol/casgrain/en/labo/R/index.html> ó RookCase, <http://aix1.uottawa.ca/academic/arts/geographie/lpcweb/sections1/software/frmssoft.htm>).

3.3. Un ejemplo

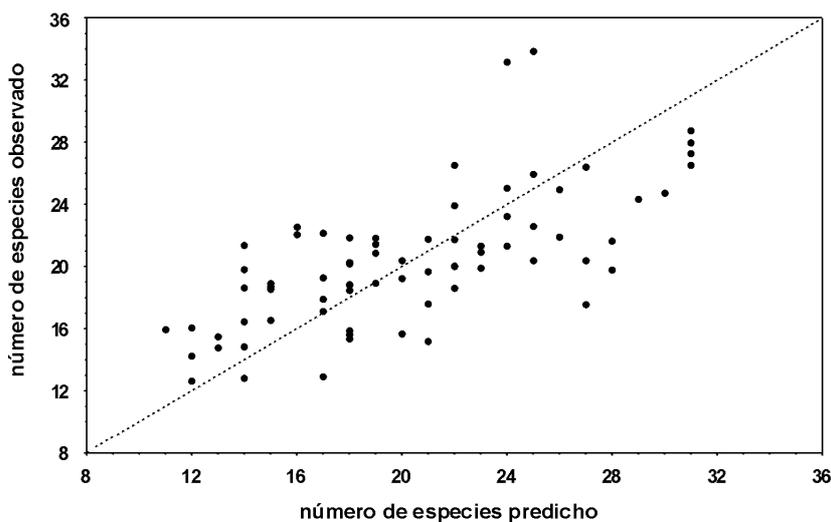
Utilizando el número de registros contenidos en una exhaustiva base de datos sobre los Scarabaeidae Ibéricos como medida

de esfuerzo, hemos predicho la distribución geográfica de la riqueza de especies en la Península Ibérica, mediante la utilización de variables ambientales. En primer lugar, se evaluó el esfuerzo de muestreo en toda la península a escala de cuadrículas UTM de 50 km. Para ello, se definieron primeramente las regiones fisioclimáticas del territorio peninsular utilizando variables climáticas (temperatura media anual, variación de la temperatura anual, días de sol al año, precipitación total anual, variación de la precipitación estival y anual) y topográficas (altitud media y rango altitudinal). A continuación, se analizó el esfuerzo de muestreo diferencial entre áreas, investigando la relación asintótica del número de especies registradas a diferentes niveles de esfuerzo de muestreo. Posteriormente, se estableció el grupo de cuadrículas UTM bien muestreadas en cada una de las subregiones, mediante una jerarquía de esfuerzos de muestreo subjetivamente definidos a diferentes tasas de incremento del número de especies. Se identificaron un total de 82 cuadrículas (32,2%) adecuadamente muestreadas (Fig. 1). Como variables explicativas se utilizaron los valores de 24 variables en esas 82 cuadrículas (Tabla I): dos espaciales, dos geográficas, cuatro topográficas, tres geológicas, ocho climáticas y cinco relativas a usos del suelo, más los correspondientes términos de interacción entre las diferentes variables ambientales ($n=231$) y entre éstas y las espaciales ($n=44$).

Para elegir el modelo predictivo se examinó primeramente la significación de cada una de las variables ambientales, comprobando si su relación con el número de especies era una función lineal ($y = x$), cuadrática ($y = x + x^2$) o cúbica ($y = x + x^2 + x^3$). La función de la variable ambiental con una mayor capacidad explicativa ingresó en primer lugar en un análisis de regresión por pasos. Una vez incluida esa primera variable, se introdujeron secuencialmente cada una de las restantes variables, eligiendo aquella que mejoraba más el modelo. Una vez incluidas todas las variables ambientales significativas, se examinaron todos los posibles pares de interacciones entre variables. Cada interacción era añadida al modelo hasta entonces resultante y una vez elegida la interacción más significativa, el proceso se repetía hasta que no se encontraba ninguna otra interacción significativa. Por último se añadieron los nueve términos del polinomio de tercer grado de la latitud y la longitud. Tras cada inclusión de una variable significativa (*forward*), se comprobó si cualquiera de las anteriormente introducidas dejaba de ser significativa (*backward*).

La elaboración completa de un primer modelo final permitió comprobar que algunas cuadrículas poseían valores reales muy superiores a los que predecía la función, mientras que otras poseían valores reales inferiores a los predichos ('*outliers*'). Se trataba de cuatro cuadrículas ubicadas en los alrededores de Madrid y Barcelona con una gran cantidad de información faunística acumulada y de las tres cuadrículas ubicadas en las Islas Baleares. Ya que utilizar el número de especies observado en estas cuadrículas para predecir la riqueza del resto, podría producir resultados erróneos, se elaboró un nuevo modelo tras eliminar estas 7 cuadrículas. El modelo final permite explicar el 62,4% de la varianza en el número de especies y contiene ocho variables (Tabla II): altitud máxima, el término cuadrático de la superficie de pastizal, la diversidad de usos del suelo, las interacciones superficie forestal x diversidad geológica, área terrestre x altitud máxima, superficie caliza x diversidad geológica y los términos cuadrático y cúbico de la latitud. Un examen de los residuos del modelo mostró la ausencia de características aberrantes y autocorrelación.

Fig. 6. Validación del modelo predictivo final (Tabla II). Se trata de la relación entre el número de especies de Escarabeidos observado en las cuadrículas UTM de 50 Km Ibéricas y el número de especies predicho (Coeficiente de correlación de Pearson; $r = 0,689$; $p < 0,001$). En este caso los valores predichos se obtuvieron eliminado una a una cada observación y calculando cada vez la función predictiva según el modelo final para esa observación. Se realizaron así 75 análisis de regresión uno para cada observación. La alta correlación significa que el modelo es fiable a la hora de predecir el número de especies en una cuadrícula no muestreada.



una especie, pero que no está habitado por ella (por ejemplo, cada región mediterránea tiene riquezas y especies distintas; Ver Ricklefs y Schluter, 1993). Cuanto mayor es la escala espacial de análisis, más frecuente es este fenómeno debido, lógicamente, a la actuación de factores de carácter único e irreplicable que condicionan la distribución. Se trata de la historia evolutiva propia de cada grupo, de la historia y los avatares propios de cada región y de las características geográficas de la misma. Evidentemente, la importancia de los factores ambientales a la hora de restringir la distribución de las especies es limitada, de modo que elaborar funciones predictivas que únicamente tengan en consideración variables ambientales producirá, generalmente, modelos incompletos. Ello será especialmente cierto cuando la escala espacial de análisis sea amplia, como exige el estudio geográfico de la variación del número de especies. Historia y geografía también juegan y es necesario incluirlas ¿Cómo podemos hacerlo?

Si los factores históricos y geográficos ejercen una manifiesta influencia sobre la distribución de las especies y si el efecto de estos factores tiene una expresión espacial, la inclusión de funciones complejas de la latitud y la longitud, puede mejorar la capacidad de predicción de las funciones obtenidas. Veamos un ejemplo. En la Figura 7 se muestra la distribución hipotética de una variable ambiental que decrece en sus valores hacia el sur y hacia el oeste y, a la vez, una distribución hipotética de la riqueza en especies de un taxón que es función de dicha variable ambiental (el número de especies es la raíz cuadrada del valor de la variable ambiental). Supongamos que por causas históricas o geográficas no asociadas a factores ambientales, el número de especies disminuye en el cuadrante noroccidental sombreado (Fig. 7A). En este caso, elaborar una función predictiva mediante un análisis de regresión, utilizando sólo la variable ambiental permite explicar el 74,6% de la variación en el número de especies. Sin embargo, si se incluye esa misma variable ambiental junto a los 9 términos del polinomio de tercer grado de la latitud y la longitud, la función predictiva permite explicar el 92,1% de la variación en el número de especies (ver Fig. 7). La longitud y la interacción entre la latitud y la longitud² resultan ser los términos espaciales que mejoran la capacidad explicativa del modelo.

Los factores de carácter único e irreplicable pueden ejercer una profunda influencia sobre la diversidad biológica, sobre todo en aquellos territorios en los que existe una yuxtaposición de taxones con distintas historias evolutivas.

Incorporar la componente espacial en estos casos, puede permitir elaborar funciones predictivas más fiables y competentes ya que, generalmente, el efecto de estos factores suele estar delimitado geográficamente. Sin embargo, en el caso de que el espacio posea una alta capacidad explicativa, no se debe atribuir ésta únicamente a factores de tipo histórico o geográfico. En la práctica, resulta difícil discernir si la capacidad explicativa del espacio se debe a factores de este tipo o a alguna otra variable ambiental no considerada (ver Legendre, 1990 y Borcard *et al.*, 1992). Sin embargo, si utilizando diversos tipos de variables ambientales, el espacio todavía posee una apreciable capacidad explicativa es lícito sospechar que ésta se deba a factores histórico-geográficos.

5.- COMENTARIO FINAL

Conocer la distribución geográfica del número de especies de un grupo taxonómico cualquiera es una tarea prácticamente inabordable a corto plazo, incluso en aquellos países con una dilatada tradición taxonómica. Esto es especialmente cierto cuando tratamos con algún grupo de insectos y con escalas espaciales geográficas. Ante esta situación, una estrategia recomendable consiste en recopilar la información taxonómica y corológica dispersa en las colecciones y la bibliografía, con el fin de delimitar los territorios insuficientemente muestreados y aquéllos en los que el inventario faunístico puede considerarse bien establecido. Cuando, como resultado de este análisis, la información sobre la distribución de las especies es insuficiente, deberá incrementarse el esfuerzo de inventario pero, en este caso, tendremos una valiosa información sobre la ubicación de las zonas a prospectar y sobre el esfuerzo suplementario a realizar. Si, por el contrario, la información biológica es suficiente, pueden aplicarse distintas aproximaciones metodológicas que, dentro de unos límites razonables, son capaces de proporcionar una estima de la diversidad existente en cada unidad espacial analizada, mediante la utilización de diversas variables ambientales. La distribución de la diversidad generada mediante este proceso debe considerarse una primera aproximación. Aproximación que puede y debe validarse y mejorarse paulatinamente mediante la realización de posteriores colectas en distintos territorios.

AGRADECIMIENTO

Trabajo realizado gracias a una beca postdoctoral de la CAM y a la financiación del proyecto PB97-1149 de la DGICYT.

A														B																										
0	0	0	1	2	4	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92	96	100
0	0	0	1	2	4	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	83	87	91	95	99
0	0	0	1	2	4	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82	86	90	94	98
0	0	0	1	3	6	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77	81	85	89	93	97
0	0	0	1	3	6	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92	96
0	0	0	1	3	6	7	7	7	8	8	8	8	8	9	9	9	9	10	10	10	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	83	87	91	95
0	0	1	1	3	6	6	7	7	8	8	8	8	8	9	9	9	9	10	10	10	18	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82	86	90	94
0	0	1	1	3	6	6	7	7	8	8	8	8	8	9	9	9	9	10	10	10	17	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77	81	85	89	93
0	1	2	2	3	6	6	7	7	7	8	8	8	8	9	9	9	9	10	10	10	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88	92
1	2	5	5	6	6	6	7	7	7	8	8	8	8	9	9	9	9	10	10	10	15	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	83	87	91
1	3	5	5	5	6	6	6	7	7	7	8	8	8	8	9	9	9	9	9	9	14	18	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82	86	90
2	3	5	5	5	6	6	6	7	7	7	8	8	8	8	9	9	9	9	9	9	13	17	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77	81	85	89
3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	9	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84	88
3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	9	11	15	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	83	87
3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	9	10	14	18	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82	86
3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	9	9	13	17	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77	81	85
3	3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80	84
3	3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	7	11	15	19	23	27	31	35	39	43	47	51	55	59	63	67	71	75	79	83
2	3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	6	10	14	18	22	26	30	34	38	42	46	50	54	58	62	66	70	74	78	82
2	3	4	4	5	5	6	6	6	7	7	7	7	8	8	8	8	9	9	9	9	5	9	13	17	21	25	29	33	37	41	45	49	53	57	61	65	69	73	77	81

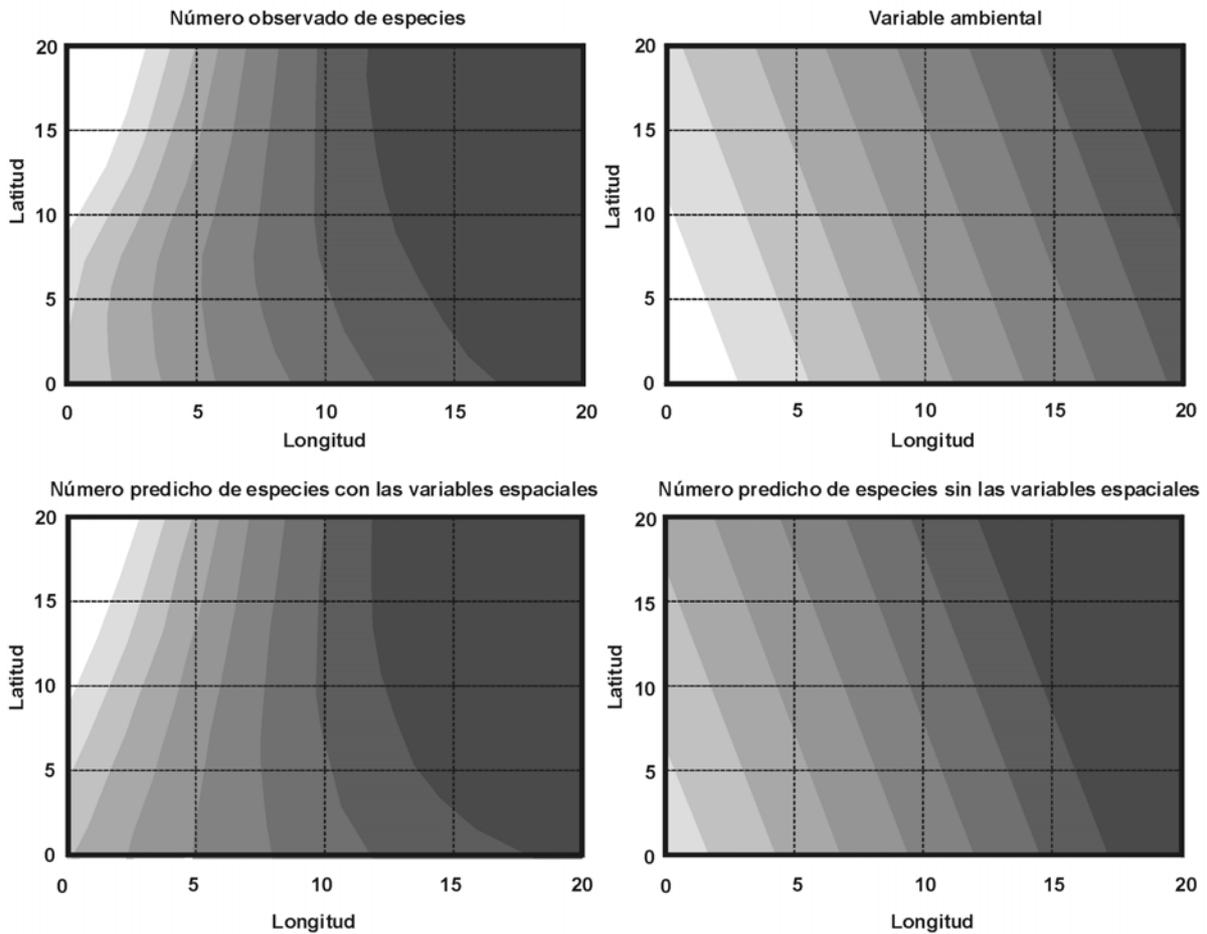


Fig. 7. (A) Distribución hipotética del número de especies. Dicho número decrece de manera similar a como lo hace la variable ambiental, cuyos valores se muestran en la matriz de la derecha. En realidad, el número de especies es la raíz cuadrada de la variable ambiental, excepto en las 54 cuadrículas sombreadas del cuadrante noroccidental, en las que la tendencia decreciente es mucho más acusada. (B) Variación de una variable ambiental hipotética que disminuye geoméricamente una unidad de norte a sur y dos unidades de este a oeste. Los cuatro gráficos de contorno expuestos reflejan la distribución del número de especies y de la variable ambiental (parte superior) y la distribución predicha según dos funciones diferentes: una que incorpora únicamente la información ambiental (parte inferior derecha) y otra que incorpora la información ambiental y espacial parte inferior izquierda. La gama de sombreados corresponde a incrementos del número de especies del blanco al gris oscuro.

BIBLIOGRAFÍA

- AHLBERG, P.E., CLACK, J.F. Y LUKSEVICS, E. 1996. Rapid braincase evolution between *Panderichthys* and the earliest tetrapods. *Nature*, **381**: 61-64.
- AUSTIN, M.P., CUNNINGHAM, R.B. Y FLEMING, P.M. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. *Vegetatio*, **55**: 11-27.
- AUSTIN, M.P., PAUSAS, J.G. Y NICHOLLS, A.O. 1996. Patterns of tree species richness in relation to environment in south-eastern New South Wales, Australia. *Australian Journal of Ecology*, **21**: 154-164.
- AVERS, C.J. 1989. *Process and Pattern in Evolution*. Oxford University Press, New York.
- BARAUD, J. 1992. *Coléoptères Scarabaeoidea d'Europe*. Fédération française des Sociétés de Sciences Naturelles et Société Linnéenne de Lyon, Lyon.
- BEAUFORT, F. DE Y MAURIN, H. 1988. *Le Secrétariat de la Faune et de la Flore et l'inventaire du patrimoine naturel: objectifs, méthodes et fonctionnement*. Secrétariat Faune-flore/MNHN, Paris.
- Baltanás, A. 1992. On the use of some methods for the estimation of species richness. *Oikos*, **65**: 484-492.
- BEERLING, D.J. Y WOODWARD, F.I. 1996. Palaeo-ecological perspectives on plant responses to global change. *Trends in Ecology and Evolution*, **11**: 20-23.
- BLONDEL, J. 1995. *Biogéographie. Approche écologique et évolutive*. Masson, Paris.
- BORCARD, D., LEGENDRE, P. Y DRAPEAU, P. 1992. Partialling out the spatial component of ecological variation. *Ecology*, **73**: 1045-1055.
- BUCKLAND, S.T. Y ELSTON, D.A. 1993. Empirical models for the spatial distribution of wildlife. *Journal of Applied Ecology*, **30**: 478-495.
- CAREY, P.D. 1996. A cellular automaton for predicting the distribution of species in a changed climate. *Global Ecology and Biogeography Letters*, **5**: 217-226.
- CHAO, A. Y LEE, S.M. 1992. Estimating the number of classes via sample coverage. *Journal of American Statistics Association*, **87**: 210-217.
- CHAZDON, R. L., COLWELL, R. K., DENSLow, J. S. Y. GUARIGUATA, M.R. 1998. Statistical methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica. En F. DALLMEIER Y J. A. COMISKEY (eds.), págs. 285-309. *Forest biodiversity research, monitoring and modeling: Conceptual background and Old World case studies*. Parthenon Publishing, Paris.
- CLIFF, A.D. Y ORD, J.K. 1981. *Spatial Processes.: Models and Applications*. Pion, London.
- COLWELL, R. K. 1997. EstimateS. Statistical estimation of species richness and shared species from samples. Version 5. User's guide and application published at <http://viceroy.eeb.uconn.edu/estimates>.
- COLWELL, R. K. Y CODDINGTON, J.A. 1995. Estimating terrestrial biodiversity through extrapolation. En HAWKSWORTH, D. L. (ed.), págs. 101-118. *Biodiversity, Measurement and Estimation*. Chapman & Hall, London.
- COOPE, G.R. 1979. Late Cenozoic fossil Coleoptera: Evolution, biogeography and ecology. *Annual Reviews of Ecology and Systematics*, **10**: 247-267.
- CRAWLEY, M.J. 1993. *GLIM for Ecologists*. Blackwell Science Ltd., Oxford.
- DANIEL, C. Y WOOD, F.S. 1980. *Fitting Equations to Data*. John Wiley & Sons, New York.
- DAVIS, F.W. 1994. Mapping and monitoring terrestrial biodiversity using geographic information systems. En PENG, C.I. Y CHOU, C.H. (eds.), págs. 461-471. *Biodiversity and Terrestrial Ecosystems*. Institute of Botany, Academia Sinica Monograph series No 14, Taipei.
- ELIAS, S.A. 1994. *Quaternary Insects and their Environments*. Smithsonian Institution Press, Washington.
- FAGAN, W.F. Y KAREIVA, P.M.. 1997. Using compiled species list to make biodiversity comparisons among regions: a test case using Oregon butterflies. *Biological Conservation*, **80**: 249-259.
- FARRELL, B.D., MITTER, C. Y FUTUYMA, D.J. 1992. Diversification at the insect-plant interface. *BioScience*, **42**: 34-42.
- GASTON, K.J. 1996. Species richness: measure and measurement. En GASTON, K.J. (ed.), págs. 77-113. *Biodiversity. A biology of numbers and difference*. Blackwell Science Ltd., Oxford.
- HANSKI, I. Y CAMBEFORT Y. (eds.). 1991. *Dung Beetle Ecology*. Princeton University Press, New Jersey.
- HARDING, P.T. Y SHEAIL, J. 1992. The Biological Records Centre: a pioneer in data gathering and retrieval. En HARDING, P.T. (ed.), págs. 5-19. *Biological Recording of Changes in British Wildlife*. Institute of Terrestrial Ecology, London.
- HEIKKINEN, R.K. Y NEUVONEN, S. 1997. Species richness of vascular plants in the subarctic landscape of northern Finland: modelling relationships to the environment. *Biodiversity and Conservation*, **6**: 1181-1201.
- HENGEVELD, R. 1990. *Dynamic Biogeography*. Cambridge University Press, Cambridge.
- HENGEVELD, R. 1997. Impact of Biogeography on a population-biological paradigm shift. *Journal of Biogeography*, **24**: 541-547.
- HULBERT, S.H. 1971. The non-concept of species diversity: A critique and alternative parameters. *Ecology*, **52**: 577-586.
- HUNTLEY, B. Y BIRKS, H.J.B. 1983. *An Atlas of Past and Present Pollen Maps for Europe: 0-13,000 Years Ago*. Cambridge University Press, Cambridge.
- IVERSON, L.R. Y PRASAD, A.M. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs*, **68**: 465-485.
- KIRKPATRICK, J.B. Y BROWN, M.J. 1994. A comparison of direct and environmental domain approaches to planning reservation of forest higher plant communities and species in Tasmania. *Conservation Biology*, **8**: 217-224.
- KREBS, C.J. 1989. *Ecological Methodology*. Harper and Row, New York.
- LATHAM, R.E. Y RICKLEFS, R.E. 1993. Continental comparisons of temperate-zone tree species diversity. En R.E. RICKLEFS Y SCHLUTER, D. (eds.), págs. 294-314. *Species Diversity in Ecological Communities*. University of Chicago Press, Chicago.
- LEGENDRE, P. 1990. Quantitative methods and biogeographic analysis. En GARBARY, D.J. Y SOUTH, R.R. (eds.), págs. 9-34. *Evolutionary Biogeography of the Marine Algae of the North Atlantic*. Nato asi series, Vol G 22, Springer-Verlag, Berlin.
- LEGENDRE, L. Y LEGENDRE, P. 1983. *Numerical Ecology*. Elsevier Scientific Publishing, Amsterdam.
- LEGENDRE, P. Y LEGENDRE, L. 1998. *Numerical Ecology* (2ª edición). Elsevier Scientific Publishing, Amsterdam.
- LOBO, J.M. 1999. Individualismo y adaptación Espacial: un nuevo enfoque para explicar la distribución geográfica de las especies. *Bol. SEA (Evolución y Filogenia de Arthropoda)*, n° **26**: 561-572.
- LOBO, J.M. Y DELOYA, C. 1993. Una predicción acerca de la diversidad de escarabeidos (Coleoptera, Scarabaeidae) en la Reserva de La Biosfera "La Michilia", Durango, México. *Boletín de la Sociedad Veracruzana Zoología*, **3**: 57-63.
- LOBO, J. M. Y MARTÍN-PIERA, F. 1991. La creación de un banco de datos zoológico sobre los Scarabaeidae (Coleoptera: Scarabaeoidea) ibero-baleares: una experiencia piloto. *Elytron*, **5**: 31-38.
- LOBO, J.M., LUMARET, J.P. Y JAY-ROBERT, P. 1997. Les atlas faunistiques comme outils d'analyse spatiale de la biodiversité. *Annales de la Société Entomologique de France (N.S.)*, **33**: 129-138.
- LUBCHENCO, J., OLSON, A.M., BRUBAKER, L.B., CARPENTER, S.B., HOLLAND, M.M., HUBBELL, S.P., LEVIN, S.A., MACMAHON, J.A., MATSON, P.A., MELILLO, J.M., MOONEY, H.A., PETERSON, C.H., PULLIAM, H.R., REAL, L.A., REGAL, P.J. Y RISSER, P.G. 1991. The sustainable biosphere initiative: an ecological research agenda. *Ecology*, **72**: 371-412.
- LUDWIG, J.A. Y REYNOLDS, J.F. 1988. *Statistical Ecology*. John Wiley & Sons, New York.
- LUMARET, J.P. 1990. *Atlas des Coleopteres Scarabéides Laparosticti*

- de France. Museum National d'Histoire Naturelle, Secretariat de la Faune et de la Flore, Paris.
- MAGURRAN, A. E. 1988. *Ecological Diversity and its Measurement*. Princeton University Press, New Jersey.
- MARGULES, C.R., NICHOLLS, A.O. Y AUSTIN, M.P. 1987. Diversity of *Eucalyptus* species predicted by a multi-variable environment gradient. *Oecologia*, **71**: 229-232.
- MARGULES, C.R. Y AUSTIN, M.P. 1994. Biological models for monitoring species decline: the construction and use of data bases. *Philosophical Transactions of the Royal Society of London, B.*, **344**: 69-75.
- MAY, R.M., LAWTON, J.H. Y STORK, N.E. 1995. Assessing extinction rates. En LAWTON, J.H. Y MAY, R.M. (eds.), págs. 1-24. *Extinction Rates*. Oxford University Press, Oxford.
- MAYR, E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge.
- MCCOY, E.D. 1990. The distribution of insects along elevational gradients. *Oikos*, **58**: 313-322.
- MCCULLAGH, P. Y NELDER, J.A. 1983. *Generalized Linear Models*. Chapman and Hall, London.
- MILLER, R.I. 1994. *Mapping the Diversity of Nature*. Chapman and Hall, London.
- MITCHELL, N.D. 1991. The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). *Journal of the Royal Society of New Zealand*, **21**: 13-24.
- MURPHY, D.D. 1990. Conservation biology and scientific method. *Conservation Biology*, **4**: 203-204.
- NICHOLLS, A.O. 1989. How to make biological surveys go further with generalised linear models. *Biological Conservation*, **50**: 51-75.
- Noss, R.F. 1991. From endangered species to biodiversity. En KOHL, K (ed), págs. 227-146. *Balancing on the Brink of Extinction: The Endangered Species Act and Lessons for the Future*. Island Press, Washington.
- ODLAND, J. 1988. *Spatial Autocorrelation*. Sage publications, London.
- OSBORNE, P.E. Y TIGAR, B.J. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. *Journal of Applied Ecology*, **29**: 55-62.
- PALMER, M.W. 1990. The estimation of species richness by extrapolation. *Ecology*, **71**: 1195-1198.
- PARKER, V. 1999. The use of logistic regression in modelling the distributions of bird species in Swaziland. *South African Journal of Zoology*, **34**: 39-47.
- PRENDERGAST, J.R., WOOD, S.N., LAWTON, J.H. Y EVERSHAM, B.C. 1993. Correcting for variation in recording effort in analyses of diversity hotspots. *Biodiversity Letters*, **1**: 39-53.
- REID, W.V., BARBER, C. Y MILLER, K.R. 1992. *Global biodiversity strategy: Guidelines for action to save, study and use Earth's biotic wealth sustainably and equitably*. World Resources Institute, New York.
- RICKLEFS, R.E. Y SCHLUTER, D. (eds.). 1993. *Species Diversity in Ecological Communities*. University of Chicago Press, Chicago.
- ROLLO, C.D. 1994. *Phenotypes*. Chapman & Hall, London.
- ROSENZWEIG, M.L. Y ABRAMSKY, Z. 1993. How are diversity and productivity related?. En RICKLEFS, R.E. Y SCHLUTER, D. (eds.), págs. 52-65. *Species Diversity in Ecological Communities*. University of Chicago Press, Chicago.
- SCOTT, J.M., DAVIS, F.W., CSUTI, B., NOSS, R., BUTTERFIELD, B., GROVES, C., ANDERSON, H., CAICCO, S., D'ERCHIA, F., EDWARDS, T.C., ULLIMAN, J. Y WRIGHT, R.G.. 1993. Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs*, **123**: 1-41.
- SHIPLEY, B. 1999. Testing causal explanations in organismal biology: causation, correlation and structural equation modelling. *Oikos*, **86**: 374-382.
- SLACK, J.M.W. 1994. *From Egg to Embryo*. Cambridge University Press, Cambridge.
- SMITH, P.A. 1994. Autocorrelation in logistic regression modelling of species' distributions. *Global Ecology and Biogeography Letters*, **4**: 47-61.
- SOBERÓN, M. J. Y LLORENTE, B.J. 1993. The use of species accumulation functions for the prediction of species richness. *Conservation Biology*, **7**: 480-488.
- SOULE, M.E. 1991. Conservation tactics for a constant crisis. *Science*, **253**: 745.
- SYSTEMATICS AGENDA 2000. 1994. *Systematics Agenda 2000: Charting the Biosphere*. Technical Report. Systematics Agenda 2000, a Consortium of the American Society of Plant Taxonomy, the Society of Systematic Biologists, and the Willi Hennig Society, in cooperation with the Association of Systematics Collections, New York.
- TER BRAAK, C.J.F. Y LOOMAN, C.W.N. 1995. Regression. En JONGMAN, R.H.G., TER BRAAK, C.J.F. Y VAN TONGEREN, O.F.R. (eds.), págs. 29-77. *Data Analysis in Community and Landscape Ecology*. Cambridge University Press, Cambridge.
- TURNER, A. Y PATERSON, H.E.H. 1991. Species and speciation: evolutionary tempo and mode in the fossil record reconsidered. *Geobios*, **24**: 761-769.
- WALKER, P.A. 1990. Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. *Journal of Biogeography*, **17**: 279-289.
- WALTER, G.H. Y H.E.H. PATERSON. 1994. The implications of palaeontological evidence for theories of ecological communities and species richness. *Australian Journal of Ecology*, **19**: 241-250.
- WEBB, T., BARTLEIN, P.J., HARRISON, S.P. Y ANDERSON, K.H. 1993. Vegetation, lake levels, and climate in eastern North America for the past 18,000 years. En WRIGHT, H.E. *et al.* (eds.), págs. 468-513. *Global Climates Since the Last Glacial Maximum*. University of Minnesota Press, Minneapolis.
- WILLIAMS, P. Y GASTON, K.J. 1994. Measuring more 'diversity': Can higher taxon richness predict wholesale species richness?. *Biological Conservation*, **67**: 211-217.
- WILSON, E.O. 1988. The current state of biological diversity. En WILSON, E.O. (ed.), págs. 3-18. *Biodiversity*. National Academy Press, Washington D.C.
- WOHLGEMUTH, T. 1998. Modelling floristic species richness on a regional scale: a case study in Switzerland. *Biodiversity and Conservation*, **7**: 159-177
- WRIGHT, D.H., CURRIE, D.J. Y MAURER, B.A. 1993. Energy supply and patterns of species richness on local and regional scales. En RICKLEFS, R.E. Y SCHLUTER, D. (eds.), págs. 66-74. *Species Diversity in Ecological Communities: Historical and Geographical Perspectives*. Chicago University Press, Chicago.