

# Con los pies en el suelo: incluyendo la estructura espacial de los datos en los análisis multivariantes

A. López de Luzuriaga<sup>1</sup>, J.M. Olano<sup>2</sup>

(1) Área de Biodiversidad y Conservación, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933 Móstoles, España.

(2) Área de Botánica, EUI Agrarias, Universidad de Valladolid, Los Pajaritos s/n, 42003 Soria (Soria).

**Con los pies en el suelo: incluyendo la estructura espacial de los datos en los análisis multivariantes.** Al igual que las especies individuales, las comunidades muestran patrones espaciales muy nítidos. Sin embargo, a diferencia de las primeras, las comunidades, no son entidades discretas, sino un conjunto de elementos. Por ello, los métodos para describir los patrones y los factores ambientales que determinan estas estructuras espaciales difieren de los empleados para las especies individuales, pues nos enfrentamos a datos multivariantes. Este trabajo presenta dos grupos de técnicas (los análisis canónicos y el test de Mantel) como estrategias para tratar este tipo de datos.

Palabras clave: Análisis Canónico de Correspondencias (CCA), Análisis de Redundancia (RDA), análisis parciales, correlograma de Mantel, datos multivariantes, patrones espaciales, test de Mantel.

**With the feet on the ground: incorporating spatial patterns in multivariate analyses.** Similarly to individual species, plant communities also show clear spatial patterns. However, plant communities are not discrete entities but an assembly of elements; consequently, different methods are needed to describe the environmental factors that determine spatial patterns in plant communities. In this paper we show two types of techniques in order to deal with multivariate data analysis: canonical analysis and Mantel analysis.

Key words: Canonical Correspondence Analysis (CCA), Mantel correlogram, Mantel test, multivariate data, partial analyses, Redundancy Analysis (RDA), spatial patterns

## Introducción

Más allá de la clásica discrepancia sobre si las comunidades son entidades discretas (Clements) o una simple suma de especies individuales (Gleason), debe reconocerse que el concepto de comunidad ha resultado una herramienta de gran utilidad para el progresivo conocimiento de nuestro entorno. Su carácter sintético ha permitido una mejor descripción del territorio, especialmente para los organismos sésiles, como las plantas, en los cuales el concepto de comunidad tiene un reflejo físico, más o menos permanente, en el espacio.

Por ello, no sorprende que el concepto de comunidad haya sido durante muchas décadas el centro de escuelas científicas que, canalizado un enorme esfuerzo, han buscado su descripción y análisis. Sin embargo, y muy probablemente debido a la naturaleza compleja de las comunidades, el uso de herramientas estadísticas ha sido tradicionalmente muy reducido. A pesar de que muchas de estas técnicas estadísticas necesarias para trabajar con los datos que caracterizan la descripción de comunidades se desarrollaron hace bastante tiempo, su implementación era sumamente compleja hasta tiempos relativamente recientes. De hecho, no es hasta las últimas décadas cuando la aparición de ordenadores personales y de software de fácil manejo, ha simplificado tremendamente la aproximación estadística a este tipo de datos.

Estas técnicas se emplearon inicialmente en la discriminación de grupos, buscando dar una respuesta a la que durante mucho tiempo había sido de las principales cuestiones del análisis de las comunidades vegetales. Posteriormente, comenzó a considerarse el efecto de los factores ambientales y temporales sobre la composición de la comunidad, lo que permitió plantear otro tipo de cuestiones (sucesión, gradientes ambientales, efecto del manejo...). Sin embargo, en este contexto, sorprende el relativamente poco esfuerzo dedicado a comprender el efecto del espacio en el análisis de comunidades, a pesar de que la importancia que siempre han dado estas escuelas a la distribución geográfica de las comunidades.

Este trabajo busca describir las principales herramientas estadísticas disponibles para relacionar datos multivariantes, como los propios de las matrices de composición/abundancia de los inventarios de comunidades (pero también de otros ámbitos de la ecología), con los datos de la posición geográfica. Nuestro objetivo es exponer del modo más sencillo posible que técnicas pueden emplearse y cual es la lógica de su funcionamiento, pero no constituir un manual. Existen en el mercado textos de gran calidad, que explican con mucho más detalle cada una de las técnicas, y en los cuales nos hemos apoyado, cuyo estudio sugerimos encarecidamente al lector que quiera profundizar en este campo (Sokal y Rohlf 1995; ter Braak y Prentice 1988; Legendre y Legendre 1998; Camarero y Rozas 2006).

Los análisis que vamos a incluir permiten 1) testar hipótesis acerca de la existencia de gradientes o patrones espaciales en datos multivariantes, 2) modelizar las relaciones comunidad-ambiente al tiempo que se tiene en cuenta la estructura espacial de los datos. Los métodos de análisis que se pueden utilizar para el análisis espacial con datos multivariantes se desarrollan a partir de dos familias de técnicas:

- a) *Análisis canónicos*: Estas técnicas se basan en ordenaciones y se trabaja con los datos originales. Se trata de establecer la relación entre los datos multivariantes, las coordenadas geográficas de los lugares y otros datos ambientales.
- b) *Test de Mantel*: Los análisis se centran en matrices de similitud/distancias derivadas de los datos originales. En este caso el análisis se realiza entre distancias ecológicas (calculadas a partir de los datos multivariantes) y las distancias geográficas (calculadas a partir de las coordenadas de los datos).

## Análisis canónicos

Los análisis canónicos consisten en el análisis simultáneo de dos o más matrices de datos multivariantes. Mediante un análisis canónico se extrae toda la varianza de una matriz respuesta ( $Y$ ) que está relacionada con una matriz explicativa ( $X$ ). Este tipo de análisis se suele aplicar para evaluar la relación entre una tabla con datos de composición de especies ( $Y$ ) y otra con datos de descriptores ambientales ( $X$ ). En este tipo de análisis, la ordenación de la matriz respuesta se constriñe de forma que los vectores de ordenación resultantes son combinaciones lineales de las variables de la matriz explicativa. Esta herramienta se ha utilizado profusamente en los estudio de comunidades. Sin duda, la existencia de programas muy accesibles, en especial el desarrollo del paquete Canoco (ter Braak y Smilauer 2002), ha sido clave en la popularidad de esta técnica.

La posibilidad de comprobar si existe una relación estadísticamente significativa entre la composición de la comunidad y una matriz de variables ambientales permite pasar de un uso meramente descriptivo de las ordenaciones a emplearlas como herramientas para el contraste de hipótesis formuladas *a priori*. La hipótesis nula de un Análisis Canónico es que no existe relación alguna entre la matriz de variables respuesta y la matriz de variables explicativas. La significación de esta hipótesis nula se puede evaluar mediante permutaciones. En este caso, el estadístico es la suma de todos los valores propios canónicos y su significación se evalúa mediante un test de F. La hipótesis alternativa establece que la suma de todos los valores propios canónicos es mayor que la que se obtiene de matrices en las que se han permutado sus filas de datos. La suma de todos los valores propios canónicos dividida por el total de variación de la matriz  $Y$  permite obtener la proporción de la variación de  $Y$  explicada por  $X$ , que es análoga al coeficiente de determinación de las regresiones múltiples ( $r^2$ ). Las posibilidades de esta técnica son aún mayores, ya que mediante el empleo de técnicas de selección por pasos puede llegar a determinarse cuáles de las diferentes variables incluidas en la matriz explicativa ( $X$ ) tienen una relación significativa con la matriz respuesta ( $Y$ ).

La inclusión del componente espacial en estos análisis es extremadamente sencilla, basta incluir las coordenadas geográficas dentro de la matriz explicativa  $X$ . De este modo se modela la variación espacial de las variables de interés como combinaciones lineales entre las variables respuesta y las coordenadas geográficas de cada lugar. En este caso disponemos de dos tipos de matrices: la matriz  $Y$ , compuesta por las variables respuesta, y la matriz  $X$ , que incluye las coordenadas geográficas (y otras que puedan ser de interés) que se van a utilizar como variables explicativas. Más allá de los problemas estadísticos, la mayor dificultad para el empleo del componente espacial en estos análisis radica en la necesidad de elaborar diseños de muestreo adecuados.

Las técnicas más empleadas están basadas en el Análisis de Componentes Principales o en el Análisis de Correspondencia, según asumamos un espacio euclídeo con respuestas lineales o un espacio basado en la distancia chi-cuadrado y respuestas unimodales de las especies. También es posible realizar ordenaciones canónicas a partir de otras técnicas, como el Análisis de Coordenadas Principales (PCoA) o escalamiento multidimensional no métrico (NMDS).

Sin embargo, al margen de que se emplean unas técnicas u otras, todos estos análisis asumen una relación lineal entre la matriz  $Y$  y la matriz  $X$ , lo que, si bien es muy eficiente para detectar tendencias lineales, lo hace inadecuada para otros tipos de estructura espacial. Una opción para tratar de solventar estas limitaciones es incluir dentro de la matriz  $X$  diferentes funciones de las coordenadas geográficas  $x$  e  $y$  ( $x^2$ ,  $y^3$ ,  $y^2x$ ,...). Sin embargo, el aumento la complejidad de las variables espaciales conlleva un incremento en la dificultad de interpretar los resultados obtenidos.

Olano *et al.* (2002) muestrearon el banco de semillas en una malla regular en un bosque mixto de hayas, abedules y robles en Urkiola (Bizkaia, **Fig. 1**). El objetivo era comprender cuáles eran los factores que estructuraban el banco. El análisis mediante un Análisis Canónico de Correspondencias (CCA) reveló que la estructura espacial explicaba un modo muy significativo ( $P < 0.001$ ) un 10,3% de la composición de especies.

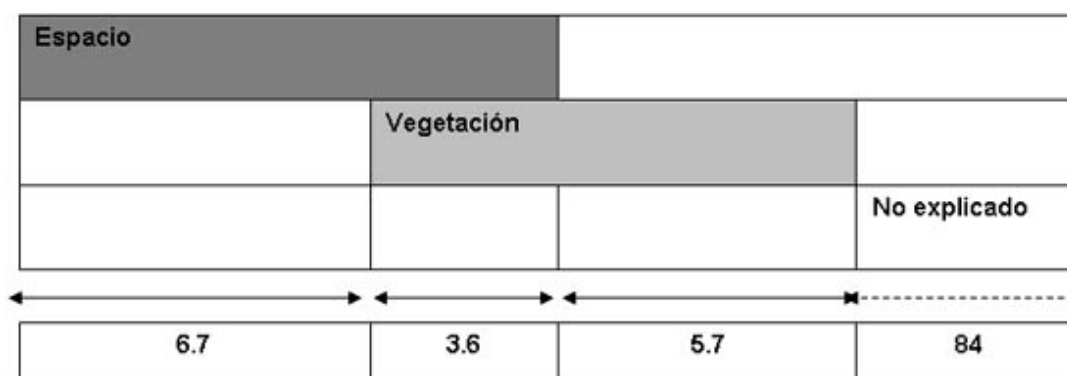


**Figura 1.** Vista del bosque mixto de Urkiola estudiado en Olano *et al.* (2002).

## Análisis canónicos parciales

El espacio en sí mismo no es un factor ecológico, sino que es una variable que integra los procesos y factores que están estructurados espacialmente (Borcard *et al.*, 1992). Por ello, es importante tratar de determinar qué parte de la estructura espacial puede ser explicada con otras variables con un efecto directo sobre los procesos ecológicos. Esto puede abordarse mediante las técnicas de partición de la varianza. Esta aproximación consiste en utilizar dos matrices de variables explicativas, por ejemplo una compuesta por variables ambientales ( $Y$ ), y otra por las coordenadas geográficas de los puntos de muestreo ( $Z$ ).

Los modelos espaciales parciales nos ofrecen una información muy valiosa sobre el reparto de la variabilidad asociado a diferentes factores. De este modo podemos cuantificar (y comprobar) qué fracción de la variabilidad explicada por las variables ambientales se debe a su estructura espacial, cuánta variabilidad ambiental no puede explicarse como una combinación lineal de la matriz espacial -lo cual no quiere decir que carezca de estructura espacial- y, por último, cuanta variabilidad sigue explicando la matriz espacial después de eliminar las variables ambientales (**Fig. 2**). Analizar esta variabilidad espacial restante es importante porque seguramente estará relacionada con variables ambientales que no se han tenido en cuenta en el modelo.



**Figura 2.** Análisis Canónico de Correspondencias parcial de la matriz de banco de semillas. Las líneas horizontales indican la variación total explicada por cada matriz (en %). La porción en común indica la varianza compartida, la parte que no es coincidente indica la varianza que no puede ser explicada por la otra matriz. El nivel de significación de estas partes es de 0,0001 (basado en 10.000 permutaciones). Adaptado de Olano *et al.* (2002).

En Olano *et al.* (2002) se tenía interés en conocer cuánto de la estructura espacial del banco de semillas estaba condicionada por la estructura de la vegetación del sotobosque. La hipótesis de partida era que la vegetación aérea, como fuente de semillas, iba a tener un efecto importante en la estructura espacial del banco. Para ello se empleó un Análisis Canónico de Correspondencias parcial (pCCA). El resultado indicó que, si bien existía un importante solapamiento entre la variabilidad explicada por la posición espacial y la relacionada con la composición del sotobosque, ambas matrices seguían explicando fracciones independientes y significativas de la variación de los datos. En el caso de la estructura espacial, una parte de ella era explicada por la estructura del dosel arbóreo y por las características del sustrato. El resto de la variabilidad del banco de semillas, relacionada con la composición del sotobosque, no pudo explicarse de un modo lineal con las variables espaciales, aunque esto no quiere decir que no estuviese estructurada espacialmente.

Con el propósito de comprender los factores que determinan la distribución de las especies y contrastar la teoría neutral frente a la teoría de nicho, Gilbert & Lechowicz (2004) compararon el efecto de la posición geográfica (equivalente a dispersión) frente al del ambiente (descriptores del nicho) en la composición de la comunidad florística en un bosque de Canadá. Para ello emplearon un pCCA, determinando la fracción de varianza debida al espacio frente a la explicada por el ambiente. El resultado fue que la mayor parte de la varianza era explicada exclusivamente por factores ambientales, mientras que el espacio (dispersión, según los autores) explicaba una porción entre 2 y 20 veces menor. Los autores concluyeron que, en su caso, la teoría del nicho explicaba más correctamente la composición florística que la teoría neutral.

## Test de Mantel

El test de Mantel estima el grado de correlación existente entre dos matrices  $X$  e  $Y$ . La hipótesis nula de esta técnica ( $H_0$ ) postula que las distancias/similitudes entre las variables de la matriz respuesta  $Y$  no están linealmente correlacionados con las correspondientes distancias/similitudes en la matriz modelo  $X$ . Se trata, por tanto, de evaluar si la asociación (positiva o negativa) es más robusta de lo que cabría esperar por puro azar. La inclusión del espacio en este proceso es sencilla, basta incluir como matriz modelo  $X$ , una matriz de distancias. De este modo puede evaluarse si la distancia en el espacio determinado por las especies está correlacionada con la distancia en el espacio geográfico (para construir esta matriz pueden emplearse diferentes medidas de distancia, no sólo la euclídea, incluyendo también la posibilidad de construir matrices no paramétricas, útiles para datos que no cumplen las condiciones de normalidad). En este sentido, es necesario remarcar la importancia de la elección de un tipo de medida u otra, así como el tipo de transformación, ya que los resultados varían en gran medida dependiendo de estos elementos.

El estadístico del test de Mantel ( $Z_M$ ) se calcula mediante la suma de los productos cruzados de los valores de las dos matrices de similitud/distancia, excluyendo la diagonal principal que sólo contiene valores triviales (0 en el caso de las matrices de distancias y 1 en el caso de las matrices de similitudes).

$$Z_M = \sum X_{ij} Y_{ij}$$

Donde  $X_{ij}$  e  $Y_{ij}$  son los elementos de las matrices  $X$  e  $Y$ , respectivamente. Para evaluar la significación de este estadístico se utiliza un test de permutaciones, en el que los elementos de una matriz se reordenan al azar. Se permutan los elementos de una de las matrices al azar y se calcula iterativamente el valor de  $Z$ . Así de la distribución de valores  $Z$  obtenidos al azar podemos evaluar cuál es la probabilidad de obtener el valor  $Z$  observado.

El estadístico  $Z$  se expresa en unidades arbitrarias y sus implicaciones son difíciles de entender, por lo que suele utilizarse el coeficiente estandarizado de Mantel. Para esto se calcula el producto cruzado de las dos matrices dividido por  $[(n(n-1)/2)-1]$  cuyo rango está entre -1 y +1. Se comporta como un coeficiente de correlación, y se denomina  $r_M$ . Su significación se evalúa también mediante permutaciones. El test de permutaciones proporciona en mismo nivel de significación para el estadístico  $Z_M$  que para  $r_M$ , ya que las transformaciones lineales, como la utilizada para estandarizar  $Z$ , afectan en la misma medida a los resultados del producto cruzado.

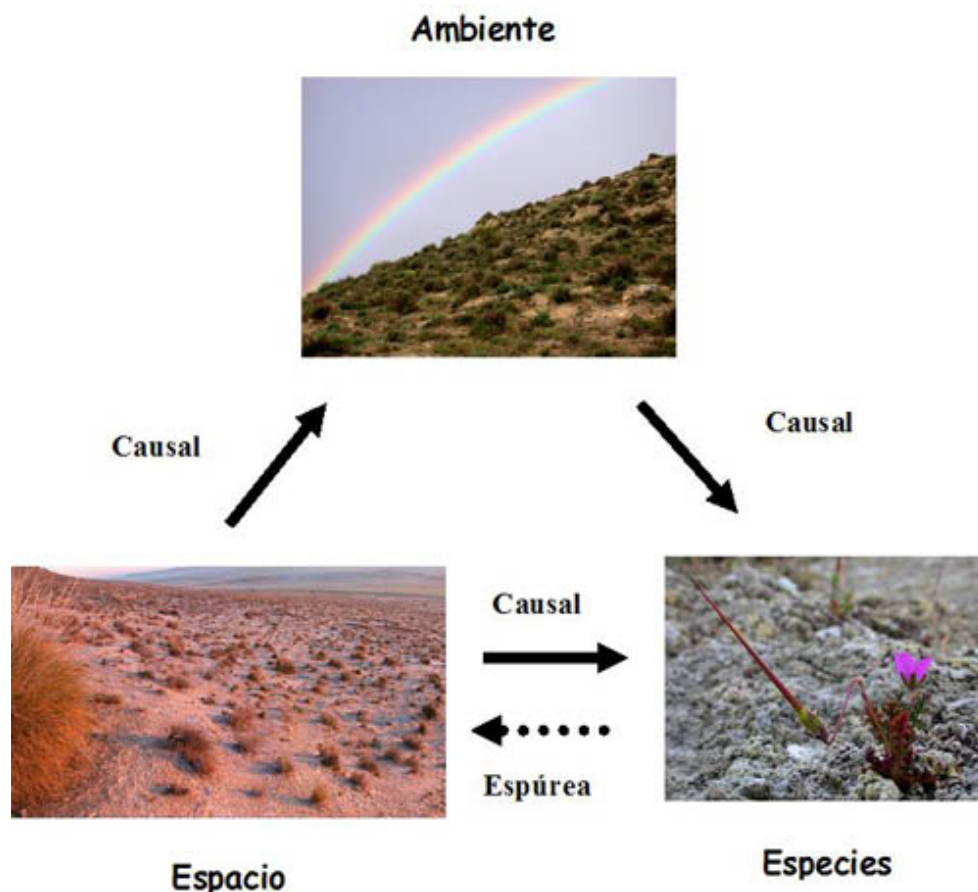
En general la interpretación es sencilla para los valores positivos de  $r_M$ . En el caso que nos ocupa, significaría que cuanto más próximo espacialmente se encuentren dos puntos, tanto más parecida es su composición de especies. Sin embargo, los valores negativos de  $r_M$  resultan mucho más difíciles de interpretar. En cualquier caso, en estos datos habría que considerar si trabajamos con matrices de distancias o de similitudes, especialmente cuando  $X$  e  $Y$  sean diferentes, ya que entonces la interpretación del signo sería la contraria.

A diferencia de las técnicas basadas en ordenaciones, el test de Mantel se basa en una evaluación global de la similitud entre las diferentes matrices, no pudiendo realizarse análisis para seleccionar cuál es el efecto de las diferentes variables de la matriz  $X$  (más allá de evaluar el efecto de matrices con un número más reducido de variables). Del mismo modo, incluir un mayor número de variables en la matriz  $X$  no incrementa la capacidad explicativa de dicha matriz (incluso puede reducirla), como sí ocurre en los análisis basados en ordenaciones.

Guerrero *et al.* (2005) utilizaron el test de Mantel para evaluar diferentes hipótesis para conocer qué factores habían determinado la composición actual de la fauna de reptiles endémicos de las Islas Canarias. Para ello compararon las composiciones de fauna existentes en la actualidad con las predichas mediante 11 hipótesis diferentes. La mejor hipótesis incluía en el modelo la distancia entre islas, así como la similitud de hábitat.

### Test de Mantel parcial

Al igual que ocurría con el caso de las ordenaciones, existe una técnica que permite controlar el la correlación entre la matriz  $Y$  y la matriz ambiental/espacial  $X$ . Este tipo de análisis parcial permite evaluar la correlación entre dos matrices ( $X$  e  $Y$ ) mientras se controla el efecto de una tercera matriz ( $Z$ ). Este tipo de análisis (**Fig. 3**) se puede responder a preguntas del tipo: ¿La estructura espacial de las variables ambientales explica completamente la estructura espacial de nuestros datos multivariantes (p. ej. composición de la vegetación)?, o ¿hay una parte de la estructura espacial de los datos multivariantes que no es explicada por las variables ambientales del modelo?



**Figura 3.** Esquema de las relaciones existentes entre espacio, ambiente y las especies. La estructura espacial determina parte de los valores de las variables ambientales, que a su vez pueden afectar a la distribución de las especies. A su vez la estructura espacial puede tener un efecto sobre las especies, si bien parte de esa relación es espúrea pues se produce a través del efecto de la estructura espacial sobre el ambiente. Mediante el empleo del test parcial de Mantel (o de correlaciones parciales en caso de datos univariados) se elimina este efecto. Basado en Urban (2003).

La correlación entre ambas matrices se calcula de forma similar al coeficiente de correlación parcial, sólo que con el estadístico estandarizado de Mantel en lugar del de Pearson. La principal diferencia estriba en que los valores provienen de matrices de similitud en lugar de provenir de vectores. Hay que tener en cuenta que la correlación entre dos matrices de distancias no es equivalente a la correlación entre dos tablas de datos: la correlación de Mantel entre las matrices  $X$  e  $Y$  es una estima de en qué medida las variaciones en las similitudes/distancias de  $X$  corresponden a las variaciones en  $Y$ . La significación del test de Mantel parcial se suele calcular también mediante permutaciones, pero en este caso permutando la matriz  $X$  mientras se mantienen constantes las matrices  $Y$  y  $Z$ . Esta técnica ha recibido críticas de algunos autores (Raufaste & Rousset, 2001), pero parece que en condiciones de correlación moderada entre las matrices independientes los resultados son satisfactorios (Castellano & Balletto 2002). Por otra parte, puede utilizarse la información de las matrices de distancias residuales que se obtienen mediante los análisis parciales de Mantel como base para ordenaciones que permitan explicar dicha variabilidad.

Del mismo modo que ocurría con los test de Mantel, en el caso de los test parciales de Mantel, debemos considerar que la inclusión de un mayor número de variables en la tercera matriz ( $Z$ ) no aumenta la capacidad explicativa esta matriz, ya que lo que se compara son las matrices de distancia/similitud y no las variables individuales como ocurre con las ordenaciones canónicas parciales. Por ello, es mejor reducir las variables ambientales/espaciales a aquellas que tienen un sentido ecológico dentro del problema que se plantea.

Jacquemyn *et al.* (2001) compararon diferentes manchas de bosque aisladas en el centro de Bélgica. El objetivo de este trabajo era determinar cuál era el papel que sobre la composición florística del bosque ejercían la distancia a otras masas y la edad del bosque. Para ello utilizaron test de Mantel y tests de Mantel parciales con tres matrices (composición del bosque,

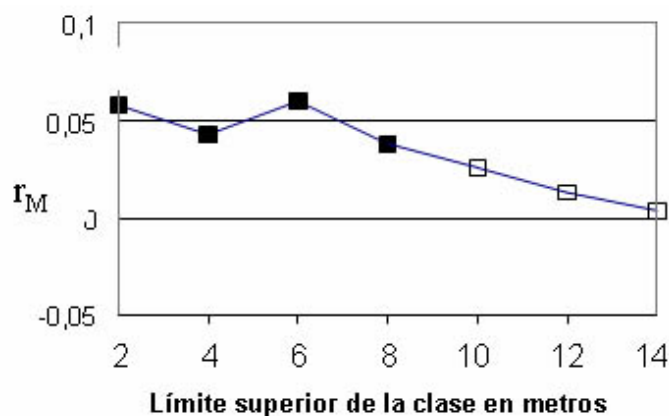
edad del bosque y una matriz de distancias entre las masas). El test de Mantel indicó que la distancia entre las manchas de bosque era más importante que la edad de las masas para explicar la composición florística, manteniéndose esta relación incluso cuando se realizaba un test parcial de Mantel con la edad de la masa como variable control. La conclusión de estos autores era que la composición el efecto de los procesos de colonización y aislamiento que ocurrían en estas masas de bosques tenían mayor peso en la composición específica que los procesos internos derivados de la edad de las masas.

### Correlograma de Mantel

Las técnicas anteriormente utilizadas permiten establecer las relaciones entre el patrón espacial y la composición de una comunidad. Sin embargo, no permiten detectar cómo varía la intensidad de dicha relación a diferentes escalas. En el caso de datos univariantes, existen una amplia panoplia de herramientas con que detectar estas estructuras espaciales (véanse los artículos de A. Gallardo y M. de la Cruz, en este volumen). Cuando se trabaja con datos multidimensionales el correlograma de Mantel aparece como la única solución sencilla para realizar análisis de autocorrelación. Un correlograma es una función de estructura que permite cuantificar la dependencia espacial, y repartirla entre clases de distancias.

Este procedimiento calcula los valores de estadístico normalizado de Mantel ( $r_M$ ) para diferentes clases de distancia ( $d$ ). Para la clase de distancias 1, se construye una matriz  $X_1$  con valores 1 para los puntos de muestreo que pertenecen a la primera clase de distancias, mientras el resto de valores de la matriz son 0. Esta matriz  $X$  se compara con la matriz  $Y$  de similitudes o distancias mediante el estadístico de Mantel ( $r_M$ ). Este proceso se repite de manera idéntica para el resto de clases de distancias, de forma que se construye una matriz modelo  $X_d$  y se recalcula el estadístico de Mantel normalizado para cada clase de distancias. El resultado se expresa de mediante un gráfico en el que se representan los valores del estadístico normalizado de Mantel ( $r_M$ ) en las ordenadas respecto a clases de distancias entre sitios en las abscisas. Dado que en esta técnica se evalúan repetidamente los mismos datos, es conveniente aplicar algún tipo de corrección de la significación (por ejemplo el test de Bonferroni).

En Olano *et al.* (2002) se analizó la estructura espacial del banco de semillas mediante un correlograma de Mantel (**Fig. 4**). Se observó que la composición tenía mayor semejanza a distancias entre 0 y 8 m. Un análisis posterior, mediante el estudio de algunas especies individuales determinó que dicho patrón se debía al banco de semillas de las diferentes especies de brezos y de *Juncus effusus* L. Si bien la mayor parte de los brezos seguían estando presentes en la vegetación aérea, los juncos ya no lo estaban, por lo que su banco debió formarse en el pasado. Este resultado podía explicar que parte de la estructura espacial del banco de semillas no pudiese ser explicada con la estructura de vegetación aérea actual, sino que estaba condicionado por la estructura espacial de la vegetación que existía en el momento en que se formó el banco.



**Figura 4.** Correlograma de Mantel para los datos de composición del banco de semillas. Los cuadrados rellenos indican correlaciones significativas ( $P < 0,05$ , tras la corrección de Bonferroni para comparaciones múltiples). El intervalo de cada clase es de 2 m, empezando desde 0. Se observa como existe un patrón de asociación positiva en los primeros 8 metros. Adaptado de Olano *et al.* (2002).

## ¿Qué técnica escoger?

Aunque tanto las técnicas de ordenación como las basadas en el test de Mantel son útiles para analizar el efecto del espacio en la variabilidad en la composición de una comunidad, sus usos y resultados no siempre son intercambiables. Ambas técnicas son muy eficientes para calcular el efecto de las distancias geográficas respecto a la variación de la composición de la comunidad. Si bien en el caso de gradientes largos en las cuales pueden aparecer respuestas unimodales, es recomendable el uso de técnicas basadas en el análisis de correspondencia. Por otra parte, los análisis canónicos disponen de herramientas que permiten determinar cuáles de los diferentes componentes espaciales son los responsables de la relación entre variables, lo cual suele hacerlos más adecuados en la mayor parte de los casos.

En este sentido, cuando disponemos de numerosas variables en la matriz  $X$ , el uso de análisis basados en el test de Mantel es recomendable únicamente cuando dichas variables tengan una relación entre sí que las unifique (distancias genéticas, distancias geográficas, composición de una comunidad). Sin embargo, cuando se tratan de variables independientes entre sí (una matriz ambiental incluyendo precipitación, sustrato, pendiente...) su uso no es recomendable ya que una gran parte de la información se pierde al trabajar exclusivamente con una única matriz de distancias/similitud, obtenida a partir de todos los datos ambientales. Se pueden realizar un test de Mantel para cada una de las variables individuales, con el fin de ver cuáles de ellas tienen efectos significativos (King *et al.*, 2004), pero no sólo se trata procedimiento sumamente trabajoso, sino que su interpretación conjunta resulta complicada. En estos casos el empleo de técnicas de análisis canónico, especialmente empleando análisis canónicos parciales, permite interpretar mucho mejor la información (Legendre *et al.*, 2005).

Sin embargo, cuando lo que comparamos son dos matrices basadas en la composición de comunidades (vegetación aérea-banco de semillas) los test de Mantel son muy útiles, en especial cuando la matriz  $X$ , que referida a la composición de una comunidad, incluye un número muy elevado de especies, ya que en estos casos los análisis canónicos no son eficientes ante una inflación de variables que inevitablemente va a arrojar correlaciones espurias. Igualmente, no conocemos ninguna técnica basada en el análisis canónico que permita obtener información de la estructura espacial de la comunidad semejante a la obtenida mediante el correlograma de Mantel.

## Software recomendado

### Análisis canónico

*Canoco*. Este programa es por su versatilidad, rapidez, facilidad de manejo, capacidad de manejar grandes matrices dispersas y abundante bibliografía, uno de los más recomendables para trabajar con ordenaciones. Desgraciadamente, su precio ha ido incrementándose parejo con su popularidad y ya no resulta tan barato como en el pasado (<http://www.microcomputerpower.com/>).

*VegAna* Este software gratuito desarrollado por X. Font, M. de Cáceres y R. García de la Universitat de Barcelona que permite realizar diferentes técnicas de ordenación e incluye el test de Mantel (<http://biodiver.bio.ub.es/vegana/catala/index.html>).

*Brodgar*. Programa de pago que permite realizar análisis canónicos, así como numerosas técnicas estadísticas adicionales como GLM o GAM. Actúan como interfaz de R utilizando rutinas de dicho programa (<http://www.brodgar.com>).

*PC-ORD* Programa de pago muy popular, sobre todo en los Estados Unidos. Permite realizar la mayor parte de las técnicas de ordenación y ofrece la posibilidad de hacer el test de Mantel (<http://home.centurytel.net/~mjm/pcordwin.htm#3>).

### Test de Mantel

*Passage*, de Michael S. Rosenberg, es un software libre capaz de realizar el test de Mantel, el test de Mantel parcial, así como numerosas otras técnicas de análisis espacial. La última versión disponible es la 1.1.3.4, si bien parece que una nueva versión mejorada está en marcha ([www.passagesoftware.net/](http://www.passagesoftware.net/)).

*Mantel Nonparametric Test Calculator for Windows Version 2.00*, de Liedloff (1999), es un programa libre de fácil manejo para hacer el test de Mantel (<http://www.sci.qut.edu.au/NRS/Mantel.htm>)

*R 4.0*. Este programa libre, desarrollado por P. Casgrain y P. Legendre, ofrece un gran número de rutinas para realizar numerosos estadísticos, incluyendo análisis espacial. El único inconveniente es que requiere un esfuerzo previo para comprender el lenguaje. En esta versión está disponible para McIntosh, aunque puede ser utilizada en PC con un emulador (<http://www.bio.umontreal.ca/casgrain/en/labo/index.html>).



## Agradecimientos

J.M. Olano ha sido financiado dentro del proyecto de investigación de la Junta de Castilla y León VA0110B05.

## Referencias

- Borcard, D., Legendre, P. y Drapeau P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 76: 2176-2183.
- Camarero, J.J. y Rozas, V. 2006. Técnicas de análisis espacial de patrones de superficies y detección de fronteras aplicado a la ecología forestal. *Investigaciones Agrarias Sistemas y Recursos Forestales* 15: 66-87.
- Castellano, S. y Balletto, E. 2002. Is the partial Mantel test inadequate? *Evolution* 56: 1871-1873.
- Gilbert, B. y Lechowicz, M.J. 2004. Neutrality, niches, and dispersal in a temperate forest understory. *PNAS* 101: 7651-7656.
- Guerrero, J.C., Vargas, J.M. y Real, R. 2005. A hypothetico-deductive analysis of the environmental factors involved in the current reptile distribution pattern in the Canary Islands . *Journal of Biogeography* 32: 1343-1351.
- Jacquemyn, H., Butaye, J., Dumortier, M., Hermy, M. y Lust, N. 2001. Effects of age and distance on the composition of mixed deciduous forest fragments in an agricultural landscape. *Journal of Vegetation Science* 12: 635-642.
- King, R.S., Richardson, C.J., Urban, D.L. y Romanowicz, E.A. 2004. Spatial dependency of vegetation-environment linkages in an anthropogenically influenced wetland ecosystem. *Ecosystems* 7: 75-97.
- Legendre, P. y Legendre, L. 1998. *Numerical ecology*. Elsevier, Ámsterdam
- Legendre, P., Borcard, P. y Peres-Neto, P.R. 2005. Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecological Monographs* 75: 435-450.
- Liedloff, A. 1999. Mantel (Version 2.0): Mantel Nonparametric Test Calculator. Queensland University of Technology , Brisbane, .
- Olano, J.M., Caballero, I., Laskurain, N.A., Loidi, J. y Escudero, A. (2002) Seed bank spatial pattern in a temperate secondary forest. *Journal of Vegetation Science* 13: 775-784.
- Raufaste, N. y Rousset, F. 2001. Are partial mantel tests adequate? *Evolution* 55: 1703-1705.
- Rosenberg, M.S. 2001. *PASSAGE. Pattern Analysis, Spatial Statistics, and Geographic Exegesis. Version 1.1.3.4*. Department of Biology, Arizona State University , Tempe, AZ.
- Sokal, R.R. y Rohlf, F.J. 1995. *Biometry: the principles and practice of statistics in biological research*. 3rd edition. W. H. Freeman and Co.: New York .
- ter Braak, C.J.R. y Prentice, I. 1988. A theory of gradient analysis. *Adv Ecol Res* 18:271-317
- ter Braak, C.J.F. y Šmilauer, P. 2002. *CANOCO Reference Manual and CanoDraw for Windows User's Guide: Software for Canonical Community Ordination (version 4.5)*. Microcomputer Power ( Ithaca NY, ).
- Urban, D. 2003. *Spatial Analysis in Ecology. Mantel's Test* (<http://www.nceas.ucsb.edu/scicomp/Dloads/SpatialAnalysisEcologists/SpatialEcologyMantelTest.pdf>)