

# Introducción al Análisis Espacial de Datos en Ecología y Ciencias Ambientales: Métodos y Aplicaciones

FERNANDO T. MAESTRE

ADRIÁN ESCUDERO

ANDREU BONET

*(editores)*



Todos los derechos reservados. Ni la totalidad ni parte de este libro, incluido el diseño de la cubierta, puede reproducirse o transmitirse por ningún procedimiento electrónico o mecánico. Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, [www.cedro.org](http://www.cedro.org)) si necesita fotocopiar o escanear algún fragmento de esta obra.

© Copyright by  
Universidad Rey Juan Carlos  
Servicio de publicaciones  
Los autores  
Madrid, 2008

Editorial DYKINSON, S.L.  
Meléndez Valdés, 61 - 28015 Madrid  
Teléfono (+34) 91544 28 46 - (+34) 91544 28 69  
e-mail: [info@dykinson.com](mailto:info@dykinson.com)  
<http://www.dykinson.es> <http://www.dykinson.com>

ISBN: 978-84-9849-308-5  
Depósito Legal:

*Preimpresión:*

Besing Servicios Gráficos, S.L.  
[besing@terra.es](mailto:besing@terra.es)

# Introducción

FERNANDO T. MAESTRE

Área de Biodiversidad y Conservación, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933 Móstoles, España. E-mail: [fernando.maestre@urjc.es](mailto:fernando.maestre@urjc.es)

ADRIÁN ESCUDERO

Área de Biodiversidad y Conservación, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933 Móstoles, España. E-mail: [adrian.escudero@urjc.es](mailto:adrian.escudero@urjc.es)

## RESUMEN

Las distribuciones espaciales de organismos, factores abióticos e interacciones ecológicas juegan un papel fundamental en el mantenimiento de la estructura, funcionamiento y dinámica de los ecosistemas, por lo que su estudio y análisis constituye una parte fundamental de la ecología. Si bien el análisis de los patrones espaciales de una especie, comunidad o proceso ecológico de interés, por sí mismo, no permite determinar qué proceso los determina, proporciona valiosa información sobre los procesos que han sido importantes en su génesis y, en numerosas ocasiones, supone un primer paso para diseñar experimentos encaminados a evaluar hipótesis concretas. En el presente capítulo se introducen algunos conceptos básicos, se revisan los principales tipos de datos ecológicos espacialmente explícitos y se presentan sucintamente las distintas técnicas disponibles para su análisis. Asimismo, se proporcionan distintos consejos prácticos y recomendaciones para llevar a buen puerto todo estudio ecológico que tenga como objetivo analizar datos espacialmente explícitos, y se realiza una breve descripción de algunas páginas web de interés general.

### 1.1. IMPORTANCIA DEL ANÁLISIS ESPACIAL EN ECOLOGÍA Y CIENCIAS AMBIENTALES

La presencia de distribuciones no aleatorias en la distribución de bacterias, hongos, invertebrados, vertebrados, plantas vasculares o de cual-

quier otro tipo de organismo es la norma, más que la excepción, tanto en ecosistemas terrestres como acuáticos (Rossi 2003, Crist 1998, Maestre *et al.* 2005a, Barnes *et al.* 2007). Del mismo modo, infinidad de estudios han puesto de manifiesto como la distribución espacial de factores abióticos como la precipitación, los nutrientes y el agua disponible para las plantas sigue complejos patrones espacio-temporales, en ocasiones asociados a los de distintos organismos (Ryel *et al.* 1996, Ettema *et al.* 1998, Bárdossy y Lehman 1998, Stenger *et al.* 2002, Gallardo *et al.* 2005). Todo ello ha permitido el desarrollo de lo que podemos denominar paradigma espacial de la ecología (Tilman y Kareiva 1997) y que básicamente reconoce que la relevancia del espacio y de las estructuras espaciales en el funcionamiento de los ecosistemas.

El análisis de estas distribuciones o patrones espaciales -como comúnmente se les denomina- es una parte fundamental de una amplia variedad de disciplinas científicas, entre las que se incluyen la ecología, la geología, la geografía, la biología evolutiva y la ingeniería, por citar sólo algunas. Siendo importante en todas ellas, el análisis de datos espacialmente explícitos -datos que contienen información sobre su localización en el espacio- cobra una especial relevancia en ecología y ciencias afines. Por definición, la ecología se encarga del estudio de las relaciones recíprocas entre los organismos y su medio ambiente (Begon *et al.* 1995). Dado que tanto los organismos como los factores abióticos que definen los ecosistemas presentan en la mayor parte de casos marcados patrones espaciales, el estudio de estos patrones es clave para poder entender dichas relaciones.

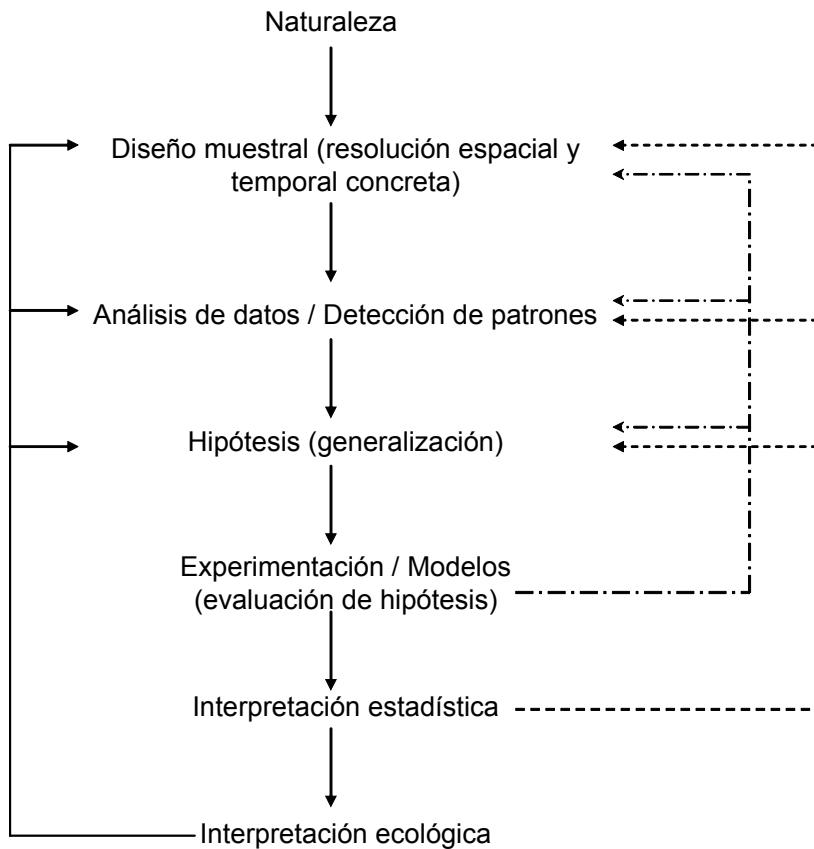
Bajo el nombre genérico de análisis espacial se engloba a un conjunto de técnicas encaminadas a analizar cuantitativamente datos espacialmente explícitos (Legendre y Fortin 1989). La utilización de este tipo de análisis en ecología y ciencias afines tiene una gran tradición. Ya durante la primera mitad del siglo XX quedó claramente establecido que la distribución de los organismos en el espacio no es aleatoria (p. ej., Blackman 1935, Watt 1947, Ashby 1948, Withford 1949), a la vez que comenzaron a desarrollarse métodos para examinar de forma cuantitativa los patrones espaciales de plantas (Watt 1947) y animales (Bliss 1941). Desde entonces, el desarrollo de técnicas analíticas ha sido un área de investigación muy fructífera, especialmente a partir del momento en el que los ordenadores han permitido abordar tareas de cálculo que con anterioridad eran sencillamente imposibles. Hoy en día existen varias decenas de métodos diferentes para

describir y evaluar cuantitativamente el patrón espacial de organismos, factores ambientales y procesos ecológicos (véase Dale 1999, Perry *et al.* 2002 y Fortin y Dale 2005 para revisiones metodológicas recientes). Junto a este desarrollo, avances en la construcción de una base teórica que tiene en cuenta los patrones espaciales de los organismos y de modelos espacialmente explícitos (Tilman y Kareiva 1997, Hassell *et al.* 1991), han contribuido notablemente a nuestra comprensión de la relación entre estructura y funcionamiento de los ecosistemas naturales.

¿Por qué tanto interés en el análisis espacial en ecología?, ¿Cuáles son las razones que han motivado los enormes avances conceptuales y metodológicos ocurridos en las últimas décadas? Varias son las razones que permiten responder a estas cuestiones. Las distribuciones espaciales de los organismos, factores abióticos e interacciones ecológicas juegan un papel fundamental en el mantenimiento de la estructura, funcionamiento y dinámica de los ecosistemas. Así, se ha comprobado que el patrón espacial de los organismos afecta a su capacidad competitiva y a su esfuerzo reproductivo (Stoll y Prati 2001, Tirado y Pugnaire 2003, Monzeglio y Stoll 2005, Internicola *et al.* 2006, Luzuriaga *et al.* 2006) y que la presencia de heterogeneidad espacial en la distribución de los factores abióticos afecta a un sinnúmero de procesos ecológicos, modificando atributos como la diversidad y composición de las comunidades (Pringle 1990, Wilson 2000, Maestre *et al.* 2005b), influyendo en interacciones bióticas como la depredación (Kaiser 1983, Webb y Wilson 1985) y el parasitismo (Nachman 1981), aumentando la estabilidad de una población (Tilman 1994), modificando los procesos de dispersión, colonización y supervivencia (Rees *et al.* 2000, Corlett 2000) y ayudando a mantener el polimorfismo genético dentro de una especie (Smith *et al.* 1983, Weider 1989, Lechowicz y Bell 1991), entre otros. Así pues, hoy en día tenemos claro que si queremos conocer cómo funcionan los ecosistemas naturales tenemos necesariamente que estudiar los patrones espaciales de sus componentes y procesos (Fig. 1.1).

Si bien el análisis de los patrones espaciales de una especie, factor abiótico, comunidad o proceso ecológico de interés, por sí mismo, no permite determinar qué proceso/s los están determinando, proporciona valiosa información sobre los factores que han podido ser importantes en su génesis (Pielou 1961, Ford y Renshaw 1984, Silvertown y Wilson 1994), en numerosas ocasiones supone un primer paso para sugerir hipótesis concretas y diseñar experimentos encaminados a evaluarlas. Asimismo, la consideración de los patrones espaciales de los organismos y factores abióticos tiene im-

portantes implicaciones prácticas para construir y validar teorías ecológicas (Tilman y Kareiva 1997), mejorar el diseño de experimentos (Legendre *et al.* 2002, 2004; Capítulo 14), gestionar especies y hábitats (Escudero *et al.* 2003, Torres *et al.* 2003, Dennis *et al.* 2002) y minimizar el impacto ambiental de actividades como la fertilización de campos agrícolas y la aplicación de pesticidas (Stenger *et al.* 2002, Wallinga *et al.* 1998), entre otros aspectos.



**Figura 1.1.** Pasos sucesivos implicados en el estudio de la naturaleza y su complejidad. Dado que los procesos naturales actúan a varias escalas espacio-temporales, el diseño muestral elegido reduce los límites espacio-temporales del proceso bajo estudio. A partir de los patrones espaciales estudiados pueden establecerse generalidades e hipótesis sobre los procesos ecológicos subyacentes, que darán pie a nuevos experimentos o modelos para evaluarlas. Dichos experimentos y modelos generarán, en última instancia, interpretaciones estadísticas y ecológicas que servirán para comprender mejor cómo funciona el mundo natural que nos rodea. Figura adaptada de Fortin y Dale (2005).

## 1.2. SOBRE LA NECESIDAD DE ESTE LIBRO

Debido a la importancia del análisis espacial en el tratamiento de datos en ecología, geografía, geología, ingeniería y ciencias afines, y a la larga tradición de su uso, no es de extrañar que existan en la actualidad numerosas monografías que revisen las distintas técnicas de análisis espacial disponibles, su base matemática y sus aplicaciones prácticas (Diggle 1983, Cliff y Ord 1981, Rossi *et al.* 1992, Upton y Fingleton 1985, Greig-Smith 1983, Isaaks y Srivastava 1989, Cressie 1993, Goovaerts 1997, Legendre y Legendre 1998, Dale 1999, Fortin y Dale 2005 y Webster y Oliver 1990, por citar algunas de las más significativas). No obstante, la mayor parte de estas obras han surgido de manera independiente en distintas disciplinas, por lo que están centradas en técnicas concretas comúnmente utilizadas en un área científica determinada, como la geostatística en el caso de las ciencias de la tierra (Goovaerts 1997) y los métodos basados en las localizaciones puntuales de las plantas en el caso de la ecología vegetal (Diggle 1983). Sólo unas pocas obras proporcionan una visión general y equilibrada de los distintos métodos de análisis espacial disponibles (Dale 1999, Fortin y Dale 2005) y todavía no se ha publicado una síntesis global, aunque se están produciendo avances en esta dirección (Dale *et al.* 2002, Fortin y Dale 2005). Asimismo, las distintas monografías disponibles difieren considerablemente en su rigor matemático y estadístico, y muchas de ellas no son fácilmente accesibles a la mayoría de usuarios potenciales en las áreas de la ecología y las ciencias ambientales (el monumental tratado de Cressie 1993, excelente por otra parte, sería un buen ejemplo de ello).

Independientemente de que abarquen un mayor menor número de técnicas, de que estén centradas en una disciplina concreta y de que su rigor matemático/estadístico las haga más o menos accesibles, una característica común a la mayoría de las obras publicadas hasta la fecha sobre el análisis de datos espacialmente explícitos es que están escritas en inglés. Existen pocas monografías y revisiones sobre el tema en español (Nolasco y Orts 1991, Gámir *et al.* 1995, Samper y Carrera 1990, Blázquez 1984, Mateu y Morell 2003, Moral 2003, 2004) y hasta la fecha ninguna de ellas proporciona una visión completa e integradora de las distintas técnicas existentes para el análisis de este tipo de datos, ni está centrada en los datos y problemas que suelen generarse en Ecología y disciplinas relacionadas. Una notable excepción a este vacío editorial lo constituye el monográfico sobre ecología espacial publicado en la revista electrónica de divulgación científica *Ecosistemas* (<http://www.revistae-cosistemas.net>) en septiembre de 2006 (número 2006/3). Dicho monográfico,

que consta de diez artículos conceptuales y metodológicos, supuso un hito importante en la divulgación en español del análisis de datos ecológicos espacialmente explícitos. No obstante, el hecho de que quedaran numerosos temas por tratar en dicho monográfico, y de que no fueran tratados en profundidad por las limitaciones necesariamente impuestas por una revista de divulgación científica, hacían necesaria la realización de una obra en la que se presentaran, de una forma detallada, las principales herramientas y conceptos de interés para el análisis espacial de datos en Ecología y disciplinas relacionadas, obra todavía inexistente en la literatura impresa en español.

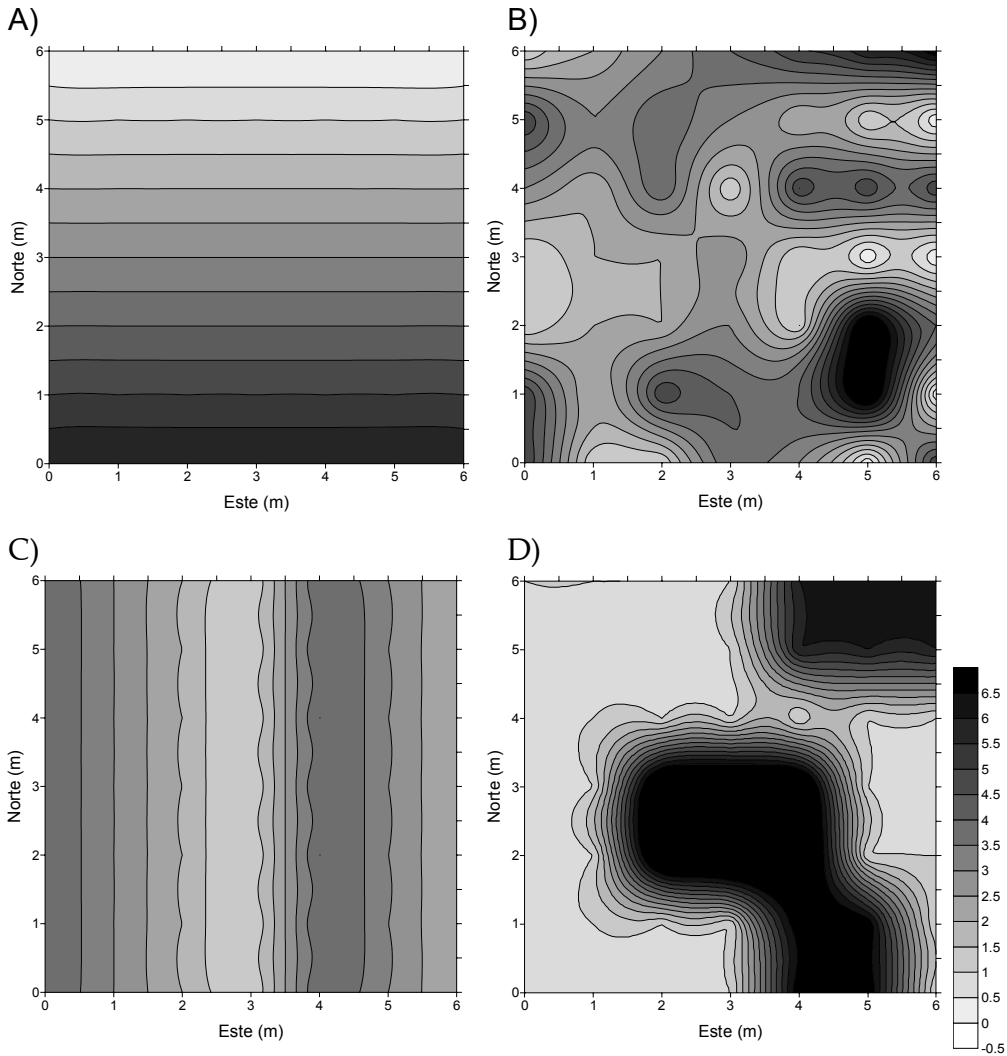
Este libro nace con la vocación de contribuir a llenar este vacío. Nuestra pretensión es que sirva tanto de introducción para el estudiante de Ecología y disciplinas relacionadas que se inicia en el mundo del análisis espacial como de referencia para el profesor, investigador o profesional que requiera una puesta al día, que quiera conocer las posibilidades analíticas existentes o que esté interesado en explorar nuevas técnicas. En este sentido, se ha puesto especial énfasis en presentar, de una manera rigurosa pero eminentemente práctica a la vez, las principales herramientas conceptuales, estadísticas e informáticas disponibles para el análisis de datos ecológicos espacialmente explícitos, ilustrando su utilización con multitud de ejemplos provenientes de estudios realizados en ecosistemas de la Península Ibérica, y en aportar toda aquella información adicional necesaria para facilitar al lector la utilización correcta de las distintas herramientas y conceptos presentados (mediante los códigos, programas y ejemplos incluidos en el CD adjunto). En cualquier caso, creemos que la obra supera también las limitaciones de buena parte de los libros disponibles en inglés en este momento, ya que reúne todas las herramientas que en la actualidad son utilizadas o tienen un elevado potencial de uso en el ámbito de la ecología, presenta un carácter eminentemente práctico y los distintos capítulos que la integran han sido escritos por investigadores de reconocido prestigio y visibilidad internacional, que además son especialistas en la utilización de las distintas herramientas y conceptos presentados.

### 1.3. ALGUNOS CONCEPTOS BÁSICOS EN ECOLOGÍA ESPACIAL

En este apartado queremos presentar algunos conceptos que serán básicos para la comprensión de buena parte de la obra. En el apartado 1.1 se han introducido dos conceptos fundamentales subyacentes a toda investigación sobre las distribuciones espacio-temporales de organismos, varia-



bles ambientales y procesos ecológicos: **patrón espacial y heterogeneidad espacial**. El primero se ha asociado de forma intuitiva a la distribución en el espacio de los organismos, aunque de una manera más formal puede definirse como la relación o dependencia espacial existente entre los valores que toma una variable en distintas localidades. Así pues, una determinada variable o proceso ecológico puede presentar, a grandes rasgos, un patrón espacial agregado, uniforme, en gradiente o aleatorio (Fig. 1.2).



**Figura 1.2.** Ejemplos de patrones espaciales en gradiente (A), aleatorio (B), regular (C) y agregado (D) de una variable continua. Los valores de la escala no tienen unidades

Respecto al concepto de heterogeneidad, debido a las numerosas acepciones que presenta y a los distintos aspectos de la misma identificados en la literatura ecológica (véanse los trabajos de Kolasa y Rollo 1991, Dutilleul y Legendre 1993, Li y Reynolds 1995 y Wiens 2000 para revisiones sobre el tema), es importante definirlo con claridad. Siguiendo a Adler *et al.* (2001), cuando se habla de la heterogeneidad desde un punto de vista genérico se la considera como sinónimo de variabilidad, entendida ésta como los cambios en los valores de un determinado descriptor en el espacio o en el tiempo (Kolasa y Rollo 1991). En este contexto, la heterogeneidad espacial hace referencia a la variabilidad en el espacio de una determinada propiedad o proceso ecológico bajo estudio, y puede cuantificarse utilizando técnicas estadísticas paramétricas como el coeficiente de variación. Cuando esta variabilidad se encuentra estructurada espacialmente, se utiliza heterogeneidad como sinónimo de patrón espacial. Así pues, y en función del contexto que se considere, heterogeneidad se emplea como sinónimo de variabilidad o de patrón espacial. Dado que el objeto principal de este libro son los datos ecológicos espacialmente explícitos, y que éstos suelen tener una fuerte estructura espacial en la mayor parte de casos, se utiliza de forma general heterogeneidad espacial como sinónimo de patrón espacial.

Otro concepto de vital importancia para el análisis de datos espacialmente explícitos es el de **escala**, que es utilizado por los ecólogos para referirse a distintos conceptos relacionados con la extensión en el espacio de los procesos bajo estudio y la resolución espacial y temporal de los datos (Fortín y Dale 2005). La detección e interpretación de los patrones espaciales de organismos y procesos ecológicos se encuentran fuertemente influidos por la escala de observación (Levin 1992; Maestre *et al.* 2005a), por lo que, junto a la utilización de una herramienta analítica adecuada, el uso de una escala que permita evaluar correctamente el fenómeno objeto de estudio es un aspecto fundamental en el análisis de datos ecológicos espacialmente explícitos. El capítulo 2 está íntegramente dedicado a la escala, por lo que se remite a todos los lectores a dicho capítulo para una discusión en profundidad sobre este concepto.

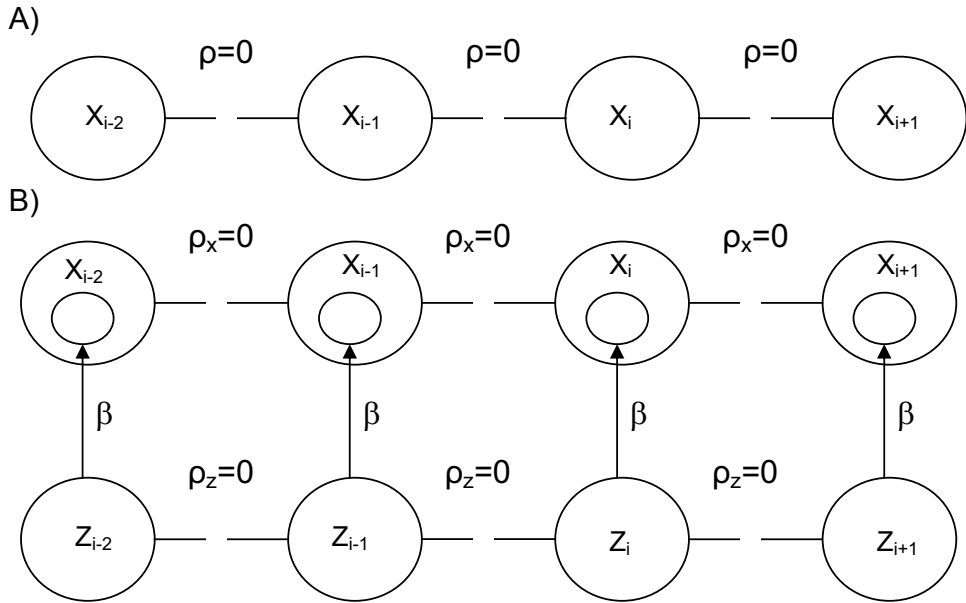
Como se podrá comprobar a lo largo de estas páginas, las principales herramientas de las que disponemos para interpretar los patrones espaciales que se observan en los sistemas naturales se basan en una serie de suposiciones sobre los procesos que han generado dichos patrones. Numerosos

métodos de análisis espacial se basan en la suposición que el proceso subyacente a un patrón espacial dado es **estacionario**. En un contexto espacial, un proceso o su modelo es estacionario si sus propiedades son independientes de su posición absoluta y dirección en el espacio (Haining 1990) o, dicho de otra forma, si sus parámetros (como su media y varianza) son los mismos en todas las partes de la zona de estudio y en todas las direcciones. La estacionaridad de los datos es un requisito de distintas técnicas analíticas que, como la geoestadística, caracterizan la estructura espacial de los datos en zonas del espacio que no son muestreadas. En el Capítulo 6 se discuten en mayor profundidad las implicaciones de la estacionaridad para el análisis de datos ecológicos utilizando técnicas geoestadísticas.

#### 1.4. LA AUTOCORRELACIÓN DE LOS DATOS ECOLÓGICOS: DEFINICIÓN Y PROBLEMÁTICA

En apartados anteriores se ha puesto de manifiesto que los organismos, las variables abióticas y los procesos ecológicos rara vez presentan distribuciones aleatorias o uniformes en la naturaleza. Pese a ello, los ecólogos hemos sido formados durante décadas en la creencia de que la naturaleza sigue las suposiciones de la estadística clásica o paramétrica, siendo una de ellas la independencia de los datos (observaciones de una población muestral son independientes cuando ninguna de ellas tiene una influencia sobre otra, Fig. 1.3). Imaginemos las consecuencias de un ecosistema en el que los componentes que lo forman y los factores abióticos que lo afectan no presentaran estructura espacial: los organismos no estarían cerca de su comida, los individuos de una especie del sexo opuesto se distribuirían al azar a lo largo del paisaje, las propiedades edáficas en las inmediaciones de una planta no serían más apropiadas para sus plántulas que aquellas situadas en otro lugar, etc. Pese a que esta visión parecerá poco o nada realista a la mayor parte de lectores a la luz de los conocimientos de los que disponemos en la actualidad, hay que destacar que constituye una suposición básica de la mayor parte de teorías y modelos que describen el funcionamiento de poblaciones y comunidades (Legendre y Legendre 1998).

Cuando los datos no son independientes se dice que presentan autocorrelación. En el contexto objeto del libro, una variable presenta autocorrelación espacial cuando sus valores en lugares separados por una cierta distan-

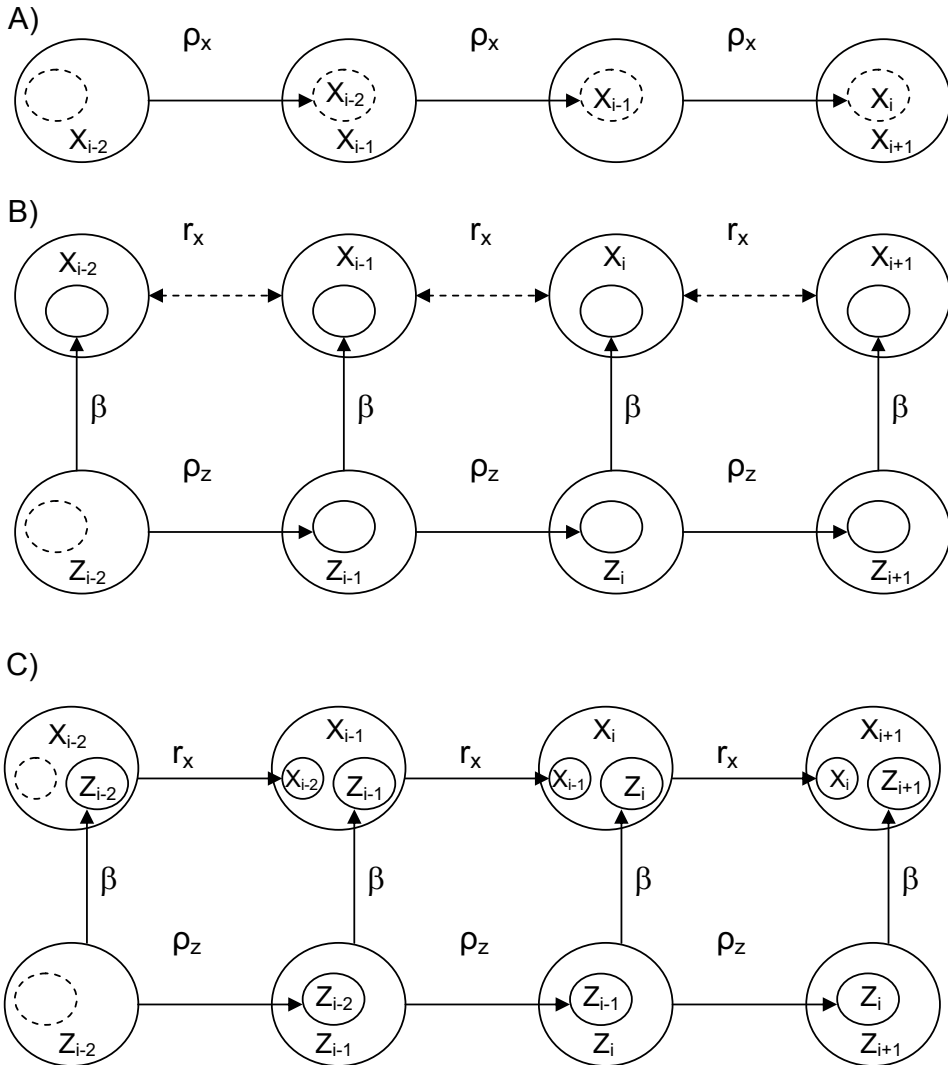


**Figura 1.3.** Representación gráfica de la independencia entre datos espacialmente adyacentes. En la parte superior (A), se muestra una serie de valores de la variable  $X$  espacialmente independientes entre sí, por lo que la correlación ( $\rho$ ) es 0. En la parte inferior (B) se muestra una serie de valores de la variable  $X$  dependientes de otra variable ( $Z$ ), cuyos valores son independientes entre sí ( $\rho_z = 0$ ); los valores de la variable  $X$  siguen siendo espacialmente independientes, por lo que la correlación esperada entre valores adyacentes ( $\rho_x$ ) es 0. Adaptada de Fortin y Dale (2005).

cia son más (autocorrelación positiva) o menos (autocorrelación negativa) similares que lo que cabría esperar según una distribución aleatoria de la misma (Fig. 1.4). Las causas de la autocorrelación en los datos ecológicos pueden ser diversas. Así, la abundancia de una determinada especie puede estar espacialmente autocorrelada como consecuencia de los mecanismos de dispersión o movilidad. Un ejemplo de ello lo constituyen los patrones de dispersión de las semillas de muchas especies vegetales, que suelen seguir patrones ajustables a funciones del tipo exponencial negativo, donde el número de semillas dispersadas disminuye de forma exponencial con la distancia a la planta madre (véase el Capítulo 12 para un ejemplo). Determinados autores denominan a este tipo de fenómeno autocorrelación inherente o verdadera (Fortin y Dale 2005; Fig. 1.4A). En otras ocasiones, la distribución de la variable que estamos estudiando depende de otra variable que presenta autocorrelación espacial, lo que deriva en que los valores de la mis-

ma no sean espacialmente independientes. Un ejemplo de ello lo constituye la mortalidad de plantaciones de arbustos en zonas semiáridas, que está fuertemente influenciado por los patrones espaciales de distintas variables edáficas relacionadas con la disponibilidad de agua en el suelo (Maestre *et al.* 2003a). A este tipo de autocorrelación se le denomina dependencia espacial inducida (Fortin y Dale 2005; Fig. 1.4B). También puede darse el caso de que una variable presente a la vez autocorrelación inherente e inducida (Fig. 1.4C). Es importante señalar que la autocorrelación es escala o distancia dependiente; es decir los valores de una variable suelen ser más (autocorrelación positiva) y menos (autocorrelación negativa) similares de lo esperado a distancias cortas y largas, respectivamente.

Como puede imaginarse, la autocorrelación espacial es “el pan nuestro de cada día” para los ecólogos. En un influyente artículo, Pierre Legendre (Legendre 1993) planteaba la pregunta de si la autocorrelación es un problema o una nueva ventana de oportunidad para el conocimiento de la importancia de la estructura espacial en los procesos ecológicos. La respuesta es clara: ambas cosas. Por una parte, la falta de independencia entre las observaciones derivada de la autocorrelación espacial plantea numerosos problemas cuando se utilizan pruebas estadísticas paramétricas y no paramétricas, como el análisis de varianza y la regresión lineal, por poner unos ejemplos (conviene mencionar que los análisis estadísticos no paramétricos, como las pruebas de Kruskal-Wallis o de Mann-Whitney, también requieren la independencia de las observaciones, aspecto frecuentemente olvidado por los investigadores que las utilizan). Cuando una variable está autocorrelacionada espacialmente, es decir, cuando un valor determinado en un punto no es completamente independiente de los valores que toma la variable en localidades vecinas, la cantidad de información en el conjunto de datos se reduce en proporción al grado de autocorrelación que presenta. En otras palabras, en presencia de autocorrelación espacial cada unidad muestral no contribuye con un grado de libertad a la prueba estadística, si no con algo menos en función del nivel de autocorrelación. A modo de ejemplo, si la autocorrelación es tan intensa que permite predecir el resto de valores de un conjunto de datos a partir de uno solo, entonces el tamaño de muestra efectivo sería realmente la unidad (esto es un ejemplo extremo que difícilmente ocurriría en la práctica). Por otra parte, no hay que olvidar que la autocorrelación es una característica intrínseca a los datos ecológicos y que si no fuera por ella difícilmente podrían hacerse predicciones (los ecólogos estaríamos en serios



**Figura 1.4.** Representación gráfica de datos espacialmente explícitos que presentan autocorrelación inherente (A), inducida (B) y de ambos tipos (C). Los valores de la serie mostrada en A) no son independientes; cada valor depende directamente de su predecesor e indirectamente de todos los que le preceden. En estos valores la correlación esperada entre valores no es 0, sino  $\rho$ . En el ejemplo mostrado en B), los valores de la variable X no son independientes, pero la autocorrelación que presentan,  $r_x$ , está inducida por su dependencia lineal con los de la variable Z, que presentan a su vez autocorrelación inducida. En la serie mostrada en C) la correlación de los valores de la variable X viene de dos fuentes: inherente en la propia variable X e inducida por la variable Z, dando lugar a una situación de autocorrelación doble (inherente e inducida). Adaptada de Fortin y Dale (2005)

apuros si los fenómenos y procesos ecológicos fueran espacial- y temporalmente independientes!).

La caracterización de la autocorrelación espacial requiere la utilización de técnicas de análisis espacial, como las descritas en los capítulos 3, 4, 5 y 6. Cuando el objetivo del estudio requiera la utilización de técnicas de análisis clásicas, como una correlación o un ANOVA, hay que tener en cuenta los efectos de dicha autocorrelación en los resultados de los análisis. Ello puede realizarse mediante la corrección del nivel de significación para tener en cuenta la pérdida de información (grados de libertad), como el método de Dutilleul (1993a) para evaluar la correlación entre variables autocorreladas o el de Dale *et al.* (1991a) para analizar tablas de contingencia múltiples en presencia de autocorrelación, mediante la utilización de técnicas de permutaciones que incorporan la estructura espacial de los datos (Fortin y Jacquez 2000) o utilizando la información espacial de los datos explícitamente en los análisis (p. ej. como covariable), tal como se describe en los capítulos 8 y 12.

La autocorrelación espacial tiene también gran importancia para el diseño y análisis de experimentos, ya que puede condicionar, e incluso invalidar, los resultados derivados de los mismos (Legendre *et al.* 2002, 2004). La heterogeneidad espacial tradicionalmente se ha tenido en cuenta a la hora de diseñar y ejecutar experimentos mediante la utilización de diseños como el de bloques aleatorizados, que intentan controlar en los correspondientes modelos esa fuente de variación (véase Quinn y Keough 2002 para una revisión sobre los principales tipos de diseños experimentales y su análisis). No obstante, en numerosas ocasiones estos diseños no son capaces de “filtrar” la autocorrelación espacial de aquellas variables ambientales que pueden afectar a la variable/organismo/proceso que estamos evaluando, normalmente por problemas de escala y de resolución del experimento. Esto puede confundir el efecto de los tratamientos de un experimento dado e invalidar sus resultados (Zas 2006a). La autocorrelación y su problemática para el diseño y análisis de experimentos son tratados con mayor detalle en los capítulos 14 y 27, por lo que se remite a los lectores interesados a los mismos para una discusión en profundidad de estos aspectos. Asimismo, para todos aquellos interesados en una completa revisión sobre la autocorrelación y su problemática en el análisis espacial de datos ecológicos, la lectura del tratado de Fortin y Dale (2005) resulta imprescindible.

## 1.5. CONSIDERACIONES SOBRE LA ADQUISICIÓN DE DATOS ECOLÓGICOS ESPACIALMENTE EXPLÍCITOS

La adquisición de datos constituye uno de los primeros pasos de todo estudio ecológico, siendo una etapa fundamental que no siempre recibe la atención que merece. Resulta curioso comprobar cómo la mayor parte de manuales de estadística y análisis espacial dedican buena parte de sus páginas a la descripción de las distintas técnicas y métodos con profusión, así como a la interpretación de los datos una vez analizados. No obstante, aspectos como la descripción de los distintos tipos de muestreo, su idoneidad e implementación en la realidad son frecuentemente olvidados en este tipo de obras de referencia (notables excepciones son los manuales de Webster y Oliver [1990], Dale [1999] y Fortín y Dale [2005]).

El establecimiento de un protocolo de muestreo adecuado es fundamental en cualquier estudio sobre el análisis espacial de organismos y procesos ecológicos, ya que muchas de las técnicas de análisis espacial requieren datos adquiridos de una forma particular y, lo que es más importante, todos los pasos y decisiones que se tomen antes y durante el muestreo afectarán a la interpretación de los patrones detectados (Fortin et al. 1989, Jelinski y Wu 1996, Qi y Wu 1996, Dungan et al. 2002). Así pues, los primeros aspectos a considerar cuando se planifica la toma de datos son la pregunta/hipótesis que pretendemos responder/evaluar con nuestro estudio, la escala a la que deseamos trabajar (véase el Capítulo 2 para información detallada sobre la importancia de la escala en el análisis espacial de datos ecológicos) y el tamaño/ecología de los organismos a estudiar (en el caso de que nuestro trabajo se centre en algún organismo en concreto). Una vez que estos aspectos han sido definidos, conviene tener en cuenta el tipo de análisis a utilizar, así como cuestiones relativas al tamaño de la zona de estudio, el número de puntos de muestreo que se pueden llevar a cabo y la distancia entre los mismos, en el caso de que se utilicen unidades de muestreo discretas, entre otros aspectos. A continuación se presentan los tres principales tipos de datos espacialmente explícitos que se suelen obtener en los estudios ecológicos, así como distintas recomendaciones para su muestreo.

### 1.5.1. Datos en forma de puntos

Los datos en forma de puntos, en los que se representa la posición de objetos discretos en una zona limitada mediante sus coordenadas (en el



caso más general bidimensionales:  $x, y$ ), constituyen una de las principales fuentes de información empleadas en el análisis espacial de datos ecológicos (Fig. 1.5). Este tipo de datos se utiliza fundamentalmente en el estudio de organismos y estructuras sésiles que pueden ser representados mediante un punto en un plano (p. ej. plantas, gasterópodos como las lapas, hormigueros, etc.). Si bien simplifican la información espacial a sus coordenadas espaciales, este tipo de datos se han utilizado ampliamente para el análisis espacial de árboles (Camarero *et al.* 2000b, Arévalo y Fernández-Palacios 2003, Camarero y Rozas 2006, Fajardo *et al.* 2006), plantas herbáceas en forma de roseta (Silander y Pacala 1985), plantas clonales con tallos erectos (Hutchings 1979) y arbustos (Haase 1995, Haase *et al.* 1995, 1996, Escudero *et al.* 2005), en los que el tronco principal puede considerarse como su posición. Este tipo de datos se analizan con técnicas de análisis de patrones de puntos, descritas con detalle en el Capítulo 3.

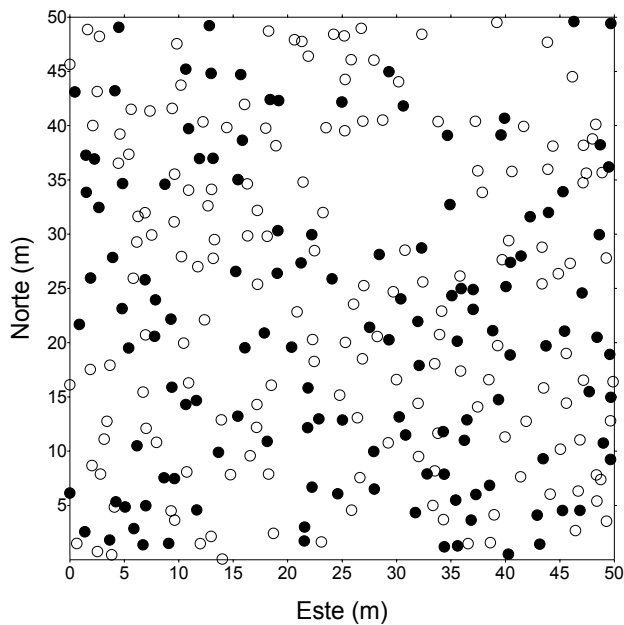
Los datos en forma de puntos presentan una ventaja importante respecto a aquellos obtenidos mediante unidades muestrales discretas, ya que, al obtener información sobre todos los individuos/estructuras a estudiar dentro de la zona de estudio, su muestreo no requiere tomar decisiones *a priori* relativas al tamaño, distancia y forma de las unidades muestrales. No obstante, las dimensiones de la zona de estudio deben ser consideradas antes de iniciar un muestreo para tomar datos en forma de puntos. Distintos autores (O'Neill *et al.* 1996, 1999) han sugerido que dichas dimensiones deberían ser entre dos y cinco veces mayores que el tamaño del proceso o fenómeno bajo estudio. Según este criterio, si estamos estudiando agregaciones entre plantas que ocurren en forma de manchas de 2 m de diámetro, nuestra zona de estudio debería de tener un tamaño mínimo de  $4 \times 4$  m. Este tipo de recomendaciones no suelen tener mucho valor práctico cuando se conduce un estudio por vez primera, ya que rara vez se conocen *a priori* estas características de los procesos que estamos estudiando, pero pueden ser de gran utilidad cuando se dispone de información preliminar.

Por lo que respecta a cuestiones prácticas relativas al muestreo, se pueden proporcionar algunas recomendaciones en función del tamaño de los individuos/estructuras a estudiar y de las dimensiones del área de estudio. Para parcelas de pequeño tamaño, inferiores a  $2 \text{ m}^2$ , puede obtenerse directamente la localización espacial de las especies utilizando acetatos transparentes (Fig. 1.6). Para parcelas de tamaño superior, posiblemente el sistema

A)



B)



**Figura 1.5.** Aspecto de una formación de matorral semiárido en Nuevo México, Estados Unidos (A). Este tipo de formaciones vegetales son idóneas para la toma de datos en forma de puntos. En B se muestra la distribución de dos especies de arbustos en una zona semiárida de Alicante (cada una representada por un círculo de cada color). En este mapa, cada punto representa la localización de un individuo. Autor de la fotografía: Fernando T. Maestre

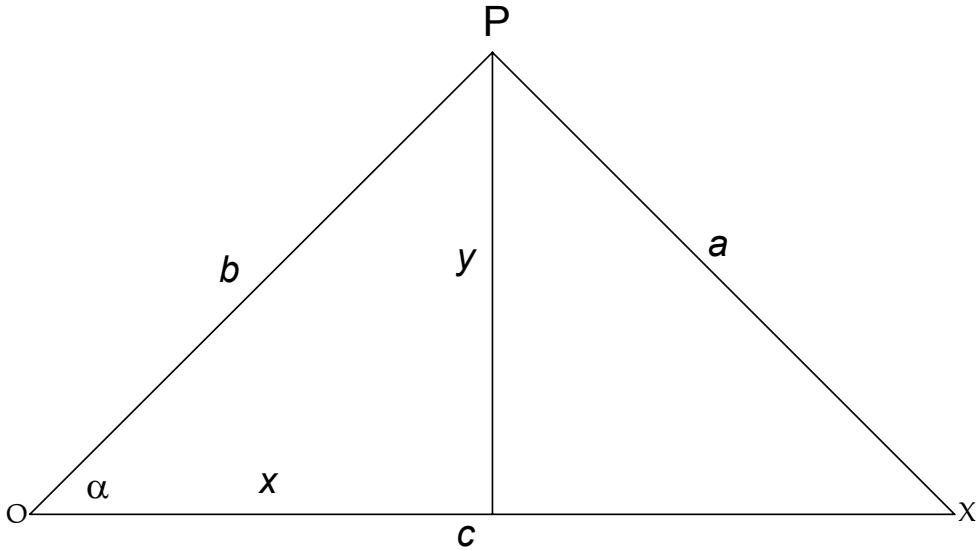
más preciso y rápido para obtener este tipo de datos sea el utilizar una estación total (instrumento electro-óptico utilizado en topografía que incorpora un medidor de distancias y un microprocesador acoplados a un teodolito electrónico) o un localizador GPS (“Global Positioning System”) de alta resolución para posicionar los individuos a estudiar. No obstante, estos instrumentos suelen presentar problemas de utilización en zonas boscosas, ya que la estación total exige que exista una línea visual entre el aparato que emite y recibe la señal láser y el prisma, que se coloca sobre cada uno de los individuos a muestrear, y por otra parte los receptores GPS pueden tener dificultades a la hora de recibir la señal de los satélites. En estos casos puede establecerse un punto de referencia y medir las distancias de los puntos muestreados entre sí y respecto a dicha referencia. Si se conoce la distancia entre dos puntos fijos, O y X, y entre éstos y el punto de referencia (P), puede utilizarse el método del coseno para obtener las coordenadas cartesianas ( $x, y$ ) de cada punto calculando el ángulo POX (Fig. 1.7):

$$\alpha = \cos^{-1}[(b^2 + c^2 - a^2)/2bc], \text{ donde } x = b \cos \alpha \text{ y } y = b \sin \alpha \quad (1.1)$$

Véase Dale (1999) para otras fórmulas alternativas. Por último, si la zona de estudio es de gran tamaño pueden utilizarse fotografías aéreas o mapas digitales para obtener las coordenadas de los individuos a estudiar (Fig. 1.8).



**Figura 1.6.** Utilización de un acetato transparente para la toma de datos en forma de puntos (A). Esta técnica es apropiada cuando se trabaja con especies en forma de roseta de pequeño tamaño. Autor de la fotografía: María José Albert.

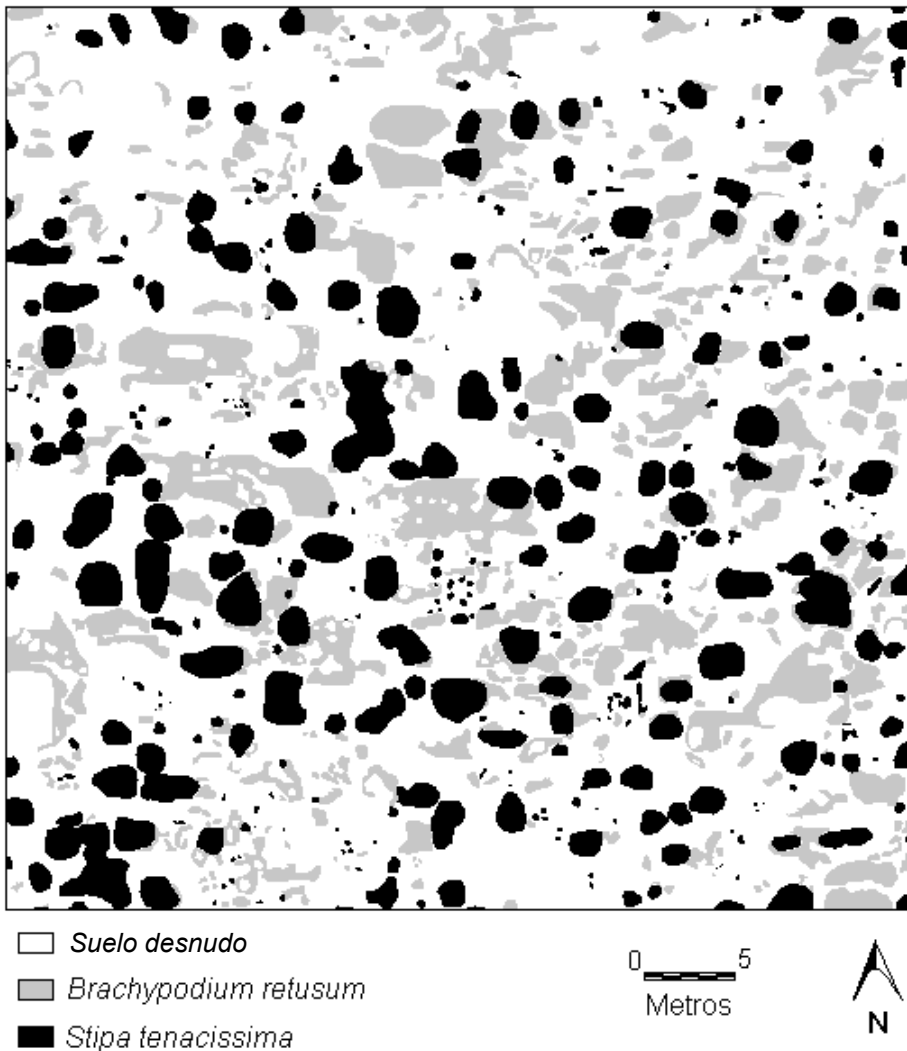


**Figura 1.7.** Obtención de las coordenadas de dos puntos de muestreo (O y X) utilizando la distancia entre ellos (de longitud  $c$ ) y un punto de referencia (P), que viene dada por las líneas  $a$  y  $b$ . El ángulo POX es  $\alpha$ . Para más detalles véase el texto. Adaptada de Dale (1999).

### 1.5.2. Datos obtenidos de unidades muestrales discretas

El establecimiento de programas de muestreo basados en unidades discretas (usualmente cuadrados o rectángulos) en los que se cuantifica la variable de interés constituye uno de los principales métodos para tomar datos ecológicos, especialmente en el caso de organismos como plantas de pequeño tamaño, líquenes y organismos bentónicos. De hecho, los métodos clásicos de análisis espacial basados en la varianza desarrollados en los años 50 y 60 (véase Greig-Smith 1983 y Dale 1999 para una revisión de los mismos) están basados en la toma de datos en cuadrados dispuestos de forma consecutiva. Con carácter general, para poder analizar mediante técnicas de análisis espacial los datos obtenidos de cuadrados de muestreo se debe de registrar la coordenada en el espacio de los mismos, de modo que cada cuadrado sería un punto de muestreo, o bien establecer un sistema de subcuadrados de muestreo dentro del cuadrado principal. Independientemente de que se opte por uno u otro sistema, en cada cuadrado/subcuadrado de muestreo se cuenta el número de individuos o su cobertura, de modo que se obtiene una estimación de la densidad de la especie

objeto de estudio. Este tipo de datos, formado por las coordenadas cartesianas ( $x, y$ ) y un valor de densidad, es la base de técnicas como el análisis espacial mediante índices de distancia (SADIE), descritas con detalle en los capítulos 4 y 5.



**Figura 1.8.** Mapa digitalizado de la vegetación en una parcela de 50 × 50 m situada en Aigües de Busot (Alicante), mostrando la distribución de las dos especies dominantes (*Stipa tenacissima* y *Brachypodium retusum*). Una versión más detallada en color puede encontrarse en el CD que acompaña al libro (carpeta "Capítulo 1", archivo "Figura 1.7.jpg"). Adaptada de Maestre y Cortina (2002).

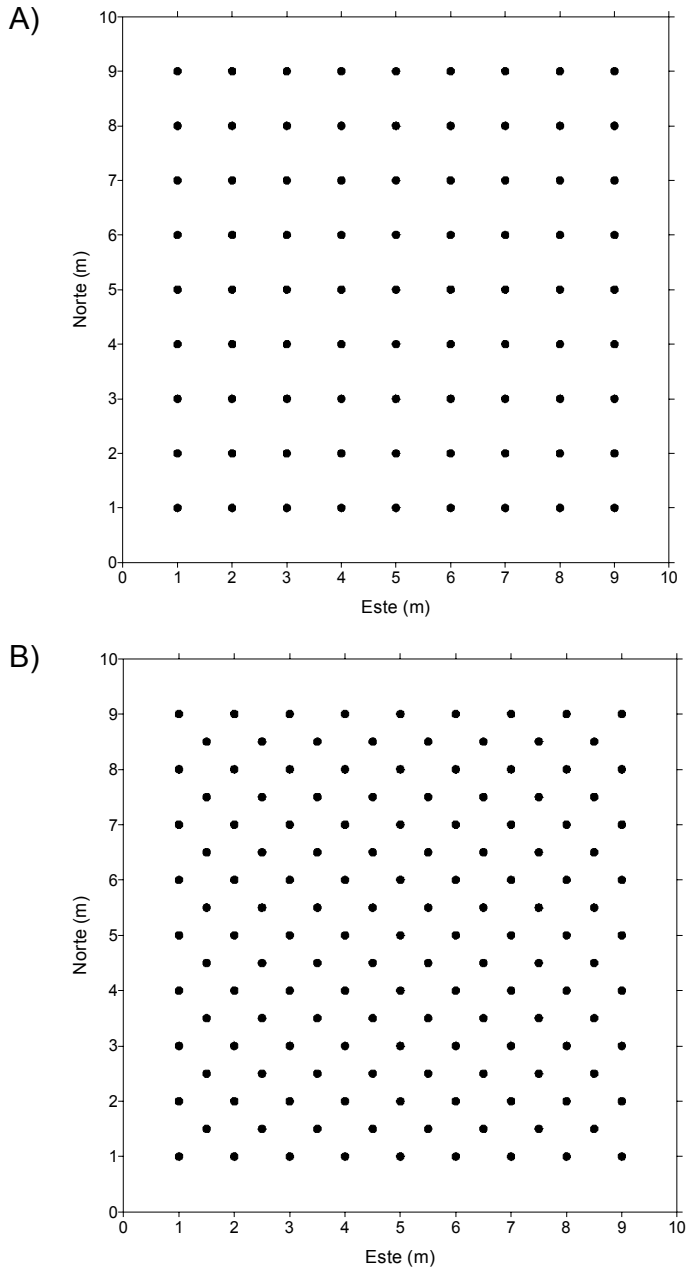
Cuando se utilizan unidades muestrales discretas hay que tener en cuenta que los patrones espaciales que observemos, así como su interpretación, van a depender de factores como el tamaño del área de estudio (los comentarios al respecto hechos en el apartado anterior son perfectamente válidos aquí), el número de unidades a muestrear, su forma, su tamaño y su disposición en el espacio (de forma aleatoria, estratificada o regular). Es por ello que estos aspectos deben tenerse muy en cuenta a la hora de establecer los protocolos de muestreo. Si bien resulta complicado establecer unas sugerencias generales que puedan ser efectivas bajo un abanico amplio de condiciones, se presentan a continuación algunas recomendaciones a tener en cuenta a la hora de diseñar un protocolo de muestreo basado en el uso de cuadrados, por ser el tipo de unidad de muestreo discreta más comúnmente utilizada en Ecología.

El tamaño del cuadrado define la resolución espacial mínima a la que se obtienen los datos y a la que una estructura espacial puede ser caracterizada. Es decir por debajo de esa escala no hay información espacial. Por ejemplo, si utilizamos cuadrados de muestreo de  $50 \times 50$  cm, no podrán detectarse procesos que ocurran a escalas inferiores a las dimensiones del cuadrado (p. ej. agregaciones de individuos a distancias inferiores a 15 cm). Así pues, la elección del tamaño es un paso fundamental en todo análisis espacial basado en datos a generar con estas unidades muestrales. Si se pretende evaluar el patrón espacial de un organismo dado, hay que tener en cuenta sus dimensiones y características ecológicas. El tamaño del cuadrado de muestreo debería ser más grande que el de un individuo de la especie de interés y del mismo tamaño, o preferiblemente menor, que las estructuras que resultan de un proceso unitario a detectar como, por ejemplo, una mancha o agrupación de organismos (Fortin *et al.* 1989). Este aspecto cobra especial relevancia si el objetivo del estudio es inferir la presencia de procesos como las interacciones bióticas a través de los patrones espaciales (Haase 2001, Maestre 2003a, Alados *et al.* 2006). Si el tamaño del cuadrado de muestreo es muy grande se detectarán asociaciones positivas debido a que especies con los mismos requerimientos de hábitat aparecerán en las mismas unidades de muestreo, no correspondiendo estas asociaciones a interacciones reales positivas o de facilitación. Si, por el contrario, es muy pequeño, la presencia de pocos individuos puede limitar la detección de asociaciones significativas entre las especies (Jonsson y Moen 1998). Cuando se estudian comunidades con especies de distinto tamaño y morfología, el uso de cuadrados de distinto tamaño (Maestre *et al.* 2005a)

es una buena opción de muestreo. No obstante, hay que tener en cuenta que la utilización de cuadrados de muestreo muy grandes no es muy práctica debido a la dificultad de obtener estimas de densidad precisas sin que el esfuerzo sea inaceptable. En el caso de que se requiera utilizar cuadrados muy grandes se recomienda subdividir estos en unidades contiguas más pequeñas, que siempre se podrán agregar *a posteriori*.

Otro aspecto importante a considerar cuando se utilizan cuadrados de muestreo es su localización en el espacio. La decisión de establecer un muestreo en una (p. ej. transecto) o dos (p. ej. superficie) dimensiones depende de los objetivos del estudio, de las variables a evaluar, de los recursos de los investigadores y de las características del área estudiada. Por ejemplo, si estamos evaluando los cambios en el patrón espacial a lo largo de un gradiente unidireccional (p. ej. la pendiente en una zona rocosa), la disposición contigua de cuadrados de muestreo en la dirección de este gradiente puede ser una buena opción de muestreo. Si no se dispone de información previa sobre la especie/variable a estudiar, o sobre si ésta presenta una determinada direccionalidad, el establecimiento de una malla bidimensional de puntos de muestreo dispuesta de forma regular suele representar el mejor compromiso de muestreo (Fig. 1.9A). Independientemente de la estrategia de localización de cuadrados que se siga, hay que tener en cuenta que la distancia entre cuadrados de muestreo debería ser menor que la distancia media entre las estructuras a detectar (e.g. agregaciones de individuos, Legendre y Legendre 1998). Cuando se desconoce dicha distancia, el establecimiento de protocolos de muestreo anidados (Fig. 1.9B) puede ser una buena opción para maximizar las posibilidades de lograr los objetivos del estudio.

Por lo que respecta al número de unidades a establecer (tamaño muestral, la famosa  $n$ ), esta cifra va a depender, nuevamente, de los objetivos del estudio, así como de cuestiones logísticas relativas a la accesibilidad, esfuerzo de muestreo, dimensiones del área de estudio, etc. En el contexto de análisis espacial debe tenerse en cuenta también los requerimientos mínimos de las técnicas analíticas que se tenga previsto utilizar. Por ejemplo, se ha recomendado un tamaño mínimo de 30 unidades de muestreo para poder detectar de forma significativa la presencia de autocorrelación espacial (Legendre & Fortin 1989). No obstante, autores como Webster & Oliver (1992) recomiendan utilizar al menos 100 puntos de muestreo si se quiere estimar funciones de autocorrelación como los semivariogramas



**Figura 1.9.** Ejemplos de mallas de muestro dispuestas de una forma regular. Las mallas representadas en A) y en B) permiten detectar patrones espaciales con una resolución mínima de 1 y 0,5 m, respectivamente.



(véase el Capítulo 6 para una descripción detallada del mismo) con cierta fiabilidad. Si bien obtener un número de réplicas elevado es costoso y en numerosas ocasiones simplemente no es posible, nuestra experiencia con la utilización de distintas técnicas de análisis es que números del orden de 30-40 puntos de muestreo resultan insuficientes para detectar patrones espaciales con ciertas garantías. No obstante, hay que resaltar la dificultad de dar una recomendación general respecto al número de muestras a utilizar que sea válida bajo un amplio abanico de situaciones. Por ejemplo si lo que queremos detectar son patrones unidimensionales, por ejemplo a lo largo de un transecto, quizás unos valores como los presentados arriba pueden ser suficientes. En la medida que sea posible, nuestro consejo es que se utilicen tantos puntos de muestreo como sea posible abordar en el estudio, empleando siempre el sentido común y teniendo en cuenta sus objetivos. Una mayor  $n$  va a proporcionarnos mayores opciones de análisis, así como mayor potencia en los análisis que realicemos.

Como se ha puesto de manifiesto en los párrafos anteriores, la utilización de cuadrados de muestreo requiere la toma de decisiones *a priori* de vital importancia para el devenir posterior de los análisis a realizar con los datos obtenidos. Para reducir la subjetividad en este proceso es necesario tener un conocimiento claro sobre la biología de los organismos a estudiar y la escala a la que operan los procesos a detectar, lo que pone de manifiesto la importancia de una buena revisión bibliográfica previa al establecimiento de un nuevo muestreo. Si no se dispone de información previa, el establecimiento de un estudio piloto puede ser de gran utilidad para obtener dicha información y maximizar que el muestreo permita responder a los objetivos del estudio.

### 1.5.3. Datos de variables continuas obtenidos en puntos

Se trata de una situación a caballo entre las dos anteriores. Es información obtenida en puntos, no en unidades muestrales discretas, pero que lleva asociada información relativa a una variable cuantitativa. A modo de ejemplo del tipo de datos que se incluirían en este apartado podemos indicar algún parámetro edáfico o el nivel de  $\text{CO}_2$  medido en una torre de *eddy covariance*. Estos datos, formado por las coordenadas cartesianas  $(x, y)$  y un valor para una variable continua, es la base de técnicas como la geoestadística.

Buena parte de los comentarios que hemos llevado a cabo para los datos procedentes de unidades muestrales discretas son aplicables para este tipo de datos, ya que en realidad en éstos la información se condensa espacialmente en un punto al que se le asigna una densidad. Es por ello que algunos autores consideran esta última categoría como un tipo de las unidades muestrales discretas, pero nosotros creemos que su naturaleza es manifiestamente diferente y que las técnicas que se han aplicado tradicionalmente a cada tipo han sido también distintas.

#### 1.6. ALGUNAS RECOMENDACIONES PRÁCTICAS PARA ABORDAR CON ÉXITO EL ANÁLISIS DE DATOS ECOLÓGICOS ESPACIALMENTE EXPLÍCITOS

Como hemos discutido en apartados anteriores, y como podrá descubrir el lector en sucesivos capítulos de esta obra, la existencia de distintos tipos de datos espacialmente explícitos, la variedad de herramientas de análisis disponibles, las distintas decisiones que hay que tomar a la hora de planificar los muestreos, el tipo de información que queremos extraer de ellos y las hipótesis a evaluar hacen difícil el establecer una serie de recomendaciones de carácter general sobre cómo debe hacerse un estudio ecológico que incluya el análisis de datos espacialmente explícitos. No obstante, sí que se puede establecer una especie de “protocolo de actuación” que pueda ser utilizado en la mayor parte de casos en las que nos enfrentemos a este tipo de estudios, y que ayude a ejecutarlos de forma satisfactoria. En nuestra opinión, cualquier estudio que tenga entre sus objetivos el análisis espacial de datos ecológicos debería de seguir estos pasos:

1) *Plantear claramente las preguntas a responder con el estudio, así como las hipótesis de trabajo.* Con frecuencia nos encontramos casos de estudiantes e investigadores que, una vez que han tomado sus datos, se preguntan qué hacer con ellos, así como estudios en los que se pretende “ajustar” los datos obtenidos a hipótesis o preguntas que no pueden ser respondidos porque no han sido obtenidos mediante protocolos de muestreo diseñados o ejecutados de forma apropiada. Ello deriva con frecuencia en la falta de publicación de muchos estudios, con el consabido gasto económico y la lógica frustración de las personas que lo han llevado a cabo, que ven como el trabajo realizado no ha dado los resultados esperados.

Este tipo de situaciones podrían evitarse fácilmente si antes de lanzarnos a tomar infinidad de datos (sin saber realmente para qué sirven o qué haremos con ellos) nos paramos a pensar con detenimiento qué es exactamente lo que pretendemos responder con nuestro estudio, identificamos claramente nuestras hipótesis de trabajo, revisamos el trabajo previo realizado sobre el tema y evaluamos los requisitos de las técnicas que pretendemos utilizar. La importancia que estos aspectos tienen en el diseño de cualquier estudio ecológico ha sido puesta de manifiesto por distintos autores (Gotelli y Ellison 2004). Asimismo, hay que destacar la tendencia entre las revistas científicas de ámbito ecológico, especialmente entre las que tienen un mayor índice de impacto (el famoso “impact factor” otorgado por el “Institute for Scientific Information”), hacia la publicación preferente de artículos que evalúan hipótesis específicas (Grogan 2005), lo que pone de manifiesto la importancia de identificar la hipótesis de trabajo previamente a la realización del mismo.

2) *Planear el muestreo conforme a los objetivos e hipótesis del estudio.* Si revisamos nuestros discos duros o cuadernos de campo, ¿cuántos datos encontraremos que no hemos utilizado en nuestras publicaciones, y que posiblemente no utilicemos jamás, no por no ser interesantes, sino más bien por ser redundantes o innecesarios? En muchos casos la toma de estos datos ha costado mucho tiempo y dinero, con el consiguiente despilfarro. La causa más probable de que esos datos no hayan acabado utilizándose es, en numerosas ocasiones, la ausencia de una planificación inicial de los muestreos.

Con el fin de optimizar recursos materiales, humanos y económicos, así como para maximizar el trabajo realizado, es fundamental establecer *a priori* el número de variables a incluir en nuestro estudio, así como las características del muestreo a realizar, haciendo uso de toda la información previa disponible. En la medida en que sea posible, conviene tomar datos únicamente de aquellas variables que sean fundamentales para responder a los objetivos de nuestra investigación, así como para evaluar las hipótesis en las que se sustenta. Si bien en determinados estudios de tipo observacional, o cuando no se dispone de información previa, no es fácil determinar un conjunto de variables clave que deban ser evaluadas, en general nos encontraremos con que existe abundante información sobre la especie/variable ambiental/proceso ecológico que estamos estudiando, por lo que una correcta planificación puede ahorrarnos muchas horas de campo, invernadero, ordenador o laboratorio. Una vez definidas las variables a

medir en nuestro estudio conviene establecer el protocolo de muestreo siguiendo las recomendaciones establecidas en el apartado 1.5.

3) *Explorar los datos antes de analizarlos.* La fase de análisis de datos en cualquier estudio ecológico, independientemente de que se utilicen datos espacialmente explícitos o no, debería comenzar con una exploración detallada de los mismos (véanse los tratados de Isaaks y Srivatsava [1989] y Gotelli y Ellison [2004] para una completa revisión sobre las distintas herramientas disponibles para el análisis exploratorio de datos). Distintos autores enfatizan la importancia de visualizar los datos antes de hacer cualquier análisis (Fortin y Dale 2005). Algo tan sencillo como una representación espacial de los valores obtenidos en campo nos va a permitir detectar posibles errores y valores anómalos en los datos y comprobar si los datos obtenidos cumplen los supuestos de la técnica analítica que queremos utilizar. Del mismo modo que un análisis de varianza requiere la normalidad de los datos y la homogeneidad de varianzas de los grupos que están siendo comparados, numerosas técnicas de análisis espacial requieren que ciertos supuestos, como los de normalidad y estacionaridad, se cumplan. En el apartado 6.2 del libro se describe con cierto detalle los pasos a seguir para explorar datos ecológicos espacialmente explícitos.

4) *Entender los fundamentos de los análisis realizados y el funcionamiento del software utilizado.* Como el lector podrá comprobar en sucesivos capítulos, los fundamentos conceptuales y matemáticos, así como la variedad y el tipo de programas disponibles para hacerlos, varían notablemente entre las distintas técnicas de análisis espacial. Pese a ello, existe una creciente disponibilidad de software bajo entorno Windows (y Macintosh), lo que, junto a los rápidos avances hechos por los fabricantes de hardware en los últimos años y las posibilidades de difusión abiertas con Internet, ha puesto al alcance de cualquiera la posibilidad de hacer análisis espaciales relativamente complejos de una forma rápida e intuitiva. Si bien estos aspectos han estimulado de forma notable el desarrollo de la ecología espacial, han traído como contrapartida el hecho de que no se requiera conocer en profundidad las técnicas utilizadas para poder aplicarlas. El saber cómo se utiliza un programa no implica que se conozcan los fundamentos de la técnica que se está utilizando y, por consiguiente, muchas veces se hacen los análisis de forma automática o rutinaria sin comprobar sus asunciones básicas, con los consiguientes errores que ello plantea. Así pues, una recomendación básica a la hora de analizar cualquier tipo de datos mediante

técnicas de análisis espacial es la de entender los fundamentos de la técnica y saber cómo los programas que estamos utilizando con nuestros datos la están ejecutando. La mayoría de nosotros somos usuarios de estas técnicas, pero necesariamente hemos de tener un conocimiento adecuado de las técnicas y algoritmos que los programas “amigables” que utilizamos manejan.

5) *Utilizar técnicas alternativas de análisis espacial cuando sea posible.* En párrafos anteriores se ha enfatizado la importancia de tener en cuenta la técnica analítica que se pretende utilizar durante las fases preliminares de cualquier estudio. No obstante, conviene tener en cuenta que no existe una única técnica de análisis espacial que permita capturar todos los atributos y matices del patrón espacial de un conjunto de datos, que numerosas técnicas son complementarias entre sí (Dale *et al.* 2002), que se puede abordar una misma cuestión desde planteamientos diferentes y que la utilización de distintas técnicas con los mismos datos puede aumentar de forma notable la cantidad de información extraída de los mismos (Goovaerts 1997, Barot *et al.* 1999, Maestre 2002, Perry y Dixon 2002, Fortin y Dale 2005). Así pues, en los casos en los que los datos puedan analizarse con técnicas alternativas se recomienda hacerlo.

Si seguimos estos pasos con éxito tendremos nuestros datos correctamente analizados, y sólo quedará ya proceder a su publicación. El tratamiento de este tema queda fuera del ámbito de esta obra, de hecho daría para escribir más de un libro como éste, por lo que no nos vamos a extender en el mismo. Únicamente nos gustaría reseñar un aspecto importante pero que no es muy cuidado a la hora de plasmar los resultados de nuestros análisis en las publicaciones: el de facilitar el entendimiento y la replicación de los mismos por parte de otros científicos. Con frecuencia nos encontramos en muchos artículos con que es imposible reproducir los análisis realizados por los autores, bien porque se omite información sobre los pasos dados, bien porque se presentan los métodos de una forma críptica o bien porque los datos no están accesibles. Estos aspectos deben evitarse a toda costa, ya que lo único que consiguen es hacer que un artículo no se entienda, minimizando así su utilización y citación por otros colegas y, en resumen, su impacto y utilidad científica. En este sentido, animamos a todos los autores a que expliquen de forma detallada los protocolos de muestreo y análisis seguidos en sus publicaciones, y a que hagan uso de las numerosas posibilidades disponibles para hacer accesibles sus datos, los códigos de programación que desarrollen y las nuevas aplicaciones de métodos existentes que lleven a ca-

bo. Conviene destacar que hay un fuerte movimiento que aboga por el acceso libre no sólo a los datos utilizados en las investigaciones ecológicas, sino también a la secuencia de análisis completa (Duke 2006, Crall *et al.* 2006, Hollister y Walter 2007), con el fin de facilitar las revisiones y meta-análisis y permitir que los análisis realizados puedan ser contrastados y verificados por otros científicos. En este sentido, la sociedad ecológica americana ha puesto en marcha un registro de datos (<http://data.esa.org/>) para que los autores de los artículos publicados en sus revistas, que incluyen las prestigiosas *Ecology*, *Ecological Monographs* y *Ecological Applications*, hagan accesibles los datos reflejados en los mismos. Asimismo, numerosas revistas permiten almacenar en sus páginas web apéndices con códigos de programación, explicaciones detalladas de los métodos e incluso los propios datos. Nosotros apoyamos este tipo de iniciativas, que sin duda redundarán en el avance de la ciencia en general y del análisis espacial en ecología en particular, y recomendamos a los lectores que hagan lo propio.

### 1.7. UNA VARIEDAD DE TÉCNICAS Y APROXIMACIONES: BREVE GUÍA PARA NO PERDERSE EN EL LABERINTO ESPACIAL

Como podrá comprobar el lector a lo largo de estas páginas, existe una gran variedad de técnicas y herramientas diseñadas para el análisis de datos ecológicos espacialmente explícitos, que permite abordar con éxito un sinnúmero de problemas ecológicos. De hecho, nuevos métodos y programas de software de análisis espacial aparecen casi mensualmente. ¿Quién no se ha sentido, cuanto menos, confundido o abrumado ante las decenas de opciones analíticas existentes en la actualidad? Esta diversidad de aproximaciones, herramientas y programas, si bien es de una utilidad innegable, constituye uno de los principales problemas a los que se enfrenta todo aquel que se aproxima por primera vez al análisis espacial de datos ecológicos. Por otra parte, la elección de una técnica de análisis espacial adecuada al problema/hipótesis que se aborda en nuestro estudio resulta fundamental para obtener conclusiones sólidas, evaluar hipótesis de una manera rigurosa y, en última instancia, poder cumplir los objetivos planteados al inicio del mismo. Con el fin de servir de guía para los que sumergen por primera vez en el mundo del análisis espacial, se presentan a continuación una serie de tablas-resumen que muestran la diversidad de técnicas analíticas disponibles, así como recomendaciones sobre cuándo y

como utilizar cada una de ellas. Esta presentación es necesariamente breve, ya que la mayoría de ellas se describen con todo detalle en los sucesivos capítulos de la obra. En cualquier caso parece necesario volver a recordar que el solapamiento (y/o la complementariedad) de estas técnicas es muy amplio y que a día de hoy estamos lejos de haber completado la exploración de todas las posibilidades de algunas de ellas.

A grandes rasgos, se pueden dividir las distintas técnicas disponibles en función del tipo de datos a analizar y de los objetivos del estudio (Tabla 1.1). Si los datos disponibles son en forma de puntos, las técnicas denominadas de análisis de densidad de puntos como, por ejemplo, las basadas en la función  $K$  de Ripley (Capítulo 3) son opción más recomendable para el análisis de datos ecológicos. Estas técnicas no sólo nos dicen si los puntos se disponen al azar, de forma regular o agregada, si no que en este último caso nos informan sobre el tamaño que tienen las manchas. Si lo que tenemos son dos tipos de puntos podemos evaluar la relación entre ellos, es decir si hay fenómenos de atracción o repulsión (así como las escalas a las que se produce esa interacción). Incluso, con la hipótesis nula adecuada, podemos evaluar cuales han podido ser las causas subyacentes a los patrones observados. En este momento disponemos también de herramientas que nos permiten evaluar de forma simultánea patrones de muchos tipos de puntos. Si, por el contrario, tenemos datos obtenidos de unidades muestrales discretas, o variables continuas medidas en puntos, resulta más complicado dar alguna recomendación sobre la técnica a utilizar. Las técnicas disponibles difieren notablemente en su complejidad matemática, los requisitos que tienen que cumplir los datos para que puedan aplicarse, la información que puede obtenerse de su aplicación y la presencia de pruebas para evaluar la significación estadística de los patrones estudiados. Así pues, la idoneidad de las distintas técnicas habrá que ponderarla siempre en función de los objetivos concretos del estudio. Si disponemos de datos en forma de conteos, y estamos interesados en estimar si el patrón espacial de la variable bajo estudio difiere significativamente de la aleatoriedad, las técnicas basadas en índices de distancia (SADIE, Capítulo 4) son muy recomendables, ya que su base matemática es sencilla y además producen numerosa información visual de gran utilidad para interpretar los resultados de los análisis. Si además queremos ver a qué escala/distancia concreta el patrón espacial difiere significativamente de la aleatoriedad, la utilización de correlogramas basados en las funciones de

Moran y Geary (Capítulo 5) es una buena opción. Si, por el contrario, queremos caracterizar atributos del patrón espacial de variables continuas (como el tamaño de las manchas y la distancia a la que dos puntos de muestreo pueden ser considerados como independientes) y realizar interpolaciones para obtener mapas de su distribución espacial en lugares no muestreados, las técnicas geoestadísticas (Capítulo 6) son altamente recomendables. Aunque se verá con más detalle en los capítulos correspondientes es importante indicar que estas técnicas no son excluyentes y que en realidad son fácilmente convertibles. En general las técnicas geoestadísticas, probablemente por cuestiones históricas, se han utilizado para construir modelos que han sido utilizados en prospección, no existiendo especial interés cuando se desarrollaron en ver si el modelo era “significativo” o no. Por el contrario en los correlogramas las preguntas básicas tenían que ver no sólo con la intensidad de la relación espacial de la variable continua si no en saber si el patrón era diferente o no del azar para alguna distancia. Si el objetivo de nuestro estudio es evaluar el patrón espacial de datos multidimensionales, como puede ser por ejemplo un listado de especies (i.e. oribátidos) en muestras de suelo para las que tenemos su ubicación espacial, las aproximaciones matriciales, como el correlograma de Mantel (Capítulo 8), son una buena opción de análisis.

## 1.8. PÁGINAS WEB DE INTERÉS

Existen multitud de páginas en internet dedicadas al análisis espacial de datos ecológicos. Así, una búsqueda en Google con los descriptores “análisis espacial” y “ecología” da un total de 348.000 páginas web, mientras que si la hacemos con sus homónimos en inglés (“spatial analysis” y “ecology”) la cifra asciende a la friolera de 1.348.000 páginas (búsquedas realizadas el 12 de abril de 2007). En sucesivos capítulos, los autores revisarán las principales páginas web centradas en la temática del mismo, así como aquellas donde puedan descargarse los programas que pueden utilizarse para realizar los distintos análisis presentados en el libro, por lo que se van a reseñar aquí algunas direcciones que albergan información genérica sobre el análisis espacial de datos ecológicos y que son de interés general.

En primer lugar nos gustaría reseñar la página web del Grupo de Ecología Espacial de la Asociación Española de Ecología Terrestre (<http://www.ecolo->



**Tabla 1.1.** Clasificación de las principales técnicas de análisis espacial disponibles en función de los objetivos del estudio. Entre paréntesis se indica el número del capítulo en el que se describe la técnica; en el caso de técnicas no abordadas en este libro se indican referencias relevantes para encontrar más información sobre las mismas. Adaptada de Fortin *et al.* (2002) y Fortin & Dale (2005)

Objetivo	Tipo de datos	
	Unidades muestrales discretas	Puntos
Exploración (caracterización de la estructura espacial)	"Join-count" (5), Moran's $I$ (5), Geary's $c$ (5)	Métodos basados en la distancia al vecino más cercano (3)
	Local Moran's $I$ , Getis's $G$ , Ord's $O$ (5)	$K$ de Ripley y métodos relacionados (3)
	Semivariogramas (6)	"Join-count" (5)
	Correlograma de Mantel¶ (8)	Método de Dixon (Dixon 2002)
	Test de Mantel (8)	Método del círculo circunscrito (Dale y Powell 1991)
	Test de Mantel parcial (8)	Dimensión fractal (Alados <i>et al.</i> 2003)
	SADIE (4)	
	Métodos basados en la varianza de cuadrados¶ (Dale 1999)	
	Método del círculo circunscrito¶ (Dale y Powell 1991)	
	Dimensión fractal¶ (Alados <i>et al.</i> 2003)	
	"Lacunarity"¶ (Plotnick <i>et al.</i> 1996)	
	Análisis espectral¶ (Renshaw 1997)	
	Métodos basados en ondas¶ ("Wavelets"; Dale y Mah 1998)	
Inferencia estadística (evaluación de hipótesis, estimación de parámetros)	Test de Mantel (8)	
	Test de Mantel parcial (8)	
	Semivariogramas (6)	
	SADIE (4)	
	Modelos ocultos de Harkov (7)	
	Modelos autoregresivos (Haining 1990)	
	Alineado condicional (Cressie 1993)	

Tabla 1.1. Continuación

Objetivo	Tipo de datos	
	Puntos	Unidades muestrales discretas
<b>Interpolación</b> (obtención de mapas)	"Krigging" (6) Análisis de tendencias superficiales (Legendre y Legendre 1998)	Polígonos de Voronoi (Fortin y Dale 2005)
<b>Segmentación</b>	Métodos para la detección de fronteras y ecotonos (9)	

¶Requiere una malla de muestreo regular

*giaespacial.es*). En ella puede encontrarse numerosa información sobre las actividades de este grupo, enlaces relacionados con la ecología espacial y publicaciones de sus miembros. En Internet pueden encontrarse distintos manuales y compendios de clases que merece la pena visitar. Uno de ellos es el compendio sobre ecología de poblaciones cuantitativa desarrollado por Alexei Sharov, de la Universidad Tecnológica de Virginia (<http://www.ento.vt.edu/~sharov/PopEcol/popecol.html>). Si bien está escrito en inglés, presenta una introducción bastante didáctica al análisis espacial de organismos y poblaciones. Otra página de interés general que merece la pena visitar es <http://www.ai-geostats.org/>. Si bien está centrada en técnicas geoestadísticas, contiene información sobre cuestiones generales relativas al análisis de datos espacialmente explícito, así como numerosa documentación, enlaces a páginas web y programas, etc. Por último, conviene destacar la página personal de Pierre Legendre (<http://www.bio.umontreal.ca/legendre/indexEnglish.html>), profesor de la Universidad de Montreal y uno de los principales expertos en el análisis estadístico de datos ecológicos, que es muy recomendable por la gran cantidad de información que contiene (programas, artículos, bases de datos, etc.).

## AGRADECIMIENTOS

FTM agradece el apoyo de un contrato Ramón y Cajal del Ministerio de Educación y Ciencia y de un Early Career Project Grant (ECPG 231/

607), financiado por la British Ecological Society. Agradecemos la financiación de los proyectos CEFEMED (URJC-RNT-063-2), financiado por la Dirección General de Universidades e Investigación de la Consejería de Educación de la Comunidad de Madrid (DGUICAM) y la Universidad Rey Juan Carlos, REMEDINAL (S-0505/AMB/0335), financiado por la DGUICAM, INTERCAMBIO (BIOCON 06\_039), financiado por la Fundación BBVA y EXTREM (CGL2006-09431) financiado por el MEC.

