

*Introducción a los Sistemas de
Información Geográfica*

Introducción a los Sistemas de Información Geográfica
Copyright ©2016 Víctor Olaya

Versión revisada el 8 de julio de 2016

Se concede permiso para copiar, distribuir o modificar esta obra bajo los términos expresados en la licencia Creative Common Atribución, la cual puede encontrarse en [www.creativecommons.org](http://creativecommons.org). La licencia se aplica a todo el texto, así como las imágenes creadas por el propio autor, que serán aquellas para las que no se especifique de modo explícito una distinta procedencia. Este libro puede descargarse y consultarse de forma libre en varios formatos, incluyendo formatos editables, en la dirección Web <http://victorolaya.com>.

Los nombre de productos o corporaciones que aparecen en el texto pueden constituir marcas registradas y se emplean sin otro afán que el meramente identificativo. Asimismo, la inclusión o no de uno de tales productos no expresa recomendación alguna por parte del autor.

VÍCTOR OLAYA

Introducción a los Sistemas de Información Geográfica

PRÓLOGO

Hace ahora más de cinco años que se publicó la primera versión de *Sistemas de Información Geográfica*, un libro libre sobre fundamentos de SIG en español, y apenas unos meses desde que apareció la segunda. El libro ha tenido una acogida excelente, y mi intención es seguir manteniéndolo actualizado en la medida que sea posible, reflejando los avances que, a buen seguro, van a producirse en el campo de los SIG.

Existe, no obstante, un obstáculo importante para que el libro alcance a todos los públicos: su tamaño. Por su completitud, y por la complejidad propia de la disciplina, el libro es un volumen de más de 800 páginas cargadas de detalle. La segunda versión se presenta en un único tomo, frente a los dos en que consistía la primera, pero aún así sigue quedando como una obra de consulta demasiado extensa para leerse de principio a fin. Para el lector que comienza a introducirse en el ámbito de los SIG y no busca especializarse, resulta un volumen intimidante y es, no hay duda, difícil de abordar.

Este libro intenta ser una alternativa a la obra completa, de tal forma que resulte más accesible para quienes desean tener una perspectiva global de la disciplina de los SIG, sin entrar en detalles demasiado específicos. Es, básicamente, una versión resumida de aquél, pensada con la idea de usarse no como libro de consulta, sino como libro de lec-

tura. Además de ser más breve, se presenta en un formato más adecuado para esta clase de propósito, con algunas modificaciones en su enfoque y con menos contenido gráfico.

He respetado en líneas generales la estructura de los capítulos, de modo que es fácil para el lector que desee profundizar en uno de ellos encontrar este en el libro completo. Desde ese punto de vista, puede entenderse este libro como una especie de «índice» de su hermano mayor, un índice, no obstante, prolífico y con suficiente información como ofrecer al lector una visión detallada del mundo de los SIG.

Este es también, por supuesto, un libro libre, que espero que progrese de una forma dinámica gracias a la contribución de sus lectores. Si encuentras cualquier error o quieres colaborar en mejorar estas páginas, no dudes en escribirme a volayaf@gmail.com.

INTRODUCCIÓN A LOS SISTEMAS DE INFORMACIÓN GEOGRÁFICA

¿QUÉ ES UN SIG?

La mayor parte de la información que manejamos en cualquier tipo de disciplina está georreferenciada. Es decir, se trata de información a la cual puede asignarse una posición geográfica, y es por tanto información que viene acompañada de otra información adicional relativa a su localización.

Un **Sistema de Información Geográfica** (SIG) es una herramienta para trabajar con información georreferenciada. En particular, un SIG es un sistema que permite la realización de las siguientes operaciones:

- **Lectura, edición, almacenamiento** y, en términos generales, **gestión** de datos espaciales.
- **Análisis** de dichos datos. Esto puede incluir desde consultas sencillas a la elaboración de complejos modelos, y puede llevarse a cabo tanto sobre la **componente espacial** de los datos (la localización de cada valor o elemento) como sobre la **componente temática** (el valor o el elemento en sí).
- Generación de **documentos** tales como mapas, informes, gráficos, etc.

Un SIG representa un paso más allá de los mapas clásicos. Mientras que un mapa es una representación de un conjunto de datos espaciales, y aunque esta representación resulta de enorme importancia, en el entorno de un SIG no es sino un

elemento más de un conjunto de ellos. El SIG no incluye solo los datos y la representación, sino también las operaciones que pueden hacerse sobre el mapa, que no son ajenas a este sino partes igualmente de todo ese sistema.

Un SIG es una herramienta versátil y de amplio alcance, y hoy día la gran mayoría de disciplinas se benefician del uso de SIG de uno u otro modo. Una de las principales razones por las que esto sucede es el **carácter integrador** de los SIG. Los siguientes son algunos de los contextos principales en los que un SIG ejerce tal función integradora.

- **SIG como integrador de información.** Un nexo común entre la mayoría de disciplinas es el hecho de que sus objetos de estudio están asociados a una localización en el espacio. Esto va a permitir combinarlas y obtener resultados a partir de un análisis común. El SIG es, en este contexto, el marco necesario en el que incorporar esa información georreferenciada y trabajar con ella.
- **SIG como integrador de tecnologías.** Una gran parte de las tecnologías que han surgido en los últimos años (y seguramente de las que surjan en los próximos) se centran en el aprovechamiento de la información espacial, y están conectadas en mayor o menor medida a un SIG para ampliar su alcance y sus capacidades. Por su posición central en el conjunto de todas las tecnologías, los SIG cumplen además un papel de unión entre ellas, conectándolas y permitiendo una relación fluida alrededor de las funcionalidades del propio SIG.
- **SIG como integrador de personas.** Las funciones básicas que un SIG ha de cumplir cubren en realidad un rango amplio de trabajo, y engloban las necesidades de usuarios que con anterioridad no tenían entre sí un marco de trabajo común tan definido. Esto tiene como consecuencia que existe una mejor coordinación entre ellos, pues es la propia herramienta quien establece las características de las relaciones existentes, y estas no dependen ya únicamente del propio ámbito de aplicación.
- **SIG como integrador de teorías y fundamentos.** En un principio, podemos entender un SIG como la unión

de dos ciencias: la geografía y la informática. Sin embargo, un análisis más detallado nos revela que un SIG incorpora elementos de muchas ciencias distintas, como pueden ser las disciplinas relacionadas con la tecnología y el manejo de información (informática, diseño de bases de datos, tratamiento digital de imágenes), las dedicadas al estudio de la Tierra desde un punto de vista físico (geología, la geología, la oceanografía, la ecología), las dedicadas al estudio de la Tierra desde un punto de vista social y humano (antropología, geografía, sociología), las dedicadas al estudio del entendimiento humano (ciencias del conocimiento, psicología) o las disciplinas que tradicionalmente han realizado una integración de conocimientos de otros ámbitos distintos, entre las que cabe destacar la geografía. El término **geomática**, formado a partir de los vocablos *geografía* e *informática*, se emplea con frecuencia para hacer mención a todo ese grupo de ciencias relacionadas con los SIG.

Con todo lo anterior, se tiene que SIG es un sistema que integra tecnología informática, personas e información geográfica, y cuya principal función es capturar, analizar, almacenar, editar y representar datos georreferenciados.

Desde otro punto de vista, un SIG puede considerarse compuesto de cinco bloques fundamentales.

- **Datos.** Los datos son necesarios para hacer que el resto de componentes de un SIG cobre sentido y puedan ejercer su papel en el sistema. La información geográfica, la verdadera razón de ser los SIG, reside en los datos, y es por ello que el conocimiento exhaustivo de los datos, su naturaleza, su procedencia, su calidad, así como su gestión y almacenamiento, resulta obligado para una buena comprensión los propios SIG.
- **Análisis.** El análisis es una las funcionalidades básicas de los SIG, y una de las razones fundamentales que llevaron al desarrollo de estos. Un ordenador es una herramienta con enorme capacidad de cálculo, y esta puede aplicarse a los datos espaciales para obtener resultados de muy diversa índole.

En mayor o menor medida, un SIG siempre incorpora una serie de formulaciones que permiten la obtención de resultados y el análisis de los datos espaciales. Las ventajas de la incorporación de todos estos procesos en una única herramienta, el SIG, van desde la **automatización de tareas** a la aparición de nuevos procesos que producen resultados que no podrían ser obtenidos de otro modo.

- **Visualización.** Cualquier tipo de información puede ser representada de forma gráfica, lo cual facilita la interpretación de dicha información o parte de esta. En el caso particular de la información geográfica, la visualización no solo es una forma más de trabajar con esa información, sino que resulta la forma principal, por ser aquella a la que estamos más acostumbrados gracias al uso de mapas. Al contrario que un mapa, que de por sí es de naturaleza gráfica, en un SIG trabajamos con datos de tipo puramente numérico. Para poder presentar una utilidad similar a la de un mapa, un SIG debe incluir capacidades que generen representaciones visuales a partir de esos datos numéricos.

La visualización de la información geográfica se rige por los mismos conceptos y principios que se emplean para la confección de cartografía impresa, y estos deben ser conocidos por el usuario de SIG, ya que una de las tareas de este es el diseño cartográfico y las preparación de los elementos de visualización para poder realizar su trabajo sobre las representaciones creadas.

- **Tecnología.** Se incluyen en este elemento tanto el *hardware* sobre el que se ejecutan las aplicaciones SIG, como dichas aplicaciones, es decir el *software* SIG. Además de la propia plataforma, el *hardware* incluye una serie de periféricos habituales en el trabajo con SIG, como son los periféricos para entrada de datos geográficos y creación de cartografía.
- **Factor organizativo.** Engloba los elementos relativos a la coordinación entre personas, datos y tecnología, o la comunicación entre ellos, entre otros aspectos. Con la creciente complejidad de los SIG, la gestión de las relaciones entre sus elementos es cada vez más importante.

A lo largo de este libro, detallaremos cada uno de estos bloques en los capítulos correspondientes.

HISTORIA DE LOS SIG

El desarrollo sufrido por los SIG desde sus orígenes hasta nuestros días es enorme. La popularización de las tecnologías y los esfuerzos de desarrollo llevados a cabo por un amplio abanico de ciencias beneficiarias de los SIG, todos han contribuido a redefinir la disciplina e incorporar elementos impensables entonces.

Podemos situar el origen de los SIG al inicio de la década de los **60**, como resultado de unos factores que convergen para dar lugar al desarrollo de estos. Estos factores son principalmente dos: la **necesidad creciente de información geográfica** y de una gestión y uso óptimo esta, y la **aparición de los primeros computadores**.

Las bases para la futura aparición de los SIG las encontramos algunos años antes de esa década de los 60, con el desarrollo de nuevos enfoques en cartografía, tales como la **geografía cuantitativa**, que parecen predecir las necesidades futuras que un manejo informatizado de esta traerá.

La primera experiencia relevante en la que se combinan geografía e informática la encontramos en 1959, cuando Waldo Tobler define los principios de un sistema denominado MIMO (map in-map out) con la finalidad de aplicar los ordenadores al campo de la cartografía. En él, establece los principios básicos para la creación de datos geográficos, su

codificación, análisis y representación dentro de un sistema informatizado.

El primer Sistema de Información Geográfica formalmente desarrollado aparece en Canadá. Este sistema, denominado CGIS (Canadian Geographical Information Systems), fue desarrollado a principios de los 60 por Roger Tomlinson, popularmente conocido como el «padre del SIG».

A mediados de los 60, las aplicaciones SYMAP y GRID sientan respectivamente las bases de los dos principales enfoques a la hora de manejar información geográfica: el enfoque **raster** y el enfoque **vectorial**. Ambas alternativas serán explicadas con detalle más adelante en este libro. Los conceptos básicos para el análisis dentro del SIG ráster los establece poco después Dana Tomlin, desarrollando lo que se conoce como **álgebra de mapas**.

Durante los años sesenta, se produce un gran desarrollo de los SIG a partir de esos elementos iniciales, y el SIG comienza a incorporarse a la comunidad cartográfica, dejando de ser una herramienta experimental.

La evolución de los SIG desde entonces recorre sucesivas etapas, avanzando muy rápidamente ante la influencia de numerosos factores externos. Esta evolución tiene lugar en el propio SIG como disciplina, en las tecnologías que lo sustentan, en los datos, así como en las técnicas y formulaciones.

LA EVOLUCIÓN DE LOS SIG COMO DISCIPLINA

Los SIG eran en origen una mera combinación de elementos de cartografía cuantitativa, enlazados con los sistemas informáticos de la época. Se trataba de un territorio propio de cartógrafos y geógrafos que intentaban adaptar sus conocimientos y necesidades a las tecnologías que por aquel entonces comenzaban a surgir. No obstante, desde aquellos orígenes los cambios han sido muy grandes, y se han incorporado al ámbito de los SIG un gran número de otras disciplinas cuya aportación e influencia puede ser equivalente o incluso superior a la de la cartografía o la geografía.

Coincidiendo con la etapa inicial del desarrollo de los SIG, empieza a aparecer una preocupación por el entorno

que tiene consecuencias muy favorables para el desarrollo de todas las ciencias relacionadas, la gran mayoría de las cuales son o serán usuarias directas de SIG. El SIG comienza a integrarse paulatinamente en las tareas de **gestión del medio**, como un apoyo imprescindible a la hora de analizar este.

Al principio de la década de los 70, siendo ya claro que los SIG son herramientas con gran futuro, aparecen no solo los esfuerzos de desarrollo y estabilización de la disciplina, sino todos los restantes que dan entidad propia a la prometedora ciencia de la información geográfica con base informática. Comienzan a celebrarse conferencias y simposios sobre SIG, y estos pasan a formar parte de los *curricula* universitarios. En los años 80, se consolidan las revistas y foros especializados, que habrán de llevar la disciplina a un público más amplio.

En el aspecto comercial, la industria del SIG se consolida también en los años 70. **ESRI** (Environmental Systems Research Institute), empresa pionera y líder del sector hasta el día de hoy, se funda en 1969, y sus productos tienen gran importancia a la hora de convertir los SIG en un elemento de consumo. El primer SIG de código abierto, **GRASS** (Geographic Resources Analysis Support System), aparece en 1985.

El mayor avance en la incorporación de los SIG a entornos no profesionales tiene lugar en la primera década del siglo XXI, con la aparición de servicios de cartografía tales como *Google Maps*. La popularización de los **navegadores GPS**, que incorporan tanto elementos de representación como de análisis propios de los SIG, son otro buen ejemplo de la proliferación del SIG entre un público no especializado.

LA EVOLUCIÓN DE LA TECNOLOGÍA

Tres son los bloques principales del desarrollo informático con una influencia más marcada en el campo de los Sistemas de Información Geográfica:

- **Salidas gráficas.** La evolución de las capacidades gráficas, intensa desde sus inicios hasta nuestros días y aún

muy activa, ha sido seguida de cerca por los SIG, que progresivamente van incorporando mejoras tanto en la representación en pantalla como en la generación de mapas impresos.

- **Almacenamiento y acceso de datos.** El aumento en el tamaño de los datos manejados en el SIG ha debido acompañarse de mejoras en la capacidad de almacenamiento, así como en la de lectura, para poder garantizar un uso fluido.
- **Entrada de datos.** Los datos geográficos utilizados en los primeros años de los SIG eran datos en papel que se digitalizaban y almacenaban mecánicamente en tarjetas perforadas en un único proceso mecánico. Desde esos sistemas mecánicos de tarjetas hasta los modernos equipos, la aparición de *scanners* de gran precisión y técnicas de digitalización automáticas, entre otros, ha cambiado completamente el ámbito de la entrada de datos para su uso en un SIG.

Además del avance de estos factores, la evolución general de los ordenadores afecta a todos los elementos de *software* que se ejecutan sobre ellos. De las grandes computadoras se pasa a los ordenadores personales, y los programas tales como los SIG realizan también esa transición de una a otra plataforma.

La elaboración y análisis de cartografía se convierte a finales de los años 80 en una tarea que puede ya llevarse a cabo en equipos personales (PC) de bajo coste, lejos de las grandes máquinas y equipos dedicados de alto coste.

La evolución de las plataformas no se detiene ahí. Las tendencias actuales apuntan a llevar los SIG de forma genérica a plataformas móviles tales como teléfonos o tabletas, especialmente indicadas para la toma de datos en campo. La combinación de estos últimos con las tecnologías de posicionamiento global como el GPS se demuestra altamente práctica en este aspecto.

La aparición de Internet es un hecho que ha modificado todos los aspectos de la sociedad actual, estén relacionados o no con ámbito científico. El primer uso relacionado con el

SIG o la distribución de cartografía lo encontramos en 1993, con la aparición de *Xerox PARC*, el primer servidor de mapas. El primer atlas digital en linea, el Atlas Nacional de Canadá, se encuentra disponible desde 1994. Mas recientemente, ya en este siglo, los concepto de la Web 2.0 se adaptan al ámbito de los SIG y facilitan la aparición de lo que se conoce como **Web Mapping**.

LA EVOLUCIÓN DE LOS DATOS

Las primeras bases de datos geográficas contenían mapas escaneados y elementos digitalizados en base a estos. A partir de este punto, van apareciendo nuevas fuentes de datos cuya estructura es más adecuada para su tratamiento informatizado, y al tiempo que los SIG se adaptan a estas, surge una relación bidireccional que resulta beneficiosa para ambos.

Un avance primordial en este sentido lo constituye el lanzamiento de los primeros **satélites de observación terrestre**. Las técnicas existentes para la toma de fotografías aéreas, desarrolladas principalmente con fines militares durante la Primera Guerra Mundial (aunque iniciadas a mitad del siglo XIX con la toma de fotografías desde globos aerostáticos), pasan a aplicarse a escala global con la aparición de satélites destinados a estos efectos. En 1980 se funda SPOT, la primera compañía mundial en ofrecer con carácter comercial imágenes procedentes de satélite para toda la superficie terrestre.

Las tecnologías de posicionamiento y localización son otra fuente de datos de primer orden. En 1981, el sistema GPS pasa a ser plenamente operativo, y en 2000 se amplía la precisión de este para uso civil.

Al igual que las aplicaciones, los distintos tipos de datos geográficos digitales se van asentando y popularizando, recibiendo progresivamente más atención y medios. El Servicio Geográfico Estadounidense (USGS) publica en 1976 los primeros **Modelos Digitales de Elevaciones** (MDE), en respuesta a la gran importancia que este tipo de dato tiene dentro del nuevo contexto del análisis geográfico. En el año

2000 se publican los datos de la *Shuttle Radar Topographic Mission* (SRTM), con información altitudinal de un 80 % de la superficie terrestre a una resolución de un segundo de arco (aproximadamente, 30 metros).

La aparición de nuevas técnicas tales como el LiDAR (ver ??) abre nuevos caminos en cuanto a la precisión que puede obtenerse en la caracterización del terreno, posibilitando nuevos usos y análisis antes no planteados.

La evolución de los datos no es solo una evolución técnica, sino también de carácter social y organizativo. Se empieza a entender que resulta necesario formular estrategias adecuadas para la gestión de los datos espaciales, y se desarrollan las denominadas **Infraestructuras de Datos Espaciales** (IDE). El ejemplo más destacado de estas es la IDE Nacional de los Estados Unidos (NSDI), de 1994. En Europa, la directiva INSPIRE, de 2007, pretende la creación de una infraestructura similar.

Muchos de estos desarrollos y actividades se adhieren a las especificaciones establecidas por el *Open GIS Consortium* (OGC), un consorcio internacional fundado en 1994 para **homogeneizar** el empleo y difusión de los datos geográficos.

LA EVOLUCIÓN DE LAS TÉCNICAS Y FORMULACIONES

Una vez que se implementan los primeros SIG y se suplen las necesidades de análisis y gestión de datos espaciales que motivaron su aparición, comienza el proceso de desarrollar nuevas técnicas y planteamientos que permiten ir más allá en dicho análisis.

Los antecedentes del análisis espacial son relativamente recientes. En 1854 John Snow realizó la que puede considerarse como una de las primeras experiencias cartográficas analíticas, al utilizar mapas de puntos para efectuar sus deducciones y localizar en Inglaterra la fuente de un brote de cólera.

En su libro *Design with Nature* (1969), Ian McHarg define los elementos básicos de la **superposición y combinación de mapas**, que, como veremos más adelante, son los que

se aplican tanto en el análisis como en la visualización de las distintas *capas* de datos geográficos en un SIG.

Un desarrollo especialmente relevante es el experimentado por el **análisis del relieve**, una disciplina que con la aparición de los SIG sufre un salto cualitativo muy importante. La orografía clásica, con un enfoque tradicionalmente sustentado en la geología y el análisis geomorfológico, va dando lugar a una ciencia cada vez más cuantitativa centrada en el análisis morfométrico del relieve.

Junto con la componente analítica, otros elementos de la práctica cartográfica evolucionan de forma similar. En 1819, Pierre Charles Dupin crea el primer **mapa de coropletas**. Con la llegada de los SIG, este tipo de mapas se convertirán en una forma de representación muy popular.

El avance en el desarrollo de las aplicaciones de diseño asistido por ordenador (CAD), y en general de las representaciones gráficas por ordenador, impulsó igualmente la aparición y evolución posterior de una nueva disciplina: la geometría computacional. Sobre esta se fundamenta el análisis vectorial dentro de un SIG, y también parte de los mecanismos que este usa para la representación gráfica de elementos.

FUNDAMENTOS CARTOGRÁFICOS Y GEODÉSICOS

CONCEPTOS GEODÉSICOS BÁSICOS

La característica principal de la información georreferenciada es que tiene una **localización en el espacio**, particularmente en el espacio terrestre. Esta localización se ha de dar por medio de unas **coordenadas** que la definan de forma adecuada, lo cual implica la necesidad de establecer un sistema en base al cual expresar dichas coordenadas.

La **geodesia** es la ciencia encargada de proveer el marco teórico en el que fundamentar lo anterior, y su objeto de estudio es **la forma de la Tierra**. La geodesia, en sus diversas ramas, proporciona métodos y conceptos que permiten la utilización rigurosa de coordenadas.

La necesidad del estudio geodésico surge por el hecho de que la Tierra no es plana, y cuando el territorio que pretendemos estudiar es lo suficientemente extenso, la curvatura de la Tierra no puede ser ignorada. Este es el caso que vamos a encontrar cuando trabajemos con un SIG, y es por ello que los SIG implementan los elementos necesarios para poder efectuar un manejo de la información geográfica riguroso y acorde con los conceptos de la geodesia.

Uno de los objetivos principales de la geodesia es establecer un sistema de referencia y definir un conjunto de puntos (conocidos como **vértices geodésicos**) cuyas coordenadas

en dicho sistema sean conocidas con una precisión elevada. Posteriormente, y en base a esos puntos, los cuales forman una **red geodésica**, se pueden calcular las coordenadas de cualquier punto en el sistema de referencia definido.

Superficies de referencia

Dos conceptos básicos para esta tarea: el **elipsoide de referencia** y el **geoide**.

La Tierra tiene forma esférica, aunque no es una esfera perfecta, sino que está achatada, constituyendo lo que se conoce como **elipsoide**. Sobre un elipsoide, el radio de la Tierra ya no es constante, y depende del emplazamiento. Asimilar la Tierra a un elipsoide es más preciso que suponer la Tierra con una forma perfectamente esférica, y es necesario a la hora de elaborar cartografía de zonas no muy extensas.

Una vez que se dispone de una expresión teórica para la forma de la Tierra, el siguiente paso es la determinación de los parámetros que definen esta. En el caso de utilizar la esfera, hay que calcular su radio. En el caso de asumir el elipsoide como forma de referencia, deben determinarse las medidas de los semiejes menor y mayor.

Por razones históricas, existen numerosos elipsoides, derivados del trabajo de los geodestas en diferentes épocas y lugares. Los primeros elipsoides generales, que permiten ser usados en toda la superficie terrestre, aparecen hace aproximadamente un siglo, con objeto de disponer de una referencia internacional que facilite el uso de cartografía en las distintas zonas del planeta. El **elipsoide WGS-84** es uno de los más empleados en la actualidad, ya que es el utilizado por el sistema GPS (apartado 5).

El geoide es la otra superficie de referencia, definida como la superficie tridimensional en cuyos puntos la atracción gravitatoria es constante. Se trata de una superficie equipotencial que resulta de suponer los océanos en reposo y a un nivel medio y prolongar estos por debajo de la superficie terrestre.

Al igual que en el caso de los elipsoides, existen diversos geoides de referencia, y estos no son constantes en el tiempo

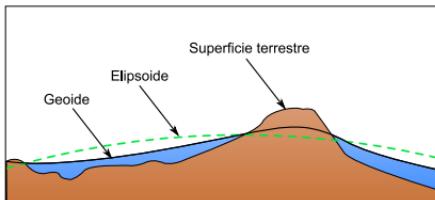


Figura 3.1: Tres superficies fundamentales: superficie real de la Tierra, geoide y elíptico (Adaptado de Wikipedia).

sino que evolucionan para adaptarse a las modificaciones que tienen lugar sobre la superficie terrestre.

La figura 3.1 muestra una comparación esquemática entre las tres superficies: superficie real de la Tierra, geoide y elíptico.

En un elíptico **general**, tanto la posición de su centro de gravedad como de su plano ecuatorial coinciden con los terrestres. Por el contrario, cuando el elíptico es **local**, estas propiedades no han de cumplirse necesariamente, y el elíptico a solas resulta insuficiente, ya que carecemos de información sobre su posicionamiento con respecto a la superficie terrestre.

Surge así el concepto de **datum**, que es el conjunto formado por una superficie de referencia (el elíptico) y un punto en el que «enlazar» este al geoide. Este punto se denomina **punto fundamental**, y en él el elíptico es tangente al geoide. La vertical al geoide y al elíptico son idénticas en el punto fundamental.

Para un mismo elíptico pueden utilizarse distintos puntos fundamentales, que darán lugar a distintos datum y a distintas coordenadas para un mismo punto.

Sistemas de coordenadas

Una vez hemos definido un modelo para definir la forma de la Tierra, podemos establecer un sistema de codificar cada una de las posiciones sobre su superficie y asignar a estas las correspondientes coordenadas. Para ello, encontramos

dos opciones: utilizar los elementos de la **geometría esférica** y con estos definir el sistema de referencia, o utilizar la *geometría plana*, para lo cual será necesario un mecanismo de **proyección** de coordenadas que permita situar los elementos de la superficie del elipsoide sobre una superficie plana.

El sistema de *coordenadas geográficas* es un sistema de coordenadas esféricas mediante el cual un punto se localiza con dos valores angulares: **latitud** y **longitud**. Las líneas de igual latitud o longitud se denominan **paralelos** y **meridianos** respectivamente.

Las coordenadas geográficas resultan de gran utilidad, especialmente cuando se trabaja con grandes regiones. No obstante, no se trata de un sistema cartesiano, y tareas como la medición de áreas o distancias es mucho más complicada. Para poder crear cartografía y simplificar gran número de operaciones posteriores, necesitamos coordenadas cartesianas. El proceso de asignar una coordenada plana a cada punto de la superficie de la Tierra (que no es plana) se conoce como **proyección cartográfica**.

La superficie de la esfera no es **desarrollable**, es decir, no puede convertirse en un plano. Por ello, es necesario disponer de una metodología para pasar puntos desde la superficie curva al plano, tal y como el que se muestra en la figura 3.2.

En el caso de la figura, los puntos se proyectan directamente sobre un plano. Otra opción es proyectarlos sobre una superficie tridimensional que, al contrario que la esfera, sea desarrollable. Las más habituales son el cilindro y el cono, que dan lugar a las **proyecciones cónicas** y **cilíndricas**.

Puede apreciarse en la figura que se producen distorsiones al realizar la proyección. Por ejemplo, la distancia entre los puntos *A* y *B* no es igual a la existente entre los puntos *a* y *b*. Con independencia de las características propias de la proyección, siempre existen distorsiones, por ser la de la esfera una superficie no desarrollable. Estas distorsiones se conocen como **anamorfosis** .

Según las propiedades métricas que se conserven, las proyecciones pueden ser **equiárea** (mantienen una escala

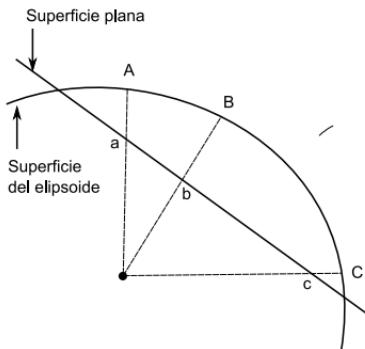


Figura 3.2: Esquema del concepto de proyección. A los puntos *A*, *B* y *C* sobre la superficie del elipsoide se les asocian equivalentes *a*, *b* y *c* sobre un plano.

constante), **conformes** (mantienen los ángulos y la forma de los objetos) o **equidistantes** (mantienen las distancias).

La elección de una u otra proyección es función de las necesidades concretas de cada caso de uso.

En la actualidad, una de las proyecciones más extendidas en todos los ámbitos es la **proyección universal transversa de Mercator**, la cual da lugar al **sistema de coordenadas UTM**. Este sistema no es simplemente una proyección, sino un sistema completo para cartografiar la práctica totalidad de la Tierra. Para ello, esta se divide en una serie de zonas rectangulares mediante una cuadricula y se aplica una proyección y unos parámetros geodésicos concretos a cada una de dichas zonas. En su forma actual, emplea un único elipsoide (WGS-84).

Con el sistema UTM, las coordenadas de un punto no se expresan como coordenadas terrestres absolutas, sino mediante la zona correspondiente y las coordenadas relativas a la zona UTM en la que nos encontramos.

La cuadricula UTM tiene un total de 60 husos numerados entre 1 y 60, cada uno de los cuales abarca una amplitud de 6° de longitud. El huso 1 se sitúa entre los 180° y 174° O, y la numeración avanza hacia el Este.

En latitud, cada huso se divide en 20 zonas, que van

desde los 80°S hasta los 84°N. Estas se codifican con letras desde la C a la X, no utilizándose las letras I y O por su similitud con los dígitos 1 y 0. Cada zona abarca 8 grados de longitud, excepto la X que se prolonga unos 4 grados adicionales.

Una zona UTM se localiza, por tanto, **con un número y una letra**, y es en función de la zona como posteriormente se dan las coordenadas que localizan un punto. Estas coordenadas se expresan en metros y expresan la distancia entre el punto y el origen de la zona UTM en concreto. El origen de la zona se sitúa en el punto de corte entre el meridiano central de la zona y el ecuador.

Para evitar la aparición de números negativos, se considera que el origen no tiene una coordenada X de 0 metros, sino de 500000, y una coordenada Y de 10000000 metros, lo cual hace que todas las coordenadas referidas a él sean positivas.

Transformación y conversión de coordenadas

Una situación muy habitual en el trabajo con un SIG es disponer de cartografía en **varios sistemas de coordenadas**, o bien en un mismo sistema pero con parámetros diferentes (por ejemplo, diferente datum). Para poder emplear toda esa cartografía de forma conjunta, resulta necesario trabajar en un sistema único y bien definido, lo cual hace necesario convertir al menos una parte de ella. Cuando el datum es distinto en los sistemas de origen y destino, la **conversión de coordenadas** se conoce como **transformación de coordenadas**.

Las operaciones de transformación y conversión aparecen en los SIG como funcionalidades que permiten modificar los datos geográficos, reemplazando sus coordenadas por coordenadas en otro sistema de coordenadas. Igualmente, aparecen como funcionalidades de representación, permitiendo la conversión «**al vuelo**», es decir, en tiempo real. En este caso, un dato en un sistema de coordenadas se puede representar en cualquier otro sin necesidad de una conversión

previa, con lo que puede usarse conjuntamente con datos en un sistema de coordenadas distinto.

Para facilitar el uso de sistemas de referencia, existen proyectos de codificación de estos, de forma que cada sistema existente puede identificarse de forma sencilla mediante un código. El más extendido de estos es el sistema de codificación **EPSG**.

CONCEPTOS CARTOGRÁFICOS BÁSICOS

De entre los conceptos fundamentales de la cartografía que todo usuario de SIG ha de conocer, destaca el de **escala**. La escala es la **relación de tamaño** existente entre el mapa que se obtiene al desarrollar nuestra superficie de proyección (de tamaño acorde con el objeto proyectado, esto es la Tierra) y el que finalmente manejamos, de tamaño más reducido. Conociendo esta relación podemos conocer las verdaderas magnitudes de los elementos que vemos en el mapa, ya que podemos convertir las medidas hechas sobre el mapa en medidas reales. Es importante recordar que esas medidas no son tan «reales», puesto que la propia proyección las ha distorsionado —lo cual no debe olvidarse—, pero sí que son medidas en la escala original del objeto cartografiado.

La escala se expresa habitualmente como un denominador que relaciona una distancia medida en un mapa y la distancia que esta medida representa en la realidad. Por ejemplo, una escala 1:50000 quiere decir que 1 centímetro en un mapa equivale a 50000 centímetros en la realidad, es decir a 500 metros. Este valor se conoce como **escala numérica**.

Independientemente del tipo de proyección, la escala es completamente cierta únicamente en determinadas partes del mapa. En otros puntos de este, la escala varía. La relación entre la escala en esos puntos y la escala numérica se conoce como **factor de escala**.

Aunque tradicionalmente se entiende la escala como un concepto asociado a la representación, los datos geográficos tienen una escala inherente que no es función de dicha representación, sino del detalle con que han sido tomados. En

este sentido es más conveniente entender la escala como un elemento relacionado con la **resolución** de los datos, es decir, con el **tamaño mínimo cartografiado**. Esta concepción no es en absoluto propia de los SIG, ya que deriva de las representaciones clásicas y los mapas impresos. Se sabe que el tamaño mínimo que el ojo humano es capaz de diferenciar es del orden de 0,2 mm. Aplicando a este valor la escala a la que queremos crear un mapa, tendremos la mínima distancia sobre el terreno que debe medirse.

Es importante ser consciente de la limitación que la escala considerada a la hora de la toma de datos (conocida como **escala operacional**) impone, especialmente en el contexto de un SIG. En un SIG, podemos aumentar el tamaño en pantalla de una cierta información geográfica, variando la escala de representación (también conocida como **escala cartográfica**), pero ello no modifica la escala operacional. Por mucho que ampliemos no vamos a ver más detalles, ya que para ello sería necesario tomar más datos.

Un tipo de datos particulares con los que se trabaja en un SIG, los datos *raster*, tienen a su vez un parámetro de resolución (el **tamaño de celda**) ligado a la escala. Veremos más al respecto en el capítulo 4.

Relacionado con el concepto de escala encontramos la denominada **generalización cartográfica**. Generalizar implicar expresar alguna idea o información de forma más resumida, de tal modo que esta sea comprensible y pueda aprovecharse de la mejor manera posible. La generalización es necesaria en un SIG para representar datos a una escala menor que su escala operacional, ya que a las limitaciones de la visión humana han de sumarse las limitaciones de resolución que los dispositivos presentan. Por ejemplo, no tiene sentido representar el callejero de una ciudad a una escala pequeña como la que se utilizaría para representar un mapa mundial, ya que cada pequeño punto de la pantalla contendría un gran número de calles. Además de obtener un resultado inservible, se consumirían recursos en efectuar todos los cálculos necesarios para producir esa representación.

En ocasiones, el proceso de generalización es necesario



Figura 3.3: Un ejemplo de generalización por agregación. Dos carreteras prácticamente paralelas y unidas se representan como dos elementos en el mapa, pero en el localizador de la parte superior izquierda, a escala de menor detalle, se generalizan como una única (Tomado de Yahoo Maps).

por razones distintas a las anteriores, y requiere operaciones también distintas. Por ejemplo, podemos crear un mapa del mundo que contenga vías de comunicación, pero no todas, sino solo las principales autopistas de cada país. En este caso, no vamos a encontrar problemas con distintas carreteras que se solapan en la representación, ni tampoco un volumen excesivo de datos, pero debemos igualmente «adaptar» la representación a la escala, es decir, efectuar algún tipo de generalización. En este caso, se representarían las carreteras con un ancho mayor del real, ya que, de otro modo, no serían apenas visibles.

La generalización, por tanto, es un proceso que tiene como objetivo la producción de una imagen cartográfica **legible y expresiva**, reduciendo el contenido del mapa a aquello que sea posible y necesario representar. Para ello, se enfatiza lo que resulta de importancia y se suprime lo que carece de ella.

Existen diversas operaciones que se emplean en el proceso de generalización. Algunas de las más relevantes son las **simplificación** (representar un elemento menos complejo),

la **agregación** (representar varios elementos como uno solo —Figura 3.3—), la **exageración** (representar elementos con mayor tamaño del que les corresponde) y el **desplazamiento** (representar en una posición modificada, para garantizar la legibilidad).

En un SIG, la generalización puede incorporarse como parte de los propios mecanismos de representación, aplicándose las transformación correspondientes en tiempo real. A partir de un juego de datos, se elaboran las representaciones según la escala a la que se estén representando. Esta solución tiene el inconveniente de producir resultados que no resultan óptimos, por ser la generalización un proceso complejo y difícil de automatizar, y, sobre todo, el de consumir gran cantidad de recursos. La generalización en este caso tiene un objetivo cartográfico, pero en lugar de hacer más fluido el trabajo con datos de gran volumen, lo hace más lento.

Una solución alternativa y más adecuada de incorporar la generalización dentro de un SIG suele basarse en un enfoque multi-escalar (Figura 3.4), en el cual se maneja información de una misma zona de estudio a diferentes escalas, y se usa en cada momento aquella que resulte más conveniente. Si se trabajara con cartografía en papel, sería equivalente a tener varios mapas de una zona a diferentes escalas.

El concepto de *capa*, que veremos en el capítulo 4 y que es vital para la idea actual de un SIG, permite este manejo simultáneo de información a distintas escalas.

En el caso de imágenes, este enfoque multi-escalar implica la creación de las denominadas **pirámides** (Figura 3.5).

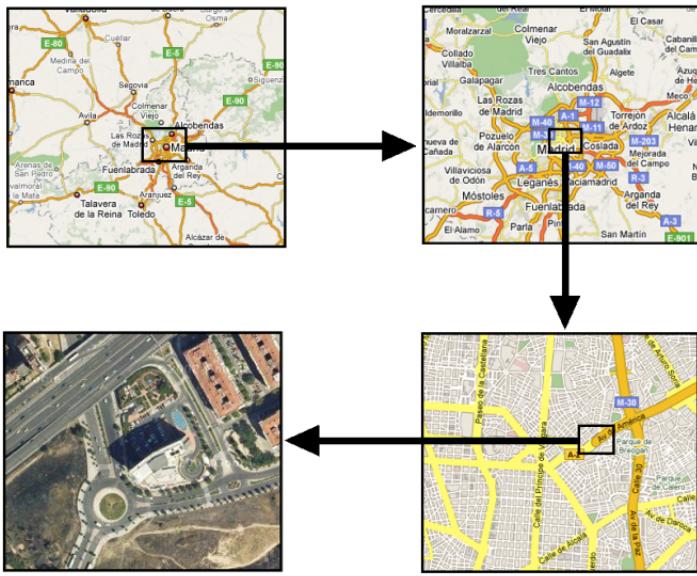


Figura 3.4: En un SIG es habitual manejar información a diferentes escalas. En función de la escala de representación, la información visualizada será una u otra.

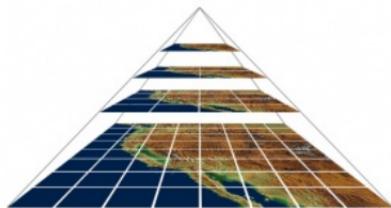


Figura 3.5: Pirámides de representación con imágenes preparadas a distintas escalas (Fuente: OSGeo).

EL DATO GEOGRÁFICO Y SU ALMACENAMIENTO

De todos los subsistemas de un SIG, el correspondiente a los datos es el pilar fundamental que pone en marcha los restantes. Los datos son el combustible que alimenta el SIG. El subsistema de datos es, a su vez, el más interrelacionado, y está conectado de forma inseparable a todos los restantes.

DATOS *vs* INFORMACIÓN. TIPOS DE INFORMACIÓN.

Existe una importante diferencia entre los conceptos de **datos** e **información**. Un SIG es un Sistema de *Información Geográfica*, pero maneja *datos* geográficos, existiendo diferencias entre estos conceptos.

Entendemos como dato al simple conjunto de valores o elementos que utilizamos para representar algo. Por ejemplo, el código 502132N es un dato.

El dato anterior podemos interpretarlo como si fuera una referencia geográfica, y cuyo significado sería entonces una latitud, en particular 50°21' 32'' Norte. Si lo interpretamos como un código que hace referencia a un documento de identificación de una persona, la información que nos aporta es en ese caso completamente distinta. El dato sería el mismo, formado por seis dígitos y una letra, pero la información que da es diferente, ya que lo entendemos e interpretamos de manera distinta.

La información es, por tanto, el resultado de un dato y una **interpretación**, y el trabajo con datos es en muchos casos un proceso enfocado a obtener de estos toda la información posible.

Comprender el significado y las diferencias entre datos e información permiten entender entre otras cosas que la relación entre los volúmenes de ambos no es necesariamente constante. Por ejemplo, los datos 502132NORTE o CINCUENTA VEINTIUNO TREINTAYDOS NORTE son mayores en volumen que 502132N, pero recogen la misma información espacial que este (suponiendo que los interpretamos como datos de latitud).

En la información geográfica se distinguen dos componentes: **espacial** y **temática**. La componente espacial hace referencia a la posición dentro de un sistema de referencia establecido, y responde a la pregunta *¿dónde?*. La componente temática responde a la pregunta *¿qué?*, y define la naturaleza del fenómeno que se produce en la localización indicada por la componente espacial.

Mientras que la componente espacial va a ser generalmente un valor numérico, pues son de esa naturaleza los sistemas de coordenadas que permiten expresar una posición concreta en referencia a un marco dado, la componente temática puede ser **numérica** o **alfanumérica** (texto). Una variable numérica puede a su vez ser de cuatro tipos: **nominal**, **ordinal**, **intervalo** o **razón**.

El tipo de variable condiciona las operaciones que pueden realizarse con un dato geográfico en función de cómo sea su componente temática.

Un concepto a tener en cuenta en relación con las componentes de la información geográfica es la **dimensión**. Los elementos que registramos pueden ir desde sencillos puntos (0D) hasta volúmenes tridimensionales (3D).

Las diferentes formas de representar y almacenar la información, que veremos más adelante en este capítulo, dependen del tipo de variable con que se trabaje.

DIVISIÓN DE LA INFORMACIÓN. CAPAS

En un SIG, la información espacial referida a una zona de estudio está dividida en varios niveles, de tal forma que, pese a coincidir sobre un mismo emplazamiento, información sobre distintas variables se encuentra recogida de forma independiente. Es decir, en función de la componente temática se establecen distintos bloques de datos espaciales. Cada uno de estos bloques temáticos se conoce como **capa** (Figura 4.1).

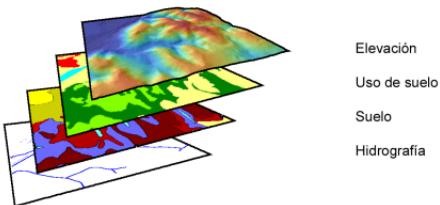


Figura 4.1: Concepto de *capa* de información geográfica dentro de un SIG

El concepto de capa es imprescindible para comprender todo SIG, y favorece la correcta estructuración de la información y el trabajo con ella. Toda la información geográfica con que trabajemos en un SIG va a ser en forma de capas. Cada una de ellas puede abrirse de forma independiente en un SIG y utilizarse por sí misma o en conjunto con otras.

Con la cartografía clásica, no es posible (o resulta difícil e impreciso) combinar distintos tipos de información, como por ejemplo la contenida en un mapa topográfico y la existente en un mapa de tipos de suelo y otro de vegetación potencial. En el caso de un SIG, los distintos tipos de información se pueden combinar de forma sencilla y limpia, y no aparecen los mismos problemas.

La relevancia del concepto de capa como elemento fundamental de un SIG es enorme, pues constituye el marco básico sobre el que se van a llevar a cabo gran parte de las operaciones. Por ejemplo, vimos en el apartado dedicado a la generalización cartográfica cómo en un SIG podemos

utilizar diferentes «versiones» de los datos correspondientes a una zona concreta, y representar una u otra de ellas en función de la escala de trabajo. Estas versiones se almacenarán como distintas capas. La capa es así la unidad fundamental no solo en términos de un área dada, sino también de una escala concreta, y permite una división de los datos óptima a todos los efectos.

La separación de la información en capas evita asimismo la redundancia de datos, ya que cada capa contiene un tipo de información concreto. En un mapa clásico se presentan siempre varias variables, algunas de ellas presentes con carácter general, tales como nombres de ciudades principales o vías más importantes de comunicación. En un SIG, al encontrarse estas variables separadas en sus correspondientes capas, es el usuario quien las combina.

El trabajo con capas permite, por tanto, una estructura **más organizada** y una **mayor atomización** de los datos, con las consecuentes ventajas en el almacenamiento, manejo y funcionalidad que esto conlleva.

Además de dividir la información geográfica en capas de acuerdo con su contenido, también dividimos esta con criterios puramente espaciales, «cortándola» en unidades menores que ocupen una región de amplitud más reducida. Este es un procedimiento similar al que encontramos en un mapa impreso, ya que el territorio de un país se encuentra cartografiado en diferentes *hojas*.

La principal cualidad de un SIG para integrar de forma transparente datos correspondientes a zonas distintas y formar un mosaico único es la **separación que existe entre datos y visualización**. Los datos son la base de la visualización, pero en un SIG estos elementos conforman partes del sistema bien diferenciadas. Esto quiere decir que los datos se emplean para crear un resultado visual pero en sí mismos no contienen valores relativos a esa visualización.

De este modo, es posible combinar los datos y después representarlos en su conjunto. Un proceso así no puede realizarse con un mapa ya impreso, pues este contiene ya elementos de visualización e incluso componentes cartográficos tales como una flecha indicando el Norte, una leyenda o

una escala. Por ello, aunque puedan combinarse, realmente no se «funde» la información de cada uno de los mapas para conformar uno único. En un SIG, por el contrario, la visualización de cuatro o más bloques de datos puede ser idéntica a la que obtendría si todos esos datos constituyeran un único bloque.

MODELOS PARA LA DE INFORMACIÓN GEOGRÁFICA

El proceso de convertir un área geográfica y la información acerca de ella en un dato susceptible de ser incorporado a un SIG puede dividirse en tres fases:

- Establecimiento de un **modelo geográfico**. Es decir, un modelo conceptual de la realidad geográfica y su comportamiento.
- Establecimiento de un **modelo de representación**. Es decir, una forma de recoger el anterior modelo conceptual y sus características propias, reduciéndolo a una serie finita de elementos.
- Establecimiento de un **modelo de almacenamiento**. Es decir, un esquema de cómo almacenar los distintos elementos del modelo de representación.

Por su mayor importancia, nos centraremos en los modelos de representación. Los modelos de representación que se utilizan principalmente en un SIG son dos: **modelo raster** y **modelo vectorial**. Las capas que utilizan estos modelos se conocen como **capas raster** y **capas vectoriales**, y esta es la terminología habitual en el ámbito de los SIG para referirse a la naturaleza de una determinada capa.

Modelo ráster

El modelo ráster se basa en una **división sistemática** del espacio. Todo el espacio queda cubierto y caracterizado como un conjunto de unidades elementales, cada una de ellas con un valor asociado.

Lo más habitual es una división en una malla de **celdas cuadradas** o rectangulares. Conociendo la orientación de la

malla y las dimensiones de cada una de las celdas, así como las coordenadas de al menos una de ellas, es posible conocer las coordenadas del resto en virtud de su **estructura regular**. Con esto, conocemos los valores de la variable en todos los puntos del espacio cubierto por la capa. El **tamaño de celda** es un parámetro relacionado con la escala de trabajo de la capa, ya que define la resolución de esta y está en función de la precisión con que se han tomado los datos correspondientes.

La figura 4.2 muestra un ejemplo de una malla ráster con valores de elevación.

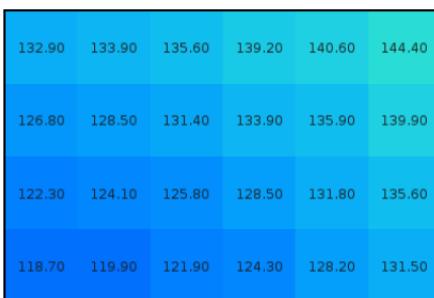


Figura 4.2: Celdas de una malla ráster con sus valores asociados.

El número de valores distintos recogidos para cada celda coincide con el número de las denominadas **bandas**. Una banda contiene un único valor en una capa raster. Puede entenderse una capa raster de más de una banda como un conjunto de capas (cada banda sería una subcapa de ese conjunto), teniendo en todas ellas la malla de celdas las mismas características espaciales, y presentándose el conjunto como un único elemento.

El ejemplo más claro de uso del modelo raster lo encontramos en las imágenes. Una imagen digital se compone de una malla de elementos (denominados **píxeles**, cada uno de los cuales tiene un color asociado). El conjunto de estos píxeles forman la imagen completa. Lo más habitual es que las imágenes contengan 3 bandas, correspondientes a las

intensidades de los colores rojo, verde y azul, las cuales al combinarse permiten obtener el color de cada píxel.

Otro uso habitual del modelo raster es en los denominados **Modelos Digitales de Elevaciones** (MDE), que recogen la topografía de un terreno. El modelo raster es el habitual de capa raster son las capas con modelos De forma general, los valores de una capa ráster, en cualquiera de sus bandas, son casi exclusivamente numéricos, no estando los SIG preparados para manejar otro tipo de valores en la componente temática de una capa ráster. De esta forma, una capa raster puede equipararse al concepto matemático de una **matriz**, con las ventajas que ello supone para aplicar sobre ella herramientas matemáticas a la hora de su análisis.

En el caso de los

Modelo vectorial

El otro modelo principal de representación es el modelo vectorial. En este modelo, no existen unidades fundamentales que dividen la zona recogida, sino que se recoge la variabilidad y características de esta mediante **entidades**, para cada una de las cuales dichas características son constantes. Las entidades se componen de **primitivas geométricas**, y estas pueden ser de tres tipos: **puntos, líneas y polígonos**.

Utilizando puntos, líneas o polígonos, puede modelizarse el espacio geográfico si se asocia a estas geometrías una serie de valores definitorios. Una entidad puede tener **varias primitivas**. Por ejemplo, en una capa de países, necesitaríamos varios conjuntos para representar España si queremos incluir tanto la península como las islas que la forman. Todos estos polígonos constituyen una única entidad, ya que todos pertenecen al mismo país y tendrán el mismo conjunto de valores asociados.

A la hora de definir las formas geométricas básicas, todas ellas pueden en última instancia **reducirse a puntos**. Así, las líneas son un conjunto de puntos interconectados en un determinado orden, y los polígonos son líneas cerradas, también expresables por tanto como una serie de puntos. Todo elemento del espacio geográfico queda definido, pues,

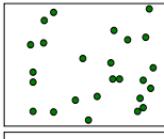
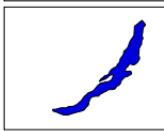
Primitiva	Entidad espacial	Representación	Atributos																		
Puntos			<table border="1"> <thead> <tr> <th>ID</th> <th>Altura</th> <th>Diametro Normal</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>17.5</td> <td>36</td> </tr> <tr> <td>2</td> <td>22</td> <td>45.8</td> </tr> <tr> <td>3</td> <td>15</td> <td>27.2</td> </tr> <tr> <td>4</td> <td>19.7</td> <td>36.1</td> </tr> <tr> <td>.</td> <td>.</td> <td>.</td> </tr> </tbody> </table>	ID	Altura	Diametro Normal	1	17.5	36	2	22	45.8	3	15	27.2	4	19.7	36.1	.	.	.
ID	Altura	Diametro Normal																			
1	17.5	36																			
2	22	45.8																			
3	15	27.2																			
4	19.7	36.1																			
.	.	.																			
Líneas			<table border="1"> <thead> <tr> <th>Ancho máx(m)</th> <th>Calado máx(m)</th> <th>Longitud(km)</th> </tr> </thead> <tbody> <tr> <td>15</td> <td>4.3</td> <td>35</td> </tr> <tr> <td>6.3</td> <td>3.9</td> <td>5.2</td> </tr> </tbody> </table>	Ancho máx(m)	Calado máx(m)	Longitud(km)	15	4.3	35	6.3	3.9	5.2									
Ancho máx(m)	Calado máx(m)	Longitud(km)																			
15	4.3	35																			
6.3	3.9	5.2																			
Polygones			<table border="1"> <thead> <tr> <th>Superficie(km²)</th> <th>Profundidad máx(m)</th> </tr> </thead> <tbody> <tr> <td>31494</td> <td>1637</td> </tr> </tbody> </table>	Superficie(km ²)	Profundidad máx(m)	31494	1637														
Superficie(km ²)	Profundidad máx(m)																				
31494	1637																				

Figura 4.3: Primitivas geométricas en el modelo de representación vectorial y ejemplos particulares de cada una de ellas con atributos asociados

por una serie de puntos que determinan sus propiedades espaciales y una serie de valores asociados.

Dentro de un SIG, una capa vectorial puede contener **un único tipo de primitiva**. Así, tenemos capas vectoriales de puntos, de líneas y de polígonos, respectivamente. Una variable puede recogerse con varios tipos de primitivas (por ejemplo, puede indicarse una ciudad con un punto o con un polígono que delimita su perímetro), y la elección de uno u otro tipo de geometría ha de ser función del tipo de fenómeno que se pretende modelizar o la precisión necesaria, entre otros factores.

La componente temática en el modelo vectorial se establece mediante los denominados **atributos**, que suelen ser múltiples, a diferencia de lo que sucede en el modelo raster, donde lo habitual es tener un único valor para cada celda. Los atributos de una capa vectorial pueden contener información de cualquier clase, siendo más versátiles que en el caso de las capas ráster, donde ya vimos que se maneja únicamente información numérica. Por su estructura particular (series de atributos asociados a una entidad), la

componente temática en el modelo vectorial se presta especialmente a representarse en tablas y almacenarse en una **base de datos**, y puede analizarse independientemente de la componente espacial.

Un elemento particular del modelo de representación vectorial es la **topología**. Una capa vectorial contiene topología si en ella se almacenan de algún modo las relaciones espaciales que existen entre sus elementos. Disponer de topología en una capa vectorial es de gran importancia a la hora de llevar a cabo ciertos tipos de análisis, así como procedimientos tales como la edición de los propios datos geográficos.

Aunque la mayoría de operaciones con una capa vectorial pueden llevarse a cabo en ausencia de topología, algunas de ellas como el **análisis de redes** no se pueden llevar a cabo sin topología. Si pensamos en una capa de vías sobre la que desarrollar ese análisis de redes, un mero conjunto de elementos geométricos (líneas en este caso), no nos da información sobre los posibles enlaces entre las vías que quedan representadas. Los puntos donde se cruzan dos vías pueden ser cruces o rotundas (es decir, puede pasarse de una vía a otra, existiendo conexión entre ellas), o bien pasos elevados o subterráneos donde una de las vías pasa por encima de la otra (y por tanto no existe comunicación entre ambas). Las circunstancias son muy distintas en función del tipo de cruce que exista, y por ello es imprescindible conocer esta información para efectuar un análisis de redes correcto (Figura 4.4)

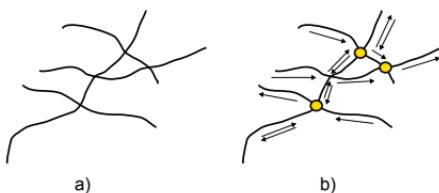


Figura 4.4: Capa de vías de comunicación sin topología (a) o con ella (b). Los puntos en este segundo caso indican conexiones entre vías, y son una representación visible de la topología existente.

El almacenamiento de entidades basado en una mera lista de coordenadas de cada entidad, sin topología, se conoce popularmente como *spaghetti*, pues si pensamos en una capa de líneas sin topología que se entrecruzan en el espacio, esta se asemejan en cierta forma a un caótico plato de *spaghetti* sin orden ni relación entre ellos.

Raster vs vectorial

Tanto el modelo raster como el vectorial pueden emplearse para recoger **cualquier tipo de información**. La figura 4.5 muestra un ejemplo de esto, representando una capa de vías según ambos modelos. Otro ejemplo para mostrar esto lo encontrarmos en las capas de elevaciones, que ya hemos visto que suelen recogerse en capas raster, en especial si se va a desarrollar sobre ellas algún tipo de análisis. No obstante, pueden recogerse también como una capa vectorial de puntos (este es un caso habitual si se obtienen los datos de un levantamiento topográfico), o bien como una capa de líneas que contenga **curvas de nivel**, entre otras opciones.

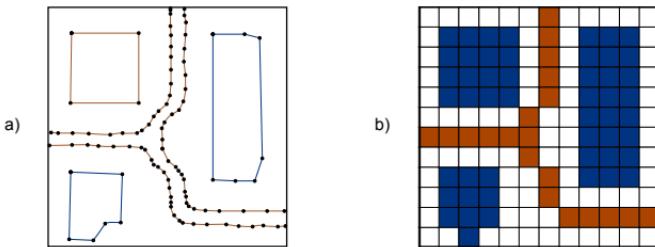


Figura 4.5: Comparación entre los esquemas del modelo de representación vectorial (a) y ráster (b).

Resulta obvio que las diferencias entre los modelos ráster y vectorial son muy notables, y que cada uno de ellos posee sus propias ventajas e inconvenientes. Algunos aspectos a los cuales puede atenderse para comparar uno y otro modelo son los siguientes:

- **Planteamiento.** El modelo ráster hace más énfasis en aquella característica del espacio que analizamos (*qué* y

cómo), mientras que el modelo vectorial da prioridad a la localización de dicha característica (*dónde*)

- **Precisión.** El modelo ráster tiene su precisión limitada por el tamaño de celda. Las entidades menores que dicho tamaño de celda no pueden recogerse, y la variación espacial que sucede dentro del espacio de la celda tampoco. Asimismo, existe una imprecisión en las formas. El detalle con el que puede recogerse la forma de una entidad geográfica según el modelo vectorial es, en la práctica, ilimitado, mientras que, como puede verse en la imagen 4.6, el modelo ráster restringe las formas a ángulos rectos, ya que la unidad base es un cuadrado.

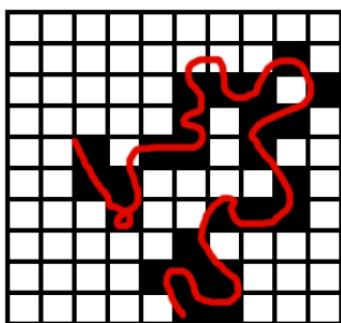


Figura 4.6: Imprecisión de forma en el modelo de representación ráster. La división del espacio en unidades cuadradas impide la representación fiel de entidades tales como curvas.

- **Complejidad.** La regularidad y sistematicidad de las mallas ráster hacen sencillo el implementar algoritmos de análisis, muy especialmente aquellos que implican el uso combinado de varias capas. Por el contrario, la irregularidad espacial de las capas vectoriales hace que la implementación de los mismos algoritmos sea sumamente más compleja si se trabaja con estas capas.

No existe un modelo de representación idóneo de forma global, sino que esta idoneidad depende de muchos factores, como por ejemplo:

- **Tipo de variable o fenómeno a recoger.** Las variables **continuas** tales como la elevación es más adecuado en general recogerlas en capas raster, para así facilitar su análisis, mientras que las variables **discretas** es preferible almacenarlas como capas vectoriales.
- **Tipo de análisis o tarea a realizar sobre dicha variable.** El uso que demos a una capa temática condiciona en gran medida el modelo de datos idóneo. Por ejemplo, en el caso de una capa de elevaciones, su análisis se lleva mejor a cabo si esta información está recogida según el modelo ráster. Sin embargo, si el objetivo principal es la visualización de esa elevación en conjunto con otras variables, unas curvas de nivel pueden resultar más adecuadas, ya que, entre otras cosas, no interfieren tanto con otros elementos a la hora de diseñar un mapa con todas esas variables.
- **Contexto de trabajo.** Por ejemplo, si queremos trabajar con imágenes, esto nos condiciona al empleo de datos ráster, ya que resulta mucho más sencillo combinarlos con las imágenes, las cuales siempre se presentan como capas ráster.

Existen procedimientos para **convertir** entre los formatos ráster y vectorial, de forma que el disponer de datos en un modelo de representación particular no implica que debamos desarrollar nuestro trabajo sobre dichos datos directamente, sino que podemos efectuar previamente una conversión. Los capítulos ?? y ?? tratan estos temas en profundidad.

FUENTES PRINCIPALES DE DATOS ESPACIALES

No hace tanto tiempo, toda la información que se manejaba dentro de un SIG tenía su origen en un mapa en papel, el cual debía *prepararse* para adaptarse a la naturaleza propia del SIG. El dato geográfico se obtenía a partir de la **digitalización** de cartografía, es decir, convertir los datos geográficos en formato impreso en datos en formato digital que un SIG pudiera manejar.

Un SIG implica una aplicación informática, y esta se alimenta en última instancia exclusivamente de datos digitales. Los datos geográficos digitales tienen una serie de ventajas frente a los analógicos (además del mero hecho de que podemos incorporarlos a nuestro SIG), y suponen, como sucede en muchos otros campos, un salto cualitativo importante. Entre estas ventajas, que son a su vez comunes a otros ámbitos, destacan la sencillez de actualización, la facilidad de distribución (en especial con la aparición de Internet), el menor espacio físico necesario para su almacenamiento, la facilidad y precisión del análisis, y la facilidad de mantenimiento (el dato digital no se degrada, lo que se degrada es su soporte, pero es sencillo replicar el dato sin pérdida de calidad)

Hoy en día las técnicas de adquisición de datos han evolucionado y permiten crear datos que pueden ser directamente

integrados en un SIG. Distinguimos así **fuentes de datos primarias y secundarias**.

Las fuentes de datos primarias son aquellas cuyos datos podemos emplear en un SIG, ya que estos, en su forma original, ya son susceptibles de ser sometidos a las operaciones de manejo y análisis que incorporan los SIG. Por su parte, las fuentes de datos secundarias generan datos que no pueden emplearse en un SIG sin un proceso de adaptación previo, siendo el dato derivado el que utilizamos en un SIG, no el original.

En este capítulo, veremos las distintas fuentes de datos con las que podemos trabajar un en SIG.

TELEDETECCIÓN

La primera fuente de datos que trataremos en este capítulo es la teledetección. Entendemos por teledetección el estudio y medida de las características de una serie de objetos (en nuestro caso elementos de la superficie terrestre) sin que exista contacto físico. Para ello, se miden las perturbaciones que el objeto provoca en su entorno, principalmente las de tipo electromagnético.

Un sistema de teledetección cuenta con los siguientes elementos (Figura 5.1):

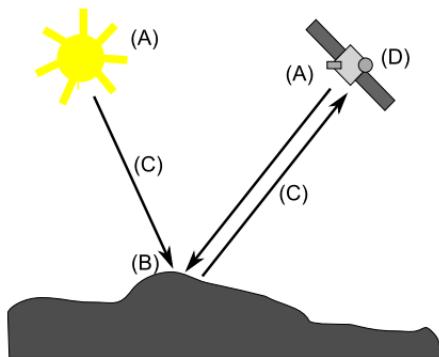


Figura 5.1: Esquema de un sistema de teledetección.

- **Una fuente de radiación (A).** Puede ser de origen natural o artificial. La radiación emitida por dicha fuente llega al terreno y sufre una perturbación causada por los elementos de este, siendo esta perturbación el objeto de estudio de la teledetección. Los propios objetos pueden ser también emisores ellos mismos de radiación.
- **Unos objetos (B) que interaccionan con la radiación** o la emiten, según lo anterior.
- **Una atmósfera (C)** por la que se desplaza la radiación, tanto desde la fuente hasta el objeto como desde el objeto hasta el receptor. La atmósfera también interactúa con la radiación, introduciendo igualmente perturbaciones en ella.
- **Un receptor (D) que recoge la radiación** una vez esta ha sido perturbada o emitida por los objetos. El receptor va a generar como producto final una imagen (en términos de un SIG, una capa ráster), en cuyas celdas o píxeles se va a contener un valor que indica la intensidad de la radiación. Estos valores son valores enteros que indican el nivel de dicha radiación dentro de una escala definida (habitualmente valores entre 1 y 256), y se conocen dentro del ámbito de la teledetección como **Niveles Digitales**.

A lo largo de este apartado veremos con detalle estos elementos. Para estudiar los dos primeros, estudiaremos los fundamentos físicos relativos a la radiación y a la interacción entre esta y la materia, mientras que para el estudio del sistema receptor analizaremos los elementos de este en dos componentes por separado: sensores y plataformas.

La interacción de la atmósfera interesa de cara a eliminar su efecto, ya que lo que resulta de interés en general son los objetos en la superficie terrestre, no la atmósfera como tal. Eliminar esta influencia de la atmósfera es parte de los procesos posteriores que se realizan con la imagen, y que se detallan en los capítulos dedicados al análisis de imágenes en lugar de en este.

Fundamentos físicos

La radiación electromagnética es producto de las alteraciones en los campos eléctrico y magnético, las cuales generan ondas correspondientes a cada uno de los campos magnético y eléctrico. Estas ondas se desplazan a la velocidad de la luz, y se pueden describir con los parámetros habituales, tales como la longitud de onda o la frecuencia. El rango de longitudes de onda cubierta por la radiación electromagnética se conoce como **espectro electromagnético**.

El espectro se subdivide en regiones en función de su longitud de onda, tales como (de menor a mayor longitud de onda) los rayos γ , los rayos X, la región ultravioleta, la región visible, la región infrarroja, o las microondas

La radiación emitida por una fuente de radiación es alterada por la presencia de los distintos objetos, ya que estos **absorben, transmiten o reflejan** esta.

Estos tres fenómenos se dan en diferente proporción en función de las características del objeto y de la radiación. La parte que interesa a efectos de la teledetección es aquella que se refleja en el objeto, ya que esta es la que posteriormente puede recogerse y emplearse para la generación de las imágenes.

Como ya se dijo en el capítulo 4, las imágenes como capas ráster presentan habitualmente la particularidad de tener varias bandas. En lugar de un único valor para cada celda, existen n valores, uno por cada banda. La imagen recoge la **intensidad de la radiación** dentro de una amplitud dada del espectro, y a su vez subdivide esta en distintas franjas. Los Niveles Digitales de cada banda corresponden a la intensidad dentro de una de esas franjas del espectro en particular.

Puesto que cada objeto refleja la radiación de diferentes longitudes de onda de modo distinto, esto puede considerarse como una propiedad del objeto. Se tiene así el concepto de **firma espectral**, que es la respuesta característica de un tipo de objeto dentro del espectro electromagnético.

Sensores y plataformas

En un sistema de teledetección, dos son los elementos tecnológicos principales que lo definen: el **sensor** y la **plataforma**.

El sensor es el elemento que incorpora la capacidad de «leer» la radiación electromagnética y registrar su intensidad dentro de una zona concreta del espectro. Puede ir desde una simple cámara fotográfica hasta un sensor más especializado.

Los sensores se denominan **pasivos** aprovechan las fuentes de radiación existentes en la naturaleza (fundamentalmente el Sol) y se limitan a recoger la radiación de dichas fuentes reflejada por los elementos del medio, o **activos** sí emiten radiación, y recogen dicha radiación tras ser reflejada por dichos elementos. Para entender este concepto de un modo de un modo sencillo, podemos decir que una cámara fotográfica es un sensor pasivo, mientras que una cámara fotográfica con *flash* es un sensor activo. La radiación emitida por los sensores activos no ha de ser necesariamente luz visible (como en el caso del *flash*), sino que pueden emitir en otras secciones del espectro.

Tecnologías como el **radar** o el **LiDAR** (similar al radar pero con pulsos de laser en lugar de ondas de radio), se basan en sensores activos. En el caso del LiDAR, permite obtener imágenes que no tiene un carácter visual, sino que contienen en sus valores la elevación de los objetos, pudiendo así emplearse para cartografiar el relieve.

La plataforma, por su parte, es el medio en el que se sitúa el sensor y desde el cual se realiza la observación. A bordo de una plataforma pueden montarse **varios sensores**.

Los dos tipos principales de plataformas son aquellas situadas **dentro de la atmósfera** terrestre (aviones en su mayoría) y aquellas situadas **fuerza de la atmósfera** (a bordo de satélites).

Los aviones tienen la ventaja de su **disponibilidad**, ya que pueden pilotarse y de este modo permiten cubrir cualquier lugar de la tierra en cualquier momento.

A diferencia de un avión, un satélite no puede dirigirse

a voluntad (no puede pilotarse), y su movimiento es una característica inherente que viene definida por una serie de parámetros. Estos parámetros se conocen como **parámetros orbitales** pues definen la órbita descrita por el satélite en torno a la Tierra.

Las órbitas pueden clasificarse en función de su **eje de rotación** así como en función de su **movimientos**. En este último caso tenemos órbitas **geosíncronas** (el satélite se sitúa sobre un punto fijo de la Tierra y su movimiento sigue al de rotación de esta) o **heliosíncronas** (mientras el satélite recorre la órbita, la Tierra efectúa su movimiento de rotación, lo cual hace que a cada vuelta de la órbita se cubran zonas distintas)

Resoluciones

Uno de los parámetros principales que definen las propiedades de un sistema de teledetección son las *resoluciones*. Estas establecen el nivel de detalle de los productos que el sistema genera, determinando este en las distintas magnitudes en las que el sistema opera. Las resoluciones dependen del sensor y de la plataforma como binomio operativo, y de las características propias de ambos. Distinguimos cuatro resoluciones, a saber:

- **Resolución espacial.** Indica la dimensión del objeto más pequeño que puede distinguirse en la imagen. Es la dimensión real que un píxel de la imagen tiene sobre el terreno.
- **Resoluciónpectral.** Indica la amplitud de cada una de las regiones del espectro que se recogen en la imagen. La región del espectro abarcada y el número de bandas son los elementos que definen la resoluciónpectral.
- **Resolución radiométrica.** Para cada una de las bandas que produce un sensor (asociada esta a una determinada región del espectro según su resoluciónpectral), el dato recogido, que constituye su Nivel Digital, indica la intensidad correspondiente a esa región. El nivel de detalle con

el que puede medirse esa intensidad es el que define la resolución radiométrica del sensor.

- **Resolución temporal.** Indica el tiempo que tarda el sensor en volver a tomar una imagen de una misma zona. Tiene sentido en el caso de sensores orbitales. La resolución temporal depende de la altura a la que se encuentra la plataforma que monta el sensor, así como la resolución espacial.

No resulta posible (por razones técnicas y teóricas) disponer de un sensor en el cual todas las anteriores regiones sean altas. Algunos sensores priman determinadas resoluciones, mientras que otros favorecen otras distintas.

A la hora de utilizar imágenes de teledetección, debe considerarse qué tipo de resolución resulta de mayor interés (por ejemplo, para localizar elementos de pequeño tamaño son necesarias imágenes con alta resolución especial). En base a esto, se escogerá una u otra clase de imágenes, que será la que ofrezca los valores de resolución más adecuados en conjunto. La utilización simultánea de datos de varios sensores en un proyecto es una alternativa para compensar este hecho.

Fotogrametría

Relacionada con la teledetección, encontramos la **fotogrametría**. La fotogrametría es la técnica para estudiar y definir con precisión la forma, dimensiones y posición en el espacio de un objeto cualquiera, utilizando medidas realizadas sobre una o varias fotografías. En el campo del SIG, es de especial interés la **fotogrametría aérea**, cuya base de trabajo tradicional son las fotografías aéreas, y que sirve principalmente para la creación de cartografía de elevaciones a partir de un proceso de **restitución**.

En lugar de imágenes individuales, la fotogrametría emplea pares de imágenes, cada una de ellas tomada desde un punto distinto. Mediante estereoscopía, resulta posible recrear el efecto que ambas imágenes tendrían para la reconstrucción tridimensional de la escena, y «engañar» al cerebro

del observador para que este pueda observar la escena con **volumen y profundidad**. Esto permite posteriormente conocer las formas del terreno, y a partir de ello crear las capas correspondientes, con información de elevaciones.

Cuando se emplean imágenes de satélite, los pares se pueden obtener con aquellas plataformas y sensores que permiten **variar el ángulo de visión**, de modo que en la misma pasada del satélite se toman imágenes de una zona desde distintos puntos.

Según su enfoque, la fotogrametría puede ser **analógica, analítica o digital**, siendo esta última la que se basa en el trabajo con imágenes digitales dentro de un entorno computerizado, y la que mayor relación tiene con los SIG.

La fotogrametría requiere de una *estación fotogramétrica*, que en el caso digital incorpora muchos elementos propios de un SIG, así como otros específicos del trabajo fotogramétrico. Entre estos, caben destacar aquellos que permiten la generación de visualizaciones con sensación de profundidad, así como los periféricos específicos tales como ratones 3D o manivelas similares a las que presentan los restituidores analíticos, facilitando así la adaptación de los operarios a este tipo de estación.

CARTOGRAFÍA IMPRESA. DIGITALIZACIÓN

Existe gran cantidad de cartografía en formato no digital, tales como mapas impresos o fotografías aéreas antiguas en formato analógico.

La **digitalización** de esta cartografía, necesaria para su uso en un SIG, supone la creación de capas raster o vectoriales a partir de ella. En este último caso, implica la **disgregación** de la información que contiene, ya que en un mapa clásico se presentan distintas informaciones que en el uso común de un SIG se almacenan en capas independientes.

La digitalización implica tres etapas: el **registro o georeferenciación** (Establecimiento del marco geográfico —sistema de coordenadas, puntos de control, etc—, de forma que los elementos digitalizados posteriormente sean correctos), la

digitalización de la componente espacial (creación de geometrías en el caso de capas vectoriales), y la digitalización de la componente temática (creación de valores de celda en el caso de capas raster o de atributos en el caso vectorial).

La digitalización puede ser **manual** o **automática**. En el primer caso, es un operario quien introduce los valores de los elementos digitalizados, mientras que en el segundo caso es un procedimiento automatizado el que se encarga de ello.

Para la creación de capas raster, lo habitual es la digitalización automática, a partir del **escaneo** del documento original.

El escaneo es el proceso de digitalización que convierte una imagen impresa (analógica) en una imagen digital. Dentro de un SIG, se emplea para digitalizar tanto mapas como fotografías aéreas, y su resultado es una capa ráster.

Aunque existen escáneres específicamente diseñados para el trabajo con documentos cartográficos, estos son dispositivos muy especializados y de muy elevado coste. Los escáneres más genéricos, pensados para el trabajo con todo tipo de imágenes y para todo tipo de usos, pueden no obstante emplearse de igual modo para escanear tanto mapas como imágenes aéreas con resultados aceptables, utilizándose con frecuencia.

Los parámetros básicos que definen las características de un escáner son la **resolución espacial** y la **resolución radiométrica**. La primera de estas se mide habitualmente en **puntos por pulgada** y nos indica el número de puntos (celdas) que el sensor es capaz de tomar por cada unidad de longitud sobre el papel. La resolución radiométrica, por su parte, indica la capacidad del sensor para distinguir entre dos colores distintos.

En relación con las resoluciones de escaneo, no deben olvidarse los fundamentos cartográficos que se explicaron en el capítulo 3. Trabajando con una resolución más elevada no hace necesariamente que estemos incorporando más información, ya que esta puede no existir en el documento original. Tendríamos un volumen de datos más elevado que el necesario para recoger toda la información del mapa.

En el caso de capas vectoriales, es más habitual la **digi-**

talización manual. En ella, un operario va definiendo las entidades, trazando la forma de estas o, en caso de ser una entidad de tipo punto, indicando su localización.

Para llevar a cabo ese trazado de la entidad, se necesita emplear algún equipo que recoja la información introducida por el operador. Existen dos alternativas principales: utilizar una **tableta digitalizadora** diseñada específicamente para la digitalización, o bien digitalizar utilizando las **funciones de edición** de un SIG, realizando todo el proceso dentro de este y sin más herramientas que el propio ordenador y un dispositivo señalador como el ratón. En este segundo caso, la digitalización se produce sobre la pantalla, por lo que es necesario tener ya una versión digital del documento (aunque no en forma de capa vectorial), la cual puede obtenerse mediante escaneo.

La digitalización automática de entidades para formar una capa vectorial, conocida como **vectorización**, resulta también posible, aunque presenta más dificultades. El principal problema reside en la necesidad de una **adecuada preparación** del documento a digitalizar, ya que las condiciones de este afectan notablemente a la calidad del resultado. La vectorización, al igual que la digitalización manual, puede llevarse a cabo mediante dispositivos especializados, o bien mediante el software apropiado, el cual se ejecuta en un ordenador y trabaja con una imagen escaneada del documento original.

Un caso particular de digitalización es la creación de **capas a partir de valores o coordenadas**. Es decir, cuando no existe un mapa o documento cartográfico, sino simplemente una serie de datos espaciales expresados de forma alfanumérica que son susceptibles de convertirse en una capa y emplearse así dentro de un SIG. Este proceso se conoce como **geocodificación** e implica la asignación de coordenadas a puntos de interés tales como muestras de campo levantamientos topográficos, inventarios de lugares donde tiene lugar un determinado fenómeno o existe un determinado elemento, u otros como el *geotagging* de fotografías, para asociar a estas la posición del enclave en el que se han tomado.

En el caso de encontrarse en formato analógico, estos datos pueden digitalizarse mediante la simple introducción manual de coordenadas a través del teclado o bien mediante algún sistema más específico como el escaneo del documento y el empleo de algún software de reconocimiento de caracteres (OCR).

En el caso de encontrarse ya en formato digital, estos datos pueden presentarse como tablas en una hoja de cálculo, datos asociados a otro dato de cualquier tipo (como en el caso del *geotagging*) o incluso simples archivo de texto. Muchos SIG incorporan métodos para leer estos archivos y después utilizar las coordenadas que contienen con el fin de crear una nueva capa, en general de puntos.

Calidad de la digitalización

Uno de los aspectos más importantes del proceso de digitalización es la **calidad del resultado** obtenido, que debe tratar de ser lo más cercano posible a la calidad original de la información que se digitaliza, es decir, del mapa o imagen original. Independientemente de la precisión del equipo utilizado o la habilidad y experiencia del operario, la digitalización no es por completo perfecta, conteniendo siempre ciertas deficiencias y errores.

Además de los errores que puedan incorporarse en las distintas fases del proceso de digitalización (sea este del tipo que sea), hay que considerar que las fuentes originales a digitalizar también pueden incluir los suyos propios. Así, el proceso de escaneado puede incorporar distorsiones geométricas, pero es posible que el mapa o fotografía aérea de partida también presente alguna distorsión como consecuencia de su deterioro, más patente cuanto más antigua sea esta.

La información contenida en el documento cartográfico puede también contener elementos problemáticos de cara a obtener un producto de calidad, que pueden ir desde líneas borradas total o parcialmente a manchas en el propio mapa derivadas de su uso habitual.

Dentro de los errores que aparecen como consecuencia

de la digitalización en sí, un tipo importante de ellos son las **discrepancias y coincidencias imperfectas** entre las distintas entidades, tal como las que se muestran en la figura 5.2

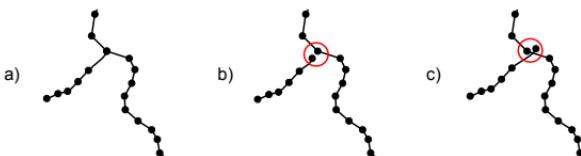


Figura 5.2: Errores derivados del proceso de digitalización. a) Versión correcta, con nodos coincidentes. b) y c) Versiones con errores que causan una falsa desconexión entre las líneas.

Debido a esto, las capacidades de edición de los SIG incorporan funcionalidades que permiten evitar estos errores en el momento de la digitalización, ayudando al operario en su tarea y permitiéndole alcanzar una exactitud y precisión imposible de lograr sin estas funcionalidades. Entre ellas, es especialmente importante el establecimiento de tolerancias y ajuste automático en función de ellas (esto se conoce con el término inglés **snapping**), que ayudan a garantizar la coincidencia entre los distintos vértices.

El hecho de que exista una completa coincidencia es especialmente importante cuando la capa vectorial que se digitaliza contiene **información topológica**. La topología exige que la coincidencia sea correcta y defina perfectamente la relación entre las entidades. La digitalización de entidades, en caso de querer recoger su topología, debe realizarse siguiendo reglas adicionales tales como la digitalización una única vez de los lados comunes de polígonos. Debe, asimismo, recoger información adicional, como por ejemplo la necesaria para la definición de nodos en los cruces entre líneas cuando existe relación entre estas.

GPS

Uno de los hitos en la aparición de nuevas fuentes de datos geográficos es la aparición de los **Sistemas Globales**

de Navegación por Satélite. Se trata de sistemas que permiten **conocer en todo momento y en cualquier punto del globo la localización exacta de dicho punto** con un margen de error del orden de unos pocos metros o menos. Para ello, se basan en el envío de señales entre un dispositivo situado en el punto concreto y una red de satélites, pudiendo establecerse la posición exacta mediante las características de dicha transmisión.

El ejemplo más extendido de estos sistemas es el **Sistema de Posicionamiento Global** (Global Positioning System, o **GPS**). El GPS cuenta con una constelación de 24 satélites activos, así como estaciones terrestres que los controlan, y su funcionamiento se basa en la **triangulación de la posición** de una unidad receptora, mediante las señales procedentes de un cierto número de satélites. Esta triangulación se basa en distancias entre la unidad receptora y dichos satélites, las cuales se calculan mediante diversos mecanismos. La posición se calcula no únicamente en sus coordenadas *x* e *y*, sino también en *z*, es decir en elevación. El sistema GPS emplea como sistema geodésico de referencia el WGS84.

El diseño de la red de satélites está pensado para garantizar que en cualquier punto de la superficie terrestre y en cualquier momento, un receptor puede localizar el número necesario de satélites para obtener con exactitud su precisión.

Existen **numerosas fuentes de error** que causa desviaciones apreciables en el cálculo de coordenadas mediante GPS. Entre ellas destacan los errores en la posición de los satélites, errores por el rebote de la señal en otros con anterioridad a alcanzar el receptor, errores por el paso de la señal por la atmósfera, así como los de precisión de los relojes empleados para el cálculo de distancias. La **disponibilidad selectiva** era un error aleatorio introducido en la señal GPS con fines militares, pero fue eliminada en el año 2000.

Entre las técnicas empleadas para corregir estas desviaciones, destaca el denominado **GPS diferencial**, pensado en origen para eliminar el error de la disponibilidad selectiva,

aunque también eficaz para corregir una buena parte los restantes errores citados anteriormente.

Para la aplicación del GPS diferencial se requiere no solo un receptor único (aquel del cual se quiere calcular su posición), sino también otro **receptor fijo** de referencia cuyas coordenadas se conocen con alta precisión. Este receptor fijo es, a su vez, un receptor de alta precisión y, además de calcular su propia posición, emite información que las unidades receptoras pueden aprovechar para corregir sus mediciones. El receptor móvil, lógicamente, tiene que estar capacitado para este tipo de correcciones, para así poder hacer uso de la señal de la estación de referencia.

El fundamento de esta técnica es que los errores que afectan al receptor móvil **también afectan al de referencia**. No obstante, la magnitud del error que afecta al receptor de referencia puede conocerse, ya que se conoce la coordenada exacta de este, y en base a eso puede eliminarse el error que afecta al receptor móvil, asumiendo que ambos errores son de similar índole.

En la actualidad, aplicando estas técnicas de corrección diferencial, un GPS puede obtener precisiones del orden de 2 metros en latitud y longitud, y 3 en altitud[?]. Sin corrección diferencial, esta precisión es de unos 10–20 metros.

La precisión del sistema GPS depende del tipo de receptor GPS (o, en el lenguaje común, GPS a secas) que se emplee, obteniéndose mayores precisiones con receptores más avanzados, siempre dentro de las posibilidades del propio sistema GPS. Existen muchas clases de receptores GPS, siendo dos de ellas las principales en relación con los SIG:

- **GPS para uso general.** Unidades **pequeñas y portátiles**, de bajo coste, para actividades al aire libre, donde no se requiere una precisión elevada sino simplemente un conocimiento de la posición aproximada. Se emplean, por ejemplo, para recoger rutas en senderismo o navegación.
- GPS para la medición topográfica. Unidades de medio tamaño, generalmente con una antena independiente que se conecta a la unidad y que el propio operario carga a la espalda. La antena garantiza **mayor precisión** y

una mejor localización de satélites en condiciones tales como zonas bajo arbolado. Están pensados para un uso profesional en levantamientos o replanteos, ofreciendo buena precisión en todas las coordenadas.

La capacidad principal de una unidad GPS en relación con un SIG es la de **recoger coordenadas**. Esta funcionalidad permite almacenar puntos o trazados completos, encontrándose el operario inmóvil o bien en movimiento a lo largo de dicho trazado. Es habitual utilizar los vocablos ingleses de la terminología GPS para denotar los distintos elementos que pueden recogerse, conociéndose a un punto de interés aislado como **waypoint** y un trazado como **track**. Una serie ordenada de *waypoints* se conoce como **route** (ruta).

En el trabajo con el receptor GPS, el operario se puede detener en un punto cualquiera y memorizar las coordenadas del mismo, añadiendo así un *waypoint* a la lista de los ya almacenados. Para crear un trazado, se suele disponer de funcionalidades de recogida automática de puntos, de tal modo que el receptor memoriza estos a intervalos fijos de tiempo. El operario simplemente ha de desplazarse por el trazado y dejar que el receptor haga su trabajo mientras tanto.

INFORMACIÓN GEOGRÁFICA VOLUNTARIA

Hemos mencionado ya que los dispositivos tales como receptores GPS de bajo coste pueden emplearse para recoger información geográfica y crear datos geográficos, y que cuando esto se une a los conceptos participativos de la denominada **Web 2.0**, surgen iniciativas de gran interés en las que el usuario de a pie, sin necesidad de una formación específica como cartógrafo, puede aportar sus datos para que otros los exploten posteriormente. Aunque no se trata de una fuente de datos como tal, y los elementos y dispositivos empleados ya los hemos visto a lo largo de este capítulo, el cambio que supone la inclusión de una filosofía acorde con las ideas de la Web 2.0 es tan notable que merece ser

tratado por separado. No se trata de un cambio en la propia toma o preparación de datos, o de una tecnología nueva que se aplique a estos, sino de un **cambio social y filosófico** que redefine el propio concepto de la información geográfica en lo que a la creación del dato geográfico respecta, y cuyas consecuencias son ciertamente importantes, ya que abren el ámbito de la creación cartográfica a un nuevo y amplio grupo de personas.

Se conoce como *Información Geográfica Voluntaria o Participativa* (en inglés Volunteered Geographical Information, VGI) al uso de Internet para crear, gestionar y difundir información geográfica aportada voluntariamente por usuarios de la propia red. El conjunto de herramientas y técnicas que emplean esos usuarios para aportar su información conforma lo que se ha dado en llamar **neogeografía**. La comparación entre proyectos de creación de VGI y la bien conocida Wikipedia sirve perfectamente para ilustrar qué es lo que entendemos por VGI y neogeografía, ya que la VGI es el resultado de aplicar los conceptos de la Web 2.0 al ámbito de la información geográfica.

En el caso particular de esta última, la neogeografía ha supuesto un profundo cambio en algunas de las ideas básicas de la cartografía, modificando asimismo la concepción tradicional de la información geográfica, sus características o el papel que esta venía desempeñando en muchos ámbitos (o incluso dándole un papel en campos donde con anterioridad el uso de información geográfica era escaso). Algunas de las ideas principales sobre la neogeografía son las siguientes:

- **Popularización y democratización.** La producción cartográfica ha estado siempre en manos de gobiernos u organismos, y en muchas ocasiones fuertemente censurada debido a su elevado valor estratégico. Con la VGI, la creación de información geográfica se democratiza y se convierte en un proceso participativo libre y sin restricciones. Se invierte el esquema «hacia abajo» de producción y uso de información geográfica.
- Los ciudadanos se convierten en **sensores** y tienen mayor conciencia de su realidad geo-espacial.

- Se elimina parte del «misticismo» de la producción de información geográfica

la neogeografía es en la actualidad un fenómeno que no debe dejarse de lado, ya que los proyectos que aglutina se están convirtiendo paulatinamente en proveedores fundamentales de datos cuya calidad en muchos casos es excelente.

El proyecto de VGI de mayor relevancia es **OpenStreet-Map** (OSM), un «proyecto colaborativo para crear mapas libres y editables».

METADATOS

Con independencia de la forma en que se hayan obtenido, los datos pueden requerir otros datos adicionales para interpretarse. Por ejemplo, si tenemos las coordenadas de un punto, para interpretarlo correctamente necesitamos conocer, entre otras cosas, el sistema de coordenadas en que vienen expresadas esas coordenadas. El dato con el que trabajamos (las coordenadas), requiere unos datos adicionales (por ejemplo, el código EPSG del sistema de referencia empleado) para cobrar verdadero sentido.

Surge así el concepto de **metadatos**. Los metadatos son **datos acerca de los datos**, y su misión es **explicar** el significado de los datos. Es decir, ayudan a los usuarios de los datos a entender mejor el significado que estos tienen y la información que guardan. Los metadatos son un documento adicional que acompaña a los datos, y que permite una mejor gestión y una utilización más precisa de ellos.

Trabajando en el entorno de un SIG, los metadatos pueden **referirse a ambas componentes** (temática y espacial)

El concepto de metadato no es algo nuevo y exclusivo de los datos digitales, ya que un mapa impreso también contiene metadatos en cierta forma. Una leyenda o un texto en un margen del mapa con información sobre la fecha en que se ha creado son también metadatos. En el caso de los datos geográficos digitales, los metadatos no forman parte

del dato directamente sino que son independientes de este. Ello permitirá **realizar operaciones separadamente** con los metadatos, tales como búsquedas, que abren nuevas posibilidades y dan un gran valor a estos.

Dos de las funciones principales de los metadatos son garantizar el uso correcto y adecuado de los datos y facilitar su gestión, localización y consulta.

Los datos espaciales, como muchos otros datos, son creados habitualmente para un determinado objetivo, y este objetivo no ha de ser necesariamente evidente o contenerse como tal en los datos mismos. Cuando se emplean esos datos para un objetivo distinto a aquel para el que fueron diseñados, pueden surgir problemas debido a que se está realizando un proceso para el que los datos con los que se trabaja presentan carencias. Por ejemplo, consultar los metadatos puede **evitar el trabajo con datos desfasados o de precisión insuficiente**, ya que conociendo estos parámetros es posible juzgar si los datos son adecuados al fin que se persigue con ellos.

Los creadores de datos deben procurar acompañar estos de metadatos precisos y suficientes, y los usuarios deben consultar estos antes de utilizar dichos datos. De este modo, se puede garantizar que un dato no es empleado de forma errónea y que los resultados que se obtendrán tendrán validez.

Los metadatos son un elemento muy importante en relación con la **calidad de los datos espaciales**.

Respecto a la gestión de los datos, los metadatos facilitan la organización y la realización de tareas tales como la búsqueda de datos dentro de una colección de estos. Los metadatos constituyen un «resumen» de las características principales de los datos, y pueden ser empleados para labores de búsqueda y localización de datos de un tipo dado. Los metadatos **facilitan y agilizan la localización** de los datos cuando estos se buscan **por criterios geográficos**. Añadiendo a los metadatos elementos como la extensión del área cubierta por los datos, este tipo de búsquedas se efectúan de forma más ágil y efectiva. En este sentido, el uso de metadatos es fundamental para el establecimiento de

catalogos de datos, los cuales ya que responden a las peticiones del usuario del catálogo en función de la información que los metadatos contienen.

Contenido y creación de metadatos

La información que contienen los metadatos para una capa de datos espaciales depende de factores tales como el **modelo de representación** empleado, el **formato** en que se almacenan los datos (formato de archivo, base de datos, etc.), la **organización, entidad o individuo responsable** o el **elemento al que se asocian** (conjunto de capas, capa individual, entidad...).

En el apartado ?? veremos las formas estandarizadas que existen para los metadatos geográficos, las cuales define directamente su contenido.

Algunos de los elementos comunes que se incorporan a los metadatos geográficos son la información **de identificación** (para identificarlo de forma única y distinguirlo de otros), la información **sobre la calidad de los datos** (incluyendo su origen y el de los datos de los que deriva, apareciendo así el concepto de **linaje** de los datos), la información **sobre la componente no espacial** o la información **sobre la distribución** (acceso, licencias, etc.)

Los metadatos puede **crearse en el origen** mismo de los datos, recogiendo la información al mismo tiempo que se producen los datos en sí. Tambien pueden **extenderse posteriormente** por parte de distribuidores, organizadores o usuarios.

La creación de los metadatos se realiza en su mayor parte **de forma manual**, existiendo aplicaciones específicas para ello. Los metadatos que pueden extraerse de las capas de datos directamente, tales como la extensión cubierta por estos, pueden crearse de forma automática.

Los metadatos se almacenan en general como **ficheros adicionales** a los ficheros de datos, o bien como **parte de una base de datos**.

HERRAMIENTAS DE ESCRITORIO

El concepto clásico de un SIG es el de una aplicación completa en la cual se implementan herramientas para llevar a cabo las tareas básicas del trabajo con datos geográficos: creación o edición, manejo y análisis. Junto a este enfoque tradicional, con el tiempo han surgido otras formas distintas de aplicaciones que también pueden considerar como parte del ámbito del SIG.

Veremos en este capítulo las características de las aplicaciones SIG, divididas en tres grupos fundamentales: herramientas de escritorio, cartografía en la web (*web-mapping*) y SIG móvil.

HERRAMIENTAS DE ESCRITORIO

Podemos dividir las funciones básicas de un SIG de escritorio en cinco bloques: **entrada y salida de datos, visualización, edición, análisis y generación de cartografía**. Una aplicación de escritorio habitual presenta todas estas capacidades en cierta medida, aunque no necesariamente con el mismo nivel de implementación.

Entrada y salida de datos

Una aplicación SIG de escritorio debe implementar capacidades para leer datos y, opcionalmente, para guardarlos.

Esta última es necesaria en el caso en que el SIG pueda generar nuevos datos geográficos (nuevas capas), pero no en aquellas aplicaciones sin capacidades de análisis o edición, donde su empleo no ha de crear nuevos datos.

La existencia de librerías y componentes de acceso a datos en los que las aplicaciones SIG de escritorio pueden apoyarse permite dar soporte al gran número de formatos de datos existentes, mejorando la **conectividad** entre estas.

Además de acceder a ficheros de datos, en la actualidad es importante poder acceder a **bases de datos** o **servicios remotos**. Hablaremos de estos últimos más adelante dentro de este capítulo.

Visualización

La visualización es una función fundamental dentro de los SIG y del trabajo con cartografía en general. Tiene importancia cuando la representación de los datos es el propósito principal de utilizar un SIG, pero también cuando el trabajo está enfocado a la edición o la realización de análisis, ya que la visualización y exploración visual de los datos de partida es un paso previo.

Esta estructura se compone fundamentalmente de un lienzo sobre el que se sitúan las distintas capas de información geográfica, y que el usuario va conformando añadiendo nuevas capas y editando la forma en la que estas se representan. Las capas se sitúan en un orden dado dentro del lienzo, lo que permite establecer una jerarquía de representación y así lograr el aspecto deseado.

Junto a este lienzo existen herramientas de navegación que permiten ampliar o reducir la escala, o bien modificar el encuadre (Figura 6.1). Asimismo, se presentan todas las funciones que permiten obtener las distintas formas de representación que veremos en la parte ??, tales como el ajuste de colores o el establecimiento de etiquetas en función de los valores asociados a las distintas entidades, entre otros.

Estas capacidades convierte a los datos geográficos en un elemento activo, pues, a diferencia de un mapa clásico donde no pueden modificarse sus características, en un SIG

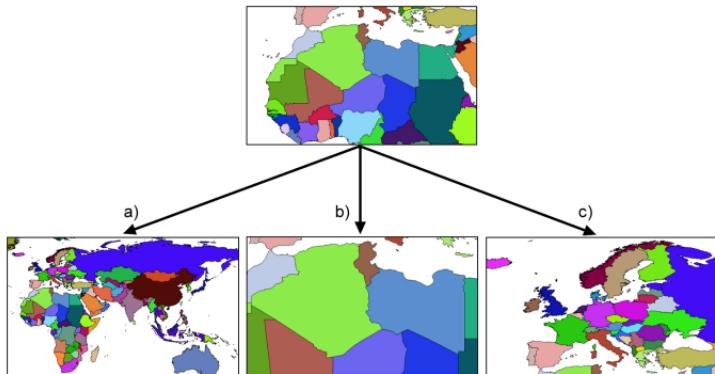


Figura 6.1: Herramientas de navegación fundamentales en el entorno gráfico de un SIG de escritorio. a) alejamiento (*zoom out*), b) acercamiento (*zoom in*), c) desplazamiento (*pan*)

el usuario puede de forma rápida y sencilla elegir *qué* ve y *cómo* lo ve.

Como parece lógico pensar, la visualización ha evolucionado mucho desde los primeros SIG, y ha ido progresivamente adquiriendo nuevas capacidades, muchas de las cuales solo son posibles con los modernos componentes gráficos de los ordenadores actuales. Así, además de ofrecer mayores posibilidades de personalización, el uso del SIG como herramienta de representación permite obtener resultados novedosos que añaden nuevas formas de explorar los datos geográficos y trabajar con ellos.

En el caso más habitual, la representación de una capa en un lienzo de un SIG es bidimensional, de la misma forma que se representa en un mapa impreso, lo cual se debe tanto a la mayor facilidad de implementación de este tipo de representaciones como a la mayor exigencia que otro tipo de representaciones presentan en lo referente al equipo (*hardware*).

No obstante, la presencia de visores tridimensionales está experimentando un gran crecimiento en los últimos años, y se van integrando paulatinamente dentro de los SIG para ofrecer una nueva forma de representación. Esta

clase de capacidades gráficas, inalcanzables en términos de rendimiento para un equipo común hace unos años, son hoy día perfectamente utilizables en un ordenador de consumo habitual, y aportan una nueva forma de trabajar con los datos geográficos.

Las funcionalidades de representación 3D aún no se integran completamente dentro del entorno de un SIG de escritorio completo, ya que se conciben por lo general como un elemento puramente visual y enfocado a la representación como tal, mientras que un lienzo 2D cumple su función tradicional de marco de trabajo sobre el que se desarrollan las restantes tareas del SIG como el análisis o la propia gestión de datos. Aun así, conforme las vistas tridimensionales se van convirtiendo en elementos habituales, los restantes componentes del SIG se van coordinando con ellas para darles a su vez la capacidad de servir como entornos de trabajo versátiles.

La presencia de una dimensión adicional hace que las herramientas de navegación sean más complejas en el caso tridimensional, existiendo ajustes relativos a la perspectiva, a los ángulos de visión o a la exageración del relieve, entre otros parámetros. Como puede verse más adelante en la figura 6.6, las herramientas de control de navegación de un visor 3D (situadas en este caso en la parte superior derecha de la figura) , son sensiblemente más complejas que las herramientas base del caso 2D, presentadas en la figura 6.1.

Esta mayor complejidad hace asimismo que puedan existir diversas formas en las que las capacidades de visualización 3D se presentan en un SIG. Las representaciones tridimensionales pueden ser simples representaciones en relieve de una capa (Figura 6.2a), o pueden incluir verdaderos elementos en tres dimensiones (Figura 6.2b). En el primer caso, la capa no contiene datos de elevación (puede ser por ejemplo una capa de usos de suelo), y la representación tridimensional se realiza basándose en la información de una capa adicional de elevaciones (que denominábamos de 2,5D por no poder representar todo tipo de formas en el espacio). La capa plana se representa en espacio «deformada» para ajustarse al relieve existente en la zona que representa.

En el caso de una representación tridimensional real, los objetos poseen información sobre su forma tridimensional, y junto con las coordenadas que delimitan su geometría plana (las geometrías básicas que conocemos: punto, línea y polígono), existen valores adicionales en el eje vertical. De este modo, pueden representarse entidades tales como edificios, o una ruta tridimensional que represente la trayectoria de un avión. Asimismo, y como puede verse en la figura, pueden añadirse elementos adicionales que aprovechan las capacidades de representación 3D, así como etiquetas o incluso elementos interactivos que también actúan en 3D (por ejemplo, para la selección de entidades).

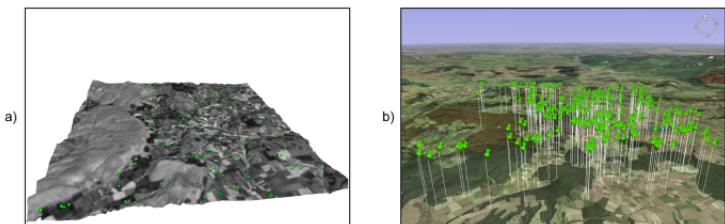


Figura 6.2: Distintas formas de representación 3D de capas de datos geográficos

Es importante recalcar que la visualización de una capa dentro de un SIG es independiente de la información que dicha capa contiene o la forma en que esta se almacena. El dato geográfico y su representación van por separado, y el dato en sí no define la representación, sino que únicamente sirve como apoyo para esta. Esto es particularmente cierto para el caso de capas vectoriales, así como para capas ráster que contengan un valor de tipo no gráfico, es decir, aquellas que no sean imágenes.

Una imagen tendrá el mismo aspecto en todos los SIG en los que se utilice, puesto que la información relativa a su representación se contiene, al menos en cierta medida, en la propia imagen. Los valores de la imagen representan una intensidad de luz y, si las bandas corresponden a la zona del visible, existe asimismo una información de colores. Otras imágenes pueden proceder del escaneo de un mapa o una

fotografía impresa, y en ese caso recogen también los colores de cada píxel.

Aun en el caso de imágenes con más de tres bandas, en las que no existe una correspondencia directa entre los valores y la representación (tal como vimos, por ejemplo, en la creación de representaciones en falso color en el apartado ??), puede existir una representación *por defecto* tomando, por ejemplo, las tres primeras bandas, y sigue siendo probable que estas imágenes también se vean de igual modo en uno u otro SIG (aunque luego ello no implica que dicha representación no pueda ajustarse a conveniencia en todos los casos).

En el caso de capas vectoriales o una capa ráster tal como un MDE, no existe ningún tipo de información acerca de la representación acompañando al dato espacial en sí. Los datos necesitan de un esquema de asignación que los convierta en elementos visuales (colores, texturas, etc.), pero este esquema es ajeno al dato en sí.

La labor del SIG relativa a la visualización consiste en «interpretar» los datos y convertirlos en representaciones, y para ello se basa en esquemas definidos por el usuario. Estos esquemas pueden ser almacenados de forma que en sucesivos usos de una capa de datos, esta se represente de una misma forma. No obstante, abrir la capa con una aplicación SIG distinta implicará en general perder las características definidas para su representación ya que, si bien los formatos de datos son relativamente interoperables, no así los formatos en que se almacenan los criterios de representación de esos datos.

Estándares para el almacenamiento de estilos como SLD(*Styled Layer Descriptor*), que veremos en detalle en el apartado ??, tienen como objeto solventar este problema.

Análisis

Si hubiéramos de ordenar cronológicamente las distintas funcionalidades que un SIG de escritorio presenta, probablemente el análisis fuera una de las primeras. Por encima de otras capacidades, los ordenadores han sido y son principalmente herramientas de cálculo capaces de realizar

operaciones y computar resultados, y este ha sido uno de los usos fundamentales relativos al manejo de datos geográficos. Otros usos, tales como la visualización, pese a ser prácticamente imprescindibles hoy en día, estaban muy limitados en los primeros SIG. No ocurría así en el ámbito del análisis ya que, aunque con capacidades como es lógico menores que las actuales, los ordenadores ofrecían una potencia de cálculo que los convertía en herramientas de análisis tan interesantes como en la actualidad.

La tendencia actual en los SIG es considerar las capacidades de análisis como herramientas modulares que se ejecutan sobre una plataforma base, la cual comprende las capacidades de visualización y entrada y salida de datos. Todas estas capacidades de análisis son independientes entre sí, aunque pueden coordinarse y emplearse en conjunto para alcanzar un resultado concreto. De otro modo, cada una de las formulaciones o algoritmos que vimos en capítulos anteriores aparece dentro del SIG como una herramienta individual que opera sobre una serie de capas y genera un resultado dado, tomando en muchas ocasiones esas capas de entre aquellas que se están representando en el SIG, e incorporando asimismo a dicha representación las nuevas capas generadas.

Las herramientas de análisis pueden aparecer igualmente como programas independientes, y el SIG de escritorio ser una herramienta aglutinadora que centraliza estas, facilitando su uso y la gestión de los datos implicados en los procesos de análisis.

Cuando las herramientas de análisis utilizan directamente la base del SIG donde se encuentran las capacidades de visualización y manejo de datos, puede existir cierto grado de interactividad. Las operaciones de consulta, por ejemplo, las cuales vimos en el capítulo 7, son en general de este tipo, ya que el usuario puede actuar sobre el lienzo para hacer una selección de modo gráfico. Más aún, esa selección puede condicionar los posteriores análisis sobre la capa cuyas entidades se seleccionan, ya que los procesos que operen sobre ella pueden restringir su alcance a aquellas entidades seleccionadas.

En caso de no existir este tipo de interacción entre elementos de análisis y elementos de visualización y exploración de datos, los procesos de análisis suelen constituir utilidades autocontenidoas que simplemente toman una serie de datos de entrada, realizan un proceso en el que el usuario no interviene, y finalmente generan un resultado con carácter definitivo. Este resultado podrá ser posteriormente visualizado o utilizado como entrada para un nuevo análisis.

A modo de ejemplo, podemos analizar el caso particular del cálculo de una ruta que conecte una serie de puntos a través de una red (los fundamentos de este análisis los vimos en el apartado??).

Una de las formas de implementar este análisis es aquella que requiere del usuario la introducción de la información necesaria (una capa de líneas con la red viaria y otra de puntos con los puntos de inicio, paso, y llegada correspondientemente ordenados) como parámetros que el proceso toma y en función a los cuales se genera una nueva capa. El proceso es una tarea perfectamente definida, con unas entradas y una salidas, y tras la selección de unas capas de entrada se genera un nuevo resultado en forma de otra capa. Este proceso puede implementarse de forma aislada del SIG, aunque coordinada con él para lograr mejores resultados y una utilización más sencilla.

Otra forma con un enfoque distinto sería presentando un proceso interactivo en el cual se introduce como único parámetro inicial la red viaria. Posteriormente, el usuario puede operar sobre el lienzo en el que esta se encuentra representada para ir definiendo y editando la lista de puntos de paso. El cálculo se va efectuando cuando se produce algún cambio en la misma y debe aplicarse de nuevo el algoritmo pertinente para adaptar el resultado —la ruta óptima— a ese cambio.

Este tipo de formulaciones interactivas son más intuitivas y agradables de usar, pero en realidad menos productivas de cara a un trabajo dentro de un proyecto SIG. Aparecen por ello en aquellas funciones que tienen una mayor componente visual o, especialmente, en las que representan un análisis puntual que se realiza de forma común como algo individual.

Estos son los análisis que se implementan en las aplicaciones que veremos más adelante como parte del grupo de SIG enfocados a la exploración visual de datos geográficos, que además de esta proveen alguna serie de utilidades de análisis, pocas en general, sin que estas estén concebidas para un trabajo completo en un proyecto SIG de cualquier índole, sino más bien para un uso ocasional.

En un proyecto SIG de cierto tamaño, lo más común es que la fase de análisis, a la que seguirá una fase también compleja de preparación e interpretación de resultados, y previa a la cual se ha debido llevar a cabo la preparación de los datos de partida, comprenda no uno, sino muchos análisis distintos. Estos análisis, a su vez, no son independientes, sino que están relacionados entre sí y lo más habitual es que definan como tales un flujo de trabajo que comienza en los datos de partida y desemboca en los resultados finales a través de una serie de procesos.

Por su naturaleza, tanto los datos espaciales como los procesos en los que estos intervienen se prestan a formar parte de estos flujos de trabajo más o menos complejos, y es por ello que en los SIG actuales una funcionalidad básica es la creación de tareas complejas que permiten simplificar todo un proceso de muchas etapas en una única que las engloba a todas. La forma anteriormente comentada en que aparecen las formulaciones dentro de un SIG, de forma atomizada y modular, facilita la creación de estos «modelos» a partir de procesos simples.

Para entender esta idea, podemos ver un ejemplo aplicado. La extracción de una red de drenaje en formato vectorial a partir de un MDE requiere una serie de procesos, a saber (para repasar los fundamentos de cada uno de estos procesos y comprender mejor la operación global, puede consultarse el capítulo ??, donde se describieron en su momento, así como la sección ??):

1. Eliminación de depresiones del MDE.
2. Cálculo de una capa de área acumulada a partir del MDE corregido

3. Extracción de capa ráster con la red de drenaje a partir de la capa anterior y un valor umbral
4. Vectorización de la capa resultante del paso anterior

Lo anterior podría simplificarse si se agruparan en una sola operación todos los procesos anteriores, de forma que tomando dos datos de entrada (un MDE y un umbral), se realizara todo el proceso de forma continua. Las capacidades de creación de modelos que implementan los SIG de escritorio (en particular, los más enfocados al análisis) permiten crear nuevos procesos compuestos utilizando entornos gráficos intuitivos, y estos procesos pasan a formar parte del conjunto de ellos de que dispone el SIG, permitiendo que cada usuario «resuma» en procesos unitarios una serie de operaciones que, de otro modo, deberían realizarse de forma individual con el coste operacional que ello implica.

Una vez que se ha creado un proceso compuesto, este puede aplicarse sobre nuevos datos de entrada, reduciéndose así el tiempo y la complejidad con que nuevos parámetros de entrada pueden ser procesados según el esquema de trabajo definido en dicho proceso.

Debe pensarse que el proceso presentado como ejemplo es muy sencillo y únicamente implica cuatro operaciones encadenadas de forma lineal. Un proceso de análisis habitual puede contener muchas más operaciones individuales, y estas disponerse de forma más compleja, con dependencia de distinta índole entre sus resultados.

La imagen 6.3 muestra el aspecto de uno de tales procesos implementado en QGIS, un software GIS con capacidad de creación de modelos. El proceso representado en la figura ?? también es un ejemplo de otro tipo de análisis que puede adaptarse ventajosamente en este tipo de herramientas de modelización.

Asimismo, las herramientas SIG que contienen funcionalidad de análisis suelen permitir el acceso a estas a través de lenguajes de *scripting*, lo cual facilita la creación de flujos de trabajo y la automatización de rutinas complejas de análisis. Si bien este trabajo no se realiza de un modo gráfico

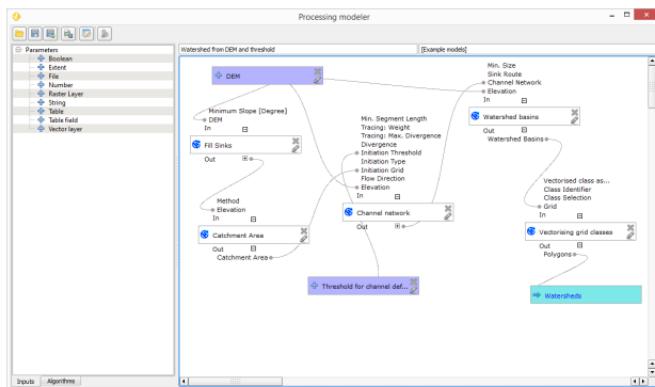


Figura 6.3: Esquema de un proceso complejo creado a partir de operaciones simples de análisis con datos SIG.

e intuitivo como en el ejemplo mostrado anteriormente, y requiere mayores conocimientos, la flexibilidad y potencia que ofrece es mucho mayor.

Edición

Los datos geográficos con los que trabajamos en un SIG no son una realidad estática. La información contenida en una capa es susceptible de ser modificada o corregida, y las funciones que permiten estas tareas son importantes para dotar al SIG de versatilidad. Sin ellas, los datos espaciales pierden gran parte de su utilidad dentro de un SIG, ya que se limitan las posibilidades de trabajo sobre estos. Las funcionalidades de edición son, por tanto, básicas en una herramienta de escritorio.

Las operaciones de edición pueden emplearse, por ejemplo, para la actualización de cartografía. Como vimos en el capítulo 5, una de las ventajas de los datos digitales frente a los analógicos es la mayor facilidad de actualización. Así, si una entidad en una capa vectorial (por ejemplo, una parcela catastral) modifica su geometría, no es necesario rehacer todo un mapa, sino simplemente editar ese elemento. A lo largo del desarrollo de un proyecto SIG, es muy probable que

sea necesario editar de un modo u otro algún dato espacial, bien sea para corregirlo, ampliarlo, mejorarlo o sencillamente adaptarlo a las necesidades del propio proyecto.

Además de la modificación de una capa ya existente, las herramientas de edición de un SIG de escritorio se emplean igualmente para la creación de capas nuevas, que pueden crearse a partir de la digitalización de imágenes como vimos en el capítulo 5, o bien en base a cualquier otra capa de la que dispongamos.

Aunque las tareas de edición más habituales son las relacionadas con la edición de geometrías, no es esta la única edición que puede realizarse dentro de un SIG. Podemos distinguir las siguientes formas de edición:

- Edición de geometrías de una capa vectorial
- Edición de atributos de una capa vectorial
- Edición de valores de una capa ráster

Las herramientas destinadas a la edición de entidades geométricas heredan sus características de los programas de diseño asistido por ordenador (CAD), cuya funcionalidad principal es precisamente la edición de elementos gráficos. Estas incluyen la adición o eliminación de nuevas geometrías, la modificación de ellas editando sus puntos (recordemos que toda entidad vectorial se reduce a un conjunto de puntos en última instancia), así como otras operaciones geométricas básicas. En la sección ?? vimos algunas de ellas a la hora de tratar la calidad de la digitalización en pantalla.

Otras funciones de edición que encontramos son las que permiten simplificar algunas tareas, tales como la división de un polígono. La figura 6.4 muestra cómo un polígono puede dividirse en dos simplemente trazando una línea divisoria. Otras funcionalidades similares incluyen la eliminación automática de polígonos espúreos (véase ??), o el ajuste automático entre entidades.

En general, el número de funcionalidades es sensiblemente menor que en el caso de los programas CAD, ya que gran parte de ellas no tiene aplicación directa en el caso de trabajar con datos geográficos. No obstante, también

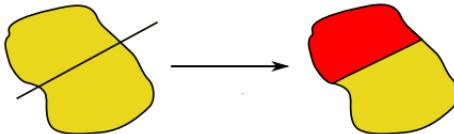


Figura 6.4: División automática de un polígono en dos nuevas entidades a partir de una línea. Funcionalidades de este tipo aparecen en los SIG para facilitar las tareas de edición.

aparecen herramientas adicionales, como sucede en el caso de que se registre información topológica, lo cual ha de considerarse por igual en el proceso de edición. Así, herramientas como la mostrada en la figura 6.4 han de tener en cuenta la existencia de una estructura topológica y un modelo de representación particular en el SIG, y no operarán igual en todas las aplicaciones, ya que, como sabemos, no todas presentan las mismas capacidades en este terreno.

Junto con las propias geometrías que pueden editarse según lo anterior, toda capa vectorial tiene asociado un conjunto de atributos, y estos deben poder editarse también desde el SIG. De hecho, la adición de una nueva geometría a una capa vectorial no está completa hasta que no se añaden igualmente sus atributos.

La edición de toda la información alfanumérica relacionada con las distintas entidades se realiza en un SIG a través de elementos tomados del ámbito de las bases de datos, siendo esto en general válido tanto en lo referente a las interfaces como en el propio acceso a datos. Las operaciones de edición de atributos abarcan tanto la modificación de valores sencillos como la de la propia estructura del conjunto de atributos (adición o eliminación de columnas —campos— en la tabla correspondiente). Por supuesto, la edición de atributos y de geometrías esta íntimamente relacionada.

Por último, la edición de capas ráster es mucho menos frecuente, y una gran mayoría de SIG no permiten la modificación directa de los valores de las celdas. Las operaciones del álgebra de mapas permiten modificar los valores de una capa y obtener nuevas capas con esos valores modificados,

pero editar directamente un valor de celda del mismo modo que se editan el valor de un atributo o la posición de un punto de una capa vectorial no es una funcionalidad tan habitual.

Este tipo de capacidades, no obstante, pueden ser de gran utilidad, especialmente en SIG orientados al manejo principal de capas ráster, donde sustituyen en cierta medida a las funcionalidades de edición vectorial equivalentes.

Generación de cartografía

A pesar de que la representación de las distintas capas de datos espaciales en un lienzo es suficientemente potente a efectos de explorar visualmente la información que estas contienen, la mayoría de los SIG incorporan capacidades de creación de cartografía impresa, generando un documento cartográfico que posteriormente puede imprimirse y emplearse como un mapa clásico. Las razones para la existencia de tales funcionalidades son muchas, pero la principal sigue siendo la necesidad general que aún existe de apoyarse en esa clase de documentos cartográficos para poder incorporarlos a proyectos o estudios como parte de anexos cartográficos.

Aunque la representación dentro de un SIG ofrece posibilidades mayores (cambio de escala, ajuste de los parámetros de visualización, etc.), disponer de una copia «estática» de cada bloque de información con el que se trabaja en un SIG es una necesidad ineludible. Más que una capacidad necesaria para la presentación adecuada de la información cartográfica, la generación de cartografía impresa es en muchos casos la principal razón para el uso de un SIG. Mientras que las capacidades de análisis o edición son en ocasiones poco o nada utilizadas, las de generación de cartografía son un elemento fundamental, y muchos usuarios la consideran erróneamente como la funcionalidad primordial de un SIG.

Llegado a este punto del libro, y tras todo lo que hemos visto, queda claro, no obstante, que un SIG es ante todo una herramienta de gestión y análisis de datos espaciales, y que las capacidades enfocadas a la producción cartográfica deben verse como una ayuda —de vital importancia, eso sí—

para la presentación final de todo el trabajo que se lleva a cabo en él.

Fundamentalmente, estas capacidades permiten la composición de documentos cartográficos de acuerdo con un diseño dado, y la impresión directa de estas en algún periférico tal como una impresora común o un *plotter* de gran formato. En la elaboración de dicho diseño, pueden emplearse todos los elementos que habitualmente podemos encontrar en un mapa: el propio mapa en sí (la representación de la información geográfica), leyenda, título, escala, etc. Con estos elementos, se crea una versión autocontenido de la información geográfica, que puede ya emplearse de modo independiente del SIG.

Las funciones de diseño que se implementan por regla general en un SIG son similares a las que pueden encontrarse en un software de maquetación genérico, permitiendo la composición gráfica del documento general y el ajuste de los distintos elementos que lo forman. Funciones más avanzadas no están presentes de modo habitual, principalmente debido a las características muy concretas y bien definidas del tipo de documento con el que se trabaja, lo cual hace posible esta simplificación.

Aunque, como se ha dicho, muchos usuarios, bien por desconocimiento o falta de formación en la herramienta, consideran que la capacidad principal de un SIG es «hacer mapas», lo cierto es que, incluso en aquellas aplicaciones más completas y avanzadas, la potencia en la producción cartográfica es muchas veces insuficiente para producir cartografía profesional. Pueden obtenerse resultados de gran calidad y sin duda de suma utilidad, pero lo más habitual es no encontrar en un SIG las capacidades que se necesitan, no ya únicamente desde el punto de vista del cartógrafo, sino desde la perspectiva del diseño.

La creación de un mapa no es solo una tarea técnica, sino asimismo una labor artística, existiendo unas necesidades en función del enfoque que prime. Como herramienta de trabajo, un SIG es un elemento técnico, y las consideraciones artísticas, aunque pueden en cierta forma aplicarse con las herramientas que este implementa, resultan más sencillas

de tratar si se dispone de aplicaciones más específicas en ese sentido.

Por ello, lograr un mapa de apariencia realmente profesional requiere unas herramientas de diseño avanzado, a la par que un conjunto de utilidades suficiente como para poder aplicar a la creación del mapa todos los conceptos sobre representación que veremos en la parte ??, no encontrándose estas en ocasiones en su totalidad dentro de un SIG. La utilización del SIG como aplicación base y el uso posterior de programas de diseño es la solución adecuada para la obtención de cartografía profesional, aunque lógicamente requiere unos mayores conocimientos y una especialización más allá de la propia práctica cartográfica.

No obstante, para el usuario técnico de SIG (el usuario al que está dirigido este libro), las herramientas de diseño cartográfico que la mayoría de aplicaciones implementan son más que suficientes, y permiten lograr resultados altamente satisfactorios.

Una de las funciones más interesantes de generación cartográfica en un SIG es la automatización del proceso y la simplificación de la producción de grandes volúmenes de cartografía. Por una parte, todas las herramientas de escritorio capaces de producir cartografía son a su vez capaces de «reutilizar» diseños, de tal modo que si un conjunto de mapas tienen unas características comunes (por ejemplo, una misma disposición de sus elementos), no es necesario elaborar todos ellos desde cero.

Esto permite, por ejemplo, crear una serie de mapas de una misma zona conteniendo cada uno de ellos información sobre distintas variables. A partir de un conjunto de capas, se elabora el diseño de un mapa y este se alimenta de dichas capas, creando mapas independientes que reflejan estas por separado o en distintas combinaciones. Esto simplifica notablemente el proceso, ya que el diseño ha de realizarse una única vez, al tiempo que se garantiza la uniformidad de los distintos resultados.

Otra aplicación en esta línea es la generación de una serie de mapas que cubren en su conjunto una amplia extensión, fragmentando esta en unidades. La gestión de los encuadres

para cada una de esas unidades, o la creación de un mapa guía en cada caso que localice la hoja concreta dentro de la extensión global del conjunto, ambas pueden automatizarse junto con las restantes operaciones de diseño. De este modo, la producción de toda una serie cartográfica se simplifica en gran medida, siendo el SIG una herramienta que supone un gran avance en términos de productividad en este tipo de tareas.

La figura 6.5 muestra un ejemplo de lo anterior.

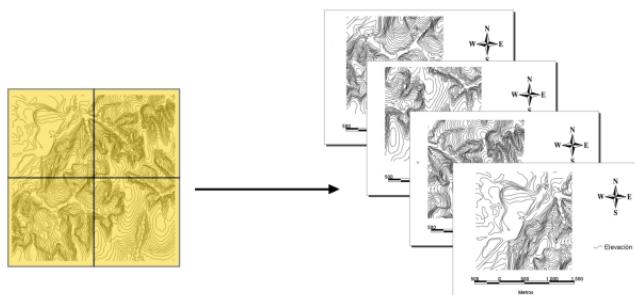


Figura 6.5: La automatización de las tareas de creación cartográfica permite simplificar la producción de grandes volúmenes de cartografía, como por ejemplo al dividir un área geográfica en una serie dada de mapas.

Estas posibilidades surgen de la separación existente en un SIG entre los datos espaciales y el diseño del documento cartográfico que los contiene, del mismo modo que ya vimos existe entre datos y parámetros de representación a la hora de visualizar los primeros.

TIPOS DE HERRAMIENTAS DE ESCRITORIO

No todas las aplicaciones de escritorio presentan las anteriores funcionalidades de igual modo. Es frecuente que ciertos Sistemas de Información Geográfica tengan una fuerte componente de análisis, pero que otras de las funciones principales, como por ejemplo la edición, no se presenten tan desarrolladas.

Un caso particular es el de aquellos SIG de escritorio que centran la gran mayoría de sus capacidades en el terreno de la visualización, permitiendo un uso de los datos geográficos similar al que corresponde a un mapa clásico, donde el trabajo con este se basa fundamentalmente en el análisis visual.

Comenzaremos por estos últimos para dar un breve repaso a los principales tipos de aplicaciones de escritorio en función de sus capacidades.

Visores y exploradores

Las aplicaciones SIG de escritorio cuya función principal es la visualización se conocen generalmente como *visores* o *exploradores*, y en la actualidad representan una fracción importante del conjunto total de herramientas SIG de escritorio.

En ocasiones, se trata de aplicaciones en el sentido habitual, las cuales presentan capacidades reducidas de análisis y edición, y cuyo objetivo no es otro que el permitir la visualización de cartografía, sin incorporar las restantes posibilidades del SIG. En otros casos, son versiones simplificadas de soluciones SIG de escritorio más complejas, desarrolladas como alternativas más asequibles (en términos de dificultad de manejo y aprendizaje, y también a veces en términos de coste).

Una forma también habitual en la que se presentan los exploradores son como herramientas de apoyo a unos datos espaciales particulares. Para entender este tipo de enfoque, debe pensarse que un mapa clásico puede visualizarse de igual modo con independencia del uso que se le pretenda dar y de la experiencia y formación de quien lo usa. Con un conjunto de datos espaciales en forma de una o varias capas, no sucede lo mismo, ya que estos datos no son un elemento «visual» de por sí. Es necesario utilizar un SIG para poder visualizarlos.

Un usuario experimentado no encontrará problemas en manejar un SIG de escritorio complejo, pero un usuario con poca experiencia que lo único que desee sea «ver» la car-

tografía y explorarla visualmente (del mismo modo que un excursionista casual puede querer emplear un mapa topográfico) encontrará el entorno de ese SIG demasiado complejo y con elementos que, en su mayoría, no le son necesarios. Con la disponibilidad creciente de cartografía y la popularización de las tecnologías SIG, este tipo de usuarios ha crecido notablemente, y las aplicaciones adaptadas a sus necesidades han ido apareciendo y popularizándose igualmente de forma progresiva.

Así, existen visores que ocupan un papel secundario como parte de un producto compuesto que incluye al propio programa y a los datos en sí. Ejemplos muy claros y muy populares son aplicaciones como Google Earth (figura 6.6), que permite que cada usuario incorpore su propia información para visualizar esta, pero cuyo mayor interés es el acceso a una enorme base de datos de imágenes de satélite con cobertura global. De esta forma, la aplicación puede utilizarse para explorar una área deseada, sin necesidad de disponer junto con ella de datos para dicha zona, puesto que por defecto la aplicación accede a una base de datos de imágenes que van indisolublemente unidos a ella.

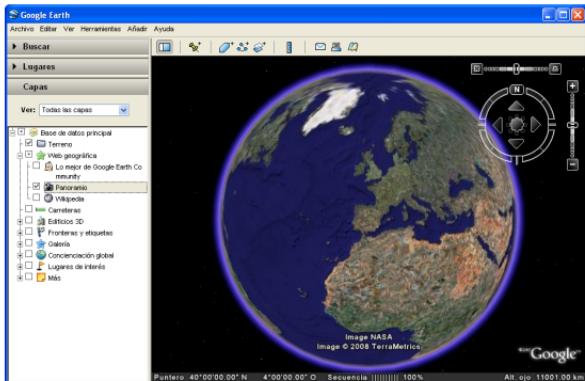


Figura 6.6: Aspecto de un globo o visor tridimensional (GoogleEarth).

El usuario puede también añadir sus propias capas y usar estos visualizadores de la misma manera que emplea

las capacidades de visualización de un SIG de escritorio más rico en funcionalidades, pero parte del interés de la aplicación no está solo en sus capacidades, sino en los datos que tiene vinculados y que permiten emplear estas.

En otros casos, un organismo o empresa puede generar un conjunto de capas bien a partir de algún tipo de análisis o de cualquier otra metodología, y opta por distribuir estas acompañadas de un visor que permita un primer acceso a los datos. Todas esas capas podrán ser empleadas dentro de un SIG que soporte los formatos de archivo en el que estas se hayan almacenado, pero aquellos usuarios que no dispongan de un SIG podrán igualmente efectuar consultas básicas y explorar la cartografía haciendo uso del visor incorporado.

En líneas generales podemos enumerar las siguientes características de los visores:

- Interfaz simple en la que tienen un peso mayoritario las herramientas de navegación.
- Capacidades de lectura de datos, pero no de escritura.
- Reducidas o nulas capacidades de edición y análisis.
- Enfocadas a usuarios no especializados.

Encontramos dos grupos básicos de visores, en función de qué tipo de visualización principal incorporan: planos y tridimensionales (globos). Mientras que una aplicación SIG de escritorio completa puede presentar los dos tipos de representación, y sobre ambas implementar las restantes funcionalidades tales como la edición o el análisis (aunque, como ya dijimos, en mayor proporción sobre las vistas bidimensionales), los visores habitualmente reducen sus capacidades de representación a una de estas variantes.

Los visores bidimensionales, aunque sin alcanzar el enfoque especializado de una aplicación completa, se orientan más al usuario con cierto conocimiento del ámbito SIG, mientras que la tendencia en los tridimensionales es a ofrecer herramientas de acceso a datos geográficos con una apariencia atractiva. Ello, no obstante, no implica que estos visores carezcan de utilidad en el ámbito científico, siendo igualmente herramientas válidas para todo tipo de investigación o

trabajo que incorpore cierta componente geográfica. De hecho, la popularización de este tipo de visores ha supuesto un gran acercamiento de los datos geográficos y las capacidades SIG a toda una amplia comunidad de usuarios, incluyendo los del mundo científico, permitiéndoles realizar sus trabajos de forma más adecuada. Esto es especialmente cierto con aquellos visores que se hallan vinculados a bases de datos particulares, como ya se ha comentado, ya que permiten explotar los datos de dichas bases y proporcionar a todo usuario un sustrato de información geográfica sobre la que trabajar.

Soluciones de escritorio completas

La aplicación SIG más habitual, y la que constituye la herramienta básica para el desarrollo de un proyecto SIG, es aquella que reúne en un único producto todas las funciones básicas que hemos visto en este capítulo. Con las lógicas diferencias en cuanto al grado de funcionalidad de estas según el enfoque de la aplicación, una aplicación SIG de escritorio completa debe permitir la lectura de los datos, la creación y modificación de estos con sus capacidades de edición, su visualización, la realización de análisis con ellos y la generación de resultados cartográficos ya sea a partir de los datos originales o de datos generados en los procesos de análisis. Con todas estas capacidades, una herramienta SIG de escritorio constituye una solución completa para todo tipo de proyectos SIG, y puede dar respuesta a todas las necesidades que en ellos se presentan.

Como ya vimos en el capítulo introductorio de esta parte, y también en el capítulo 7 dedicado a las bases de datos, la forma de abordar la implementación de las distintas capacidades ha ido variando a medida que se iban desarrollando los SIG. En la actualidad, encontramos tanto SIG de escritorio que implementan en una sola aplicación central todas las funcionalidades base, como grupos de aplicaciones muy interrelacionadas que implementan por separado cada una de dichas funcionalidades. Tanto en uno como en otro caso, existen elementos sobre los que las distintas herramientas

del SIG se apoyan, especialmente en lo relativo al acceso a datos.

Pese a que incorporan toda la gama de funcionalidades base, las soluciones de escritorio completas no cubren las necesidades de todo usuario de SIG. La convergencia a la que tienen este tipo de aplicaciones, comentada en el capítulo introductorio de esta parte del libro, no ha sido, como ya entonces se mencionó, completamente lograda en la actualidad. El problema no es ya un problema de integración de tecnologías, sino una dificultad relativa a la gran amplitud de la ciencia de los SIG. Resulta imposible reunir en una sola herramienta todas las capacidades que un SIG de escritorio puede incluir, y es por ello que todas las soluciones de escritorio presentan algún tipo de especialización, dando prioridad a algún área respecto a las restantes.

Una división que perdura, aunque en mucho menor medida que en los primeros SIG, es la existente entre SIG ráster y SIG vectoriales, especialmente en lo relativo al análisis. Un usuario que trabaje con datos espaciales como los correspondientes, por ejemplo, a información catastral, utilizará para su labor un SIG de escritorio distinto al que usara alguien cuyo trabajo implique mayoritariamente la realización de análisis del terreno o la creación de modelos geográficos, a pesar de que ambas soluciones probablemente incorporen capacidades tanto en el ámbito ráster como en el vectorial. El alcance de estas, no obstante, será distinto en cada caso.

Incluso en caso de presentar una orientación principalmente hacia los datos de tipo ráster, existen también enfoques distintos dependiendo principalmente del tipo información que se vaya a manejar. Una división fundamental es la existente entre aquellas aplicaciones destinadas al manejo de imágenes y aquellas cuyo elemento principal de trabajo son las capas ráster con otro tipo de valores, tales como MDE o capas similares.

Las imágenes, especialmente si se trata de imágenes de satélite, van a constar de una serie de bandas cuyo número puede ser muy elevado, lo que requiere unas herramientas particulares para su manejo. Asimismo, gran parte de las funciones que vimos en el capítulo ?? tales como las relativas

a la corrección de imágenes, no tienen aplicación para otro tipo de capas (no tiene sentido aplicar una corrección geométrica a un MDE o una capa de pendientes, ya que estos datos necesariamente provienen de fuentes ya conveniente corregidas), por lo que en cierto modo pueden considerarse específicas de este campo, el de las imágenes, si bien es cierto que se trata de un campo de gran amplitud.

Por el contrario, capas como el propio MDE u otras similares solo contienen una única banda, y el tipo de operaciones que se desarrollan sobre ellos es bien distinto. Con objeto de simplificar estas operaciones, la estructura de estas aplicaciones ha de enfocarse hacia alguna de estas variantes, dándole prioridad sobre la otra. Por ello, las herramientas de escritorio que se orientan al trabajo con imágenes incorporan en general pocas o nulas herramientas en áreas como el análisis del terreno, mientras que aquellas que sí tratan estos análisis no incluyen salvo las funciones más simples para el manejo de imágenes (realces, ajuste de contraste, etc.), pero no las más específicas.

En realidad, una aplicación de escritorio global que cubriera todas estas funcionalidades no sería práctica desde el punto de vista de su uso, pues sería excesivamente compleja. Es poco probable, igualmente, que un mismo usuario requiera un entorno profesional productivo en todas ellas, y más habitual sin embargo que centre su trabajo en un área concreta.

RESUMEN

Las herramientas de escritorio son la forma más clásica de los SIG. Entendemos como tales a aquellas herramientas ciertamente complejas que permiten llevar a cabo las tareas básicas de un SIG en sentido tradicional, como son el manejo de datos espaciales y el trabajo con los mismos.

Podemos distinguir cuatro funcionalidades básicas que aparecen representadas en mayor o menor medida en un SIG de escritorio: visualización, edición, análisis y generación de cartografía.

En función del grado de desarrollo e implementación en que las anteriores funcionalidades se encuentren en un SIG de escritorio, distinguimos distintas formas de estas herramientas. La división más genérica es aquella que distingue las herramientas pensadas para un trabajo completo en todas las distintas fases de un proyecto SIG de aquellas orientadas a la representación y exploración visual de los datos geográficos. Estas últimas representan un enfoque más reciente, y en la actualidad están contribuyendo de manera muy notable a la expansión de las tecnologías SIG fuera del ámbito más especializado.

BASES DE DATOS

Las bases de datos son un elemento fundamental en el entorno informático hoy en día y tienen aplicación en la práctica totalidad de campos. Concebidas con un propósito general, son de utilidad para toda disciplina o área de aplicación en la que exista una necesidad de gestionar datos, tanto más cuanto más voluminosos sean estos. En el ámbito particular de los SIG, los datos son cada día más voluminosos, debido no solo a una mayor cantidad de información, sino también a una mayor precisión en esta, la cual implica un mayor volumen de datos. Además, presentan otra serie de características (uso múltiple, necesidad de acceso eficiente para análisis, necesidad de indexación, etc.), haciendo todas ellas que sea recomendable el uso de bases de datos y tecnologías específicas para su manejo.

Entendemos como *Base de Datos* un **conjunto de datos estructurado y almacenado de forma sistemática** con objeto de facilitar su posterior utilización. Los elementos clave de la base de datos son esa estructuración y sistematicidad, pues ambas son las responsables de las características que hacen de la base de datos un enfoque superior a la hora de gestionar datos.

Las ventajas de utilizar una base de datos frente a una gestión no organizada de estos las encontramos tanto en los propios datos como en el uso que se hace de ellos. Algunas ventajas que afectan directamente a los datos son las

siguientes:

- **Mayor independencia.** Los datos son independientes de las aplicaciones que los usan, así como de los usuarios.
- **Mayor disponibilidad.** Se facilita el acceso a los datos desde contextos, aplicaciones y medios distintos, haciéndolos útiles para un mayor número de usuarios.
- **Mayor seguridad (protección de los datos).** Mayor facilidad de replicación de datos y mejor sincronización.
- **Menor redundancia.** Con el consiguiente menor volumen de datos y mayor rapidez de acceso.
- **Mayor eficiencia en la captura, codificación y entrada de datos.**

Esto tiene una consecuencia directa sobre los resultados que se obtienen de la explotación de la base de datos, presentándose al respecto ventajas como, por ejemplo:

- **Mayor coherencia.** La mayor calidad de los datos que se deriva de su mejor gestión deriva en mayor calidad de los resultados.
- **Mayor eficiencia.** Facilitando el acceso a los datos y haciendo más sencilla su explotación, la obtención de resultados es más eficiente.
- **Mayor valor informativo.** Resulta más sencillo extraer la información que los datos contienen, ya que uno de los cometidos de la base de datos es aumentar el valor de estos como fuente de información.

Por último, los usuarios de la base de datos también obtienen ventajas al trabajar con estas, entre los que cabe citar:

- **Mayor facilidad y sencillez de acceso.** El usuario de la base de datos se debe preocupar únicamente de *usar* los datos, disponiendo para ello de las herramientas adecuadas y de una estructura sólida sobre la que apoyarse.
- **Facilidad para reutilización de datos.** Esto es, facilidad para compartir.

De forma resumida, puede decirse que la principal bondad de una base de datos es la **centralización** que supone de todos los datos con los que se trabaja en un contexto determinado, con las consecuencias que ello tiene para una mejor gestión, acceso o estructuración de estos.

BASES DE DATOS RELACIONALES

De entre los distintos modelos existentes para plantear una base de datos, el más habitual, tanto dentro como fuera del ámbito de los SIG, es de las denominadas **bases de datos relacionales**. Este modelo utiliza un esquema basado en **tablas**, que resulta a la vez sencillo de comprender y fácil de utilizar para el análisis y la consulta de los datos. Las tablas contienen un número dado de **registros** (equivalentes a las filas en la tabla), así como **campos** (columnas),

La tabla en sí se conoce como **relación**, ya que recoge la relación existente entre sus elementos, y constituye así el eje central del modelo relacional. Las columnas representan los distintos **atributos** asociados a la entidad, mientras que las filas conforman los distintos **registros**. Una fila se forma con un conjunto de n atributos, constituyendo una **tupla**.

Una base de datos contiene normalmente más de una tabla, ya que suelen ser muchos los tipos de datos a almacenar y resulta conveniente dividirlos en distintas tablas. Además de las relaciones que la tabla en sí implica, es necesario definir interrelaciones entre las distintas tablas, y para ello se emplean los denominados atributos **clave**. Un atributo clave es aquel que tiene valor **único e invariable** para cada tupla, pudiendo servir para representar a esta plenamente. Por ejemplo, en una tabla con nombres de personas e información adicional sobre ellas, el número de su DNI puede servir como atributo clave.

Cuando trabajamos con datos espaciales, es habitual emplear la **componente espacial como clave**, ya que esta suele ser única.

Las interrelaciones entre tablas pueden ser de distintos tipos, según el número de entidades de una tabla con los que se relacionen las entidades de la otra. Tenemos así relaciones

de **uno a muchos**, de **uno a uno** o de **muchos a muchos**. Por ejemplo, en una tabla con entidades que representan personas y otra con entidades que representan ciudades, si establecemos una interrelación *vive en*, se tratará de una interrelación de uno a muchos, ya que en una ciudad pueden habitar varias personas.

SISTEMAS GESTORES DE BASES DE DATOS

Junto con las bases de datos, el elemento fundamental para el aprovechamiento de estas son los **Sistemas Gestores de Bases de Datos** (SGDB o DBMS, del inglés *DataBase Management System*). Estos sistemas representan un **elemento intermedio entre los propios datos y los programas** que van a hacer uso de ellos, facilitando las operaciones a realizar sobre aquellos. Los programas tales como un SIG no acceden directamente a la base de datos, sino que lo hacen **a través de un SGBD**.

Algunas características que ha de tener un SGBD son las siguientes:

- **Acceso transparente a los datos.** El SGBD debe crear una abstracción de los datos que haga el trabajo con estos más sencillo, ocultando aspectos internos que no sean relevantes para dicho trabajo.
Procedimientos como las **consultas**, que veremos en la sección 7, se realizan a través del SGBD, que es quien se encarga de interpretar dichas consultas, aplicarlas sobre la base de datos y devolver el resultado correspondiente. El SIG no accede a los datos, sino que se comunica con el SGBD y deja en manos de este el proceso de consulta en sí.
- **Protección de los datos.** Si la base de datos almacena información sensible, el SGBD debe **controlar el acceso** a esta, restringiendo el acceso cuando corresponda (por ejemplo, estableciendo distintos permisos de acceso para distintos tipos de usuarios) e implementando los mecanismos de protección necesarios.

- **Eficiencia.** El SGBD debe ser capaz de gestionar de forma fluida **grandes volúmenes de datos o de operaciones** (por ejemplo, muchos usuarios accediendo simultáneamente), de modo que dé una respuesta rápida a las peticiones de los usuarios de la base de datos.
- **Gestión de transacciones.** Las operaciones sobre la base de datos tales como la adición o borrado de un registro se realizan mediante transacciones. El SGBD ha de encargarse de gestionarlas de manera eficiente y segura para que todos los usuarios de la base de datos puedan hacer su trabajo de forma transparente. Se denomina **transacional** al SGBD capaz de garantizar la integridad de los datos, no permitiendo que las transacciones puedan quedar en un estado intermedio.

En este sentido, resulta de especial importancia la existencia de **lenguajes de consulta**, que son los que una aplicación utilizará para comunicarse con el SGBD y expresar las operaciones que desea realizar con los datos de la base de datos. El lenguaje de consulta mas extendido es **SQL** (Standard Query Language)

Bases de datos espaciales

Todo cuanto hemos visto en los puntos anteriores constituye el conjunto de ideas fundamentales sobre las que se asienta la creación y uso de bases de datos de cualquier índole. La inclusión de datos espaciales en este esquema no es en absoluto obvia, y presenta una complejidad adicional que requiere de nuevos planteamientos. Para que una base de datos pueda considerarse espacial, debe adaptarse y añadir elementos adicionales.

En primer lugar, el dato espacial debe poder **almacenarse de forma nativa** en la base de datos. Esto quiere decir que una geometría debe poder almacenarse asociada a una entidad en una tabla, del mismo modo que sucede con otros tipos de datos tales como valores numéricos o cadenas de texto. El dato no solo debe poder almacenarse, sino tambien **entenderse** por parte del SGBD, para así

comprender la naturaleza su naturaleza espacial y poder responder a peticiones relativas a esta por parte del usuario. Esto se denomina **almacenamiento transparente**, frente al almacenamiento **opaco**, en el cual la base de datos es capaz de recoger cualquier tipo de valor, pero sin ser capaz de entender su naturaleza y utilizarlo.

Aunque existen planteamientos para almacenar datos raster en bases de datos, en general **las bases de datos espaciales trabajan con datos vectoriales** y están mejor adaptadas a estos. Las geometrías del dato vectorial se incluyen, como ya se ha dicho, dentro de los valores de un registro de la tabla (que se corresponde con una entidad en el modelo de datos vectoriales). Por su parte, la componente temática del dato espacial puede almacenarse sin problema en la base de datos, sin necesidad de ningún tipo de adaptación.

Cuando la base de datos está preparada para almacenar y trabajar con los datos espaciales, es necesario **adaptar el lenguaje de consulta**. A las operaciones habituales que un SGBD es capaz de realizar en función de los datos, se añaden otras que utilizan las propiedades espaciales del dato espacial. Se tiene así un **lenguaje de consulta espacial**, que permite consultas con componente espacial.

CONSULTAS

Entendemos por consulta una operación en la cual *preguntamos* a los datos geográficos algún tipo de cuestión simple, generalmente basada en conceptos formales sencillos. Este tipo de análisis, aunque no implica el uso de conceptos analíticos complejos, es uno de los elementos clave de los SIG, pues es parte básica del empleo diario de estos.

Aunque las consultas no son algo exclusivo de las bases de datos, es mediante el concurso de estas que adquieren toda su potencia y permite efectuar gran parte de las operaciones habituales de un SIG.

En el contexto espacial, una consulta representa un uso similar al que damos a un mapa clásico, cuando en base a este respondemos a preguntas como *¿qué hay en la localización*

X? o *¿qué ríos pasan por la provincia Y?* No obstante, no debemos olvidar que los datos espaciales tienen dos componentes: una espacial y otra temática. Preguntas como las anteriores hacen referencia a la componente espacial, pero igualmente pueden efectuarse consultas que se apliquen sobre la parte temática. Y más aún, pueden efectuarse consultas conjuntas que interroguen a los datos geográficos acerca de los atributos espaciales y temáticos que estos contienen.

Un ejemplo muy sencillo de consulta es lo que conocemos en un SIG como **selección**. De todos los registros de la tabla de datos, aquellos que cumplen el criterio indicado se marcan como seleccionados, y posteriormente pueden utilizarse únicamente estos como base de otro análisis, o simplemente el usuario puede ver cuáles han sido los seleccionados para así obtener la respuesta a su consulta. La figura 7.1 muestra un ejemplo típico de esto, en el que el usuario define un área rectangular y se seleccionan las entidades que cumplen la condición de intersecar con él. Los criterios de selección pueden ser sencillos como este, o más elaborados y complejos.

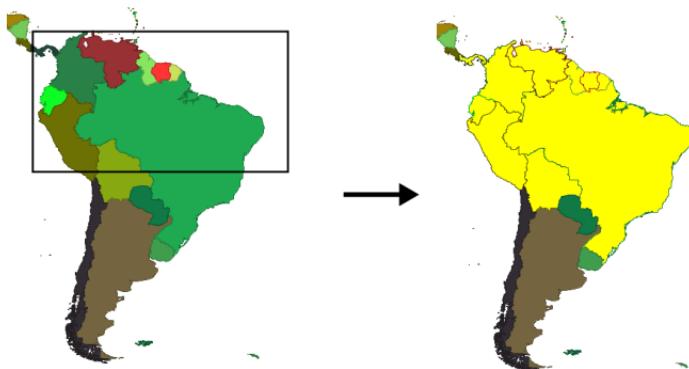


Figura 7.1: Una consulta sencilla mediante la definición gráfica de un rectángulo. Todas las entidades dentro de este son el resultado de la consulta y quedan seleccionadas en el SIG.

Una consulta nos vale también para extraer información

de una base de datos de acuerdo a nuestras necesidades, y para crear posteriormente y a partir de dicha información una nueva capa. Esta operación es útil cuando la base de datos de la que disponemos es muy voluminosa y solo resulta de interés para nuestro trabajo una parte de ella. Puede tratarse de una parte en el sentido espacial (la base de datos contiene datos a nivel mundial y se quiere trabajar a nivel estatal), en el sentido temático (la base de datos contiene mucha información de cada entidad y solo interesan algunos campos), o en una combinación de ambas. Para extraer dicha parte y trabajar únicamente con ella, utilizaremos una consulta.

Veamos algunos ejemplos de consultas. Sea una capa con los distintos países del mundo y una serie de valores económicos y sociales asociados a cada uno de ellos. Consideremos las siguientes preguntas:

- ¿Qué países tienen un Producto Interior Bruto mayor que el de España?
- ¿Qué países han experimentado un crecimiento económico en el último año?
- ¿Cuántos países tienen más de 200 millones de habitantes?

En todos estos casos estamos haciendo referencia a países, los cuales, como sabemos, estarán asociados a elementos geométricos que definen sus propiedades espaciales, es decir, a una componente espacial. Esta componente es la que permite que, además de poder plantear las consultas anteriores, podamos representar cada país en la pantalla y visualizarlo, o saber cuáles de ellos se encuentran en el hemisferio norte (esta sería una consulta espacial, de las que más adelante en este mismo capítulo veremos).

Sin embargo, cuando realizamos consultas como las tres anteriores, no acudimos para nada a la componente espacial. Consultas como estas podrían resolverse si en lugar de una capa dentro de un SIG tuviéramos, por ejemplo, un simple anuario estadístico lleno de tablas con datos correspondientes a cada país.

Las consultas pueden incluir varios criterios en una sola pregunta. Por ejemplo:

- ¿Qué países de la zona euro tienen más de 40 millones de habitantes?
- ¿En qué países de habla inglesa aumentó la población durante el último año?

Para expresar esas consultas se han de incluir elementos de la denominada **lógica booleana**¹. Esta implica el uso de **operadores lógicos**, mediante los cuales se reescribirían las consultas anteriores de la siguiente manera:

- ¿Qué países tienen como moneda el euro *y* a la vez tienen más de 40 millones de habitantes?
- ¿Que países hablan inglés *y* sufrieron un aumento de población durante el último año?

Los lenguajes de consulta que se emplean para transmitir estas operaciones a un SGBD, permiten el uso de tales operadores para formular consultas.

Si el SGBD es de tipo espacial y *entiende* que algunas de las columnas de una tabla contiene información espacial (es decir, que cada entidad no solo tiene información temática), pueden plantearse consultas que hacen uso de esta, tales como las siguientes.

- ¿Qué países comparten frontera con Alemania?
- ¿Cuántos países se encuentran completamente en el hemisferio sur?
- ¿Qué países están a menos de 2000 km de España?

Para dar respuesta a esas cuestiones, basta analizar la componente espacial y no necesitamos para nada los datos con los que hemos trabajado anteriormente. Son consultas puramente espaciales. Aunque estas consultas amplían lo que ya conocemos, en realidad no abren ninguna nueva vía de estudio de los datos geográficos. Son consultas a las que podríamos responder utilizando un mero mapa impreso, sin aprovechar el hecho de que dentro de un SIG las componentes espacial y temática se hallan íntimamente vinculadas.

¹Denominada así por el matemático irlandés George Boole(1815, 1864)

La verdadera potencia de las consultas espaciales la encontramos en la combinación de estas consultas sobre la componente espacial y las que vimos anteriormente sobre la componente temática. Así, se pueden plantear, por ejemplo, cuestiones como:

- ¿Qué países del hemisferio norte tiene una densidad de población mayor que la de Perú?
- ¿Cuántos países con más de 10 millones de habitantes se encuentran a menos de 1000 km de la frontera de Rusia?

Estas consultas incorporan elementos que hacen necesario acudir a la tabla de atributos, y otros que requieren analizar la componente espacial, estudiando las relaciones espaciales y topológicas de las geometrías asociadas.

Las consultas pueden incluir **varias capas**. Por ejemplo, si además de la capa de países disponemos de una capa de ríos del mundo, podríamos responder a la pregunta *¿qué países atraviesa el Nilo?*

Igualmente, las uniones entre tablas que hemos visto para el caso de la componente temática pueden establecerse mediante un criterio espacial. Se tiene así una **unión espacial**.

Un ejemplo muy sencillo de unión espacial es el que encontramos si combinamos la capa de países del mundo que venimos utilizando con una capa de ciudades del mundo. Podemos unir a la tabla de esta segunda capa todos los valores que caracterizan al país al que pertenece cada ciudad. Si existe un campo común entre ambas tablas de atributos (por ejemplo, el nombre del país), esto serviría para efectuar esta unión. No obstante, esto no es necesario, ya que existe otro elemento común que no se encuentra almacenado dentro de la tabla, pero que puede tomarse de la componente espacial: toda ciudad debe estar situada dentro de los límites del país al que pertenece. Esto sirve para establecer la relación entre las tablas, y cada ciudad debe relacionarse con aquella entidad dentro de cuya geometría se encuentre el punto que la representa.

Índices espaciales

Si realizamos una consulta a una base de datos, el resultado es un subconjunto de esta con los elementos que cumplen el criterio expresado en la consulta. Si se implementa de forma *directa* dicha consulta, esta operación implica comprobar todos los elementos de la base de datos y ver cuáles son los que cumplen con el citado criterio. Teniendo en cuenta que una base de datos puede tener un gran tamaño, esta forma de proceder no es la óptima.

Los índices nos permiten *alcanzar* los elementos que constituyen la respuesta a nuestra consulta, haciéndolo de la forma más rápida y llegando hasta ellos sin tener que pasar por todos los restantes.

Un ejemplo fácil de entender es el de una guía telefónica en la que los nombre están ordenados alfabéticamente. Gracias a ese orden y a que se conoce el alfabeto, se puede buscar rápidamente un nombre sin necesidad de leer todos ellos.

Al utilizar una base de datos, si no disponemos de un índice deberemos recorrer toda ella para dar respuesta a nuestras consultas. No sabemos *dónde* buscar las respuestas a nuestras consultas, del mismo modo que si en una guía telefónica no supiéramos que carece de sentido buscar en la letra F el número telefónico del señor Pérez.

Ademas de índices para datos de tipo numérico o texto, en los que resulta obvio establecer un orden natural, en el ámbito de los SIG tienen importancia los denominados **índices espaciales**. Aunque sus fundamentos teóricos son distintos, el concepto es similar al de índices de bases de datos no espaciales: elementos que permiten **optimizar las consultas mediante una correcta estructuración** de los datos, en particular en este caso de su componente espacial.

Puede entenderse la idea de un índice espacial mediante un sencillo ejemplo de cómo empleamos ideas parecidas a los índices espaciales de forma natural cuando tratamos de resolver una consulta espacial sin la ayuda de un SIG. Supongamos que tenemos nuestro mapa de países del mun-

do y queremos averiguar qué países tienen su frontera a menos de 3000 kilómetros de la frontera de España. ¿Cómo operaríamos de manera natural para dar respuesta a esta consulta?

La solución más inmediata es medir la distancia entre España y todos los países restantes, y después tomar aquellos que hayan arrojado un resultado de distancia menor a 3000. La operación daría el resultado esperado, pero implicaría un gran número de mediciones, y no sería una forma óptima de operar. Más probable es que no efectuemos mediciones con los países de América, pues un conocimiento básico de geografía basta para saber que todos ellos se encuentran a más de 3000 kilómetros. No sabemos exactamente a qué distancia se encuentran, pero sabemos que no van a cumplir el criterio establecido en la consulta.

Ese conocimiento básico de geografía que tenemos es en realidad una especie de índice espacial. No sirve para saber las distancias exactas ni resolver la consulta por completo, pero sirve para **dar una aproximación y facilitar el trabajo**. Descartamos un buen numero de países de forma casi inmediata, y luego solo realizamos las operaciones costosas (la medición) con un subconjunto del total.

De modo similar, los índices espaciales nos permiten obtener resultados en un área concreta sin necesidad de analizar todo el espacio ocupado por el total de los datos. Gracias a ello, hacen las consultas mas efectivas y permiten trabajar con grandes volúmenes de datos.

Los índices espaciales se almacenan junto con los datos a los que hacen referencia, bien en ficheros adicionales o dentro de la propia base de datos, en case de utilizarse una. Los SGBD espaciales tienen **capacidades para calcular estos índices espaciales** y almacenarlos en la base de datos, recurriendo a ellos cuando se realiza una consulta que requiera su uso.

ANÁLISIS ESPACIAL. FUNDAMENTOS

El análisis espacial es una de las tareas fundamentales sin las cuales el concepto de SIG no alcanza su verdadero significado.

El análisis espacial es el **estudio cuantitativo de aquellos fenómenos que se manifiestan en el espacio**. Ello indica una importancia clave de la posición, la superficie, la distancia y la interacción a través del propio espacio.

Ejemplos de análisis que realizamos con cartografía fuera de un SIG son el buscar en un mapa dónde se sitúa el pico más alto, ver la elevación concreta a la que se encuentra un elemento dado tal como una población, o planificar una jornada turística viendo qué lugares de interés podemos visitar o cómo llegar desde uno a otro de estos lugares haciéndolo por las mejores carreteras o de la forma más rápida. Estas actividades habituales son ejemplos de análisis geográficos que podemos igualmente realizar dentro de un SIG.

Mediante el análisis podemos generar nuevos datos que pueden ser **nuevas capas de datos geográficos, tablas de datos, valores escalares o vectores**.

En ocasiones, los resultados expresan **la misma variable** que el dato de partida (por ejemplo, el cálculo de una media), y en otros las variables de entrada y salida son **diferentes** (por ejemplo, si a partir de una capa de elevaciones calculamos una de pendientes).

Asimismo, todo análisis espacial parte de un conjunto de datos espaciales, pudiendo estos ser **de un único tipo, o de varios distintos** que se combinan en un procedimiento concreto. Por ejemplo, en el caso de calcular la localización del punto más alto, el resultado es una sencilla coordenada, y tan solo se utiliza la variable elevación. En el caso de la altura media de una ciudad, se utilizan dos entradas: por un lado, la elevación, y por otro, el emplazamiento de la ciudad. Aunque un mapa clásico contiene toda esa información en una única hoja, en realidad son dos elementos distintos combinados a la hora de representarlos. En términos más acordes con un SIG, podemos decir que tenemos dos capas distintas que utilizamos como entradas.

El análisis dentro un SIG nos permite tanto formular como responder a cuestiones. Estas cuestiones pueden ser:

- Relativas a posición y extensión
- Relativas a la forma y distribución
- Relativas a la asociación espacial
- Relativas a la interacción espacial
- Relativas a la variación espacial

Algunos ejemplos de análisis geográfico son los siguientes.

- **Consulta espacial.** Vimos las consultas en detalle dentro del capítulo dedicado a las bases de datos.
- **Análisis topológico.** Pueden plantearse consultas referidas no solo a la posición de los elementos geográficos, sino a la **relación con otros elementos**. La existencia de topología puede emplearse para la realización de consultas que respondan a cuestiones como:

¿Cómo llegar desde mi posición actual hasta una coordenada concreta por la red viaria existente?

¿Qué comunidades autónomas comparten límite con Madrid?

- **Medición.** La existencia de una referencia espacial para cada uno de los elementos con los que trabajamos en el análisis dentro de un SIG hace que podamos **cuantificar** otra serie de parámetros también espaciales. Entre las mediciones más básicas, encontramos las distancias,

áreas, perímetros o factores de forma. Mas elaboradas, encontramos otras como pendientes o indices derivados de medidas simples.

- **Combinación.** Uno de los procedimientos más habituales y más característicos dentro del uso de un SIG es la **combinación o superposición** de varias capas de información. La propia estructura de la información geográfica en capas facilita notablemente estos procedimientos y convierte a los SIG en plataformas ideales para llevar a cabo análisis donde se combina información sobre diversas variables.
- **Transformaciones.** Podemos englobar dentro de este grupo una amplia serie de procedimientos que modifican los elementos de entrada de diversas formas. Entre los más habituales, encontramos la **transformación de coordenadas**, la **simplificación de geometrías**, o la **creación de áreas de influencia** o la **reclasificación de valores**. Estas transformación pueden afectar tanto a la componente espacial como a la componente temática del dato.
- **Análisis de superficies.** El análisis de superficies es uno de los más potentes de cuantos encontramos en un SIG. Desde parámetros básicos como la **pendiente** o la **orientación** hasta parámetros morfométricos muy específicos, pasando por todas las herramientas del **análisis hidrológico**, la batería de operaciones disponibles es muy amplia.
- **Estadística descriptiva.** Los elementos de la estadística clásica tienen sus equivalentes en los datos espaciales, y nos permiten **calificar cuantitativamente** los datos con los que trabajamos. Se incluyen aquí descriptores de centralidad y dispersión, de dependencia espacial o el estudio de patrones espaciales, entre otros muchos. Estos pueden a su vez usarse para el contraste de hipótesis que contengan una cierta componente espacial.
Por ejemplo, estos estadísticos nos permiten dar respuesta a cuestiones del tipo

¿Es constante la media de altura a lo largo de toda la geografía de mi país?

¿Existe alguna dirección predominante en los movimientos de individuos de una especie o se desplazan erráticamente?

- **Inferencia.** Otro análisis estadístico de gran importancia en los SIG es el que permite inferir comportamientos de las distintas variables y estudiar, por ejemplo, la forma en que estas van a evolucionar a lo largo del tiempo.

El establecimiento de modelos de cambio y variación representa una de las herramientas más actuales en el campo de los SIG, y un campo en abundante desarrollo.

- **Toma de decisiones y optimización.** La estructura de la información geográfica en capas dentro de un SIG, favorable como ya vimos para la superposición de capas, lo es igualmente para estudiar de forma combinada los efectos de distintos factores. Este estudio nos permite luego responder a cuestiones como, por ejemplo,

¿Cuál es el mejor lugar para emplazar una nueva construcción en función de su impacto sobre el medio?

¿Dónde situar un nuevo hospital para que el servicio en la comarca mejore lo máximo posible?

- **Modelización.** La creación de modelos espaciales dentro de un SIG es una tarea aún pendiente de mucho desarrollo. No obstante, existe un gran número de modelos en los más diversos campos, y la arquitectura de datos y procesos de los SIG es propicia para la implementación de otros nuevos.

PARTICULARIDADES DE LOS DATOS ESPACIALES PARA SU ANÁLISIS

Las características propias de los datos espaciales dotan a estos de una **gran potencialidad** de análisis, al tiempo que **condicionan o limitan** otras operaciones. Algunas de estas características representan problemas que han de tenerse presentes en el análisis; otros son simplemente conceptos básicos que deben conocerse pero no han de implicar necesariamente una dificultad asociada.

Escala

A la hora de estudiar la información geográfica, podemos hacerlo a **distintos niveles** y, dependiendo del nivel elegido, los resultados serán de una u otra naturaleza. Debido a esto, además de considerar la escala cartográfica para la representación y gestión de datos en un SIG, es necesario considerar la **escala de análisis**.

La escala de análisis depende del **dato en sí** (precisión, tipo, etc.), así como del **análisis que se va a realizar** con él.

Como se muestra en la figura 8.1, si para definir las formas de relieve en un punto dado lo hacemos considerando dicho punto y los valores de elevación a su alrededor, la caracterización que hagamos varía en función de la dimensión de esa zona alrededor (que es la que define la escala de análisis). Para valores pequeños de dicha zona de análisis, el punto analizado puede definirse como una cima, mientras que aumentando la escala de análisis se advierte que el punto se sitúa en el fondo de un valle.

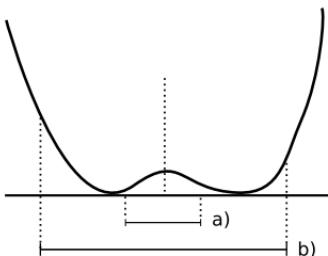


Figura 8.1: Dependiendo de la escala de análisis, un mismo relieve puede ser caracterizado como cima (a) o fondo de valle (b)

Por tanto, debemos observar el relieve desde la distancia correcta a la cual la información que nos proporciona es la más adecuada para un análisis dado. Además de existir una escala de mayor relevancia para un análisis concreto, es de interés el trabajar a **múltiples escalas** y combinar los resultados.

Otro ejemplo de cómo la escala de análisis condiciona los resultados obtenidos lo encontramos en el caso de efectuar **mediciones**. Como puede verse en la figura 8.2, la unidad de medida empleada provoca que se obtengan resultados distintos.

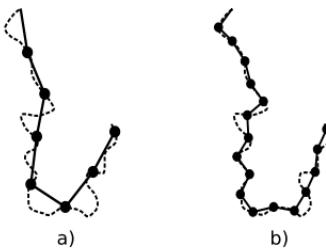


Figura 8.2: La unidad de medida empleada modifica el resultado obtenido.

La unión del valor resultante con la escala a la que se ha obtenido tiene en conjunto pleno significado, pero ese valor por sí mismo carece de dicho significado.

El concepto de **fractal** tiene una implicación directa en este hecho.

El propio **formato de almacenamiento** condiciona el efecto de la escala, ya que puede imponer límites. Tal es el caso cuando se trabaja con capas raster, en las que el **tamaño de celda** delimita la precisión que puede obtenerse en el análisis.

El Problema de la Unidad de Área Modificable

Muchas de las variables con las que trabajamos dentro de un SIG **no pueden medirse de forma puntual**, y por ello han de **estudiarse para un área dada**. Ejemplos de este tipo de variables son el porcentaje de población en un rango de edad determinado o la densidad media de población.

Las áreas que se definen para poder trabajar con las variables de esta índole son **esencialmente arbitrarias**, tales como países, regiones o distritos, que se establece sin

ningún criterio propio del análisis espacial. La utilización de una u otra unidad **altera los resultados** extraídos de las variables estudiadas.

Este problema, por tener relación con la elección de la unidad de agregación de la información, se conoce como *Problema de la Unidad de Área Modificable*(PUAM).

Un problema particular relacionado con el PUAM es la denominada **falacia ecológica**, que consiste en asumir que los valores calculados para una unidad de área pueden aplicarse a los individuos de la población existente en dicha área. Sólo en el caso de que exista una completa homogeneidad para la variable analizada, lo cual raramente sucede, la anterior suposición sería cierta.

Autocorrelación espacial

Se denomina **autocorrelación espacial** a la existencia de una **correlación de la variable consigo misma**, de tal modo que los valores de esta variable en un punto guardan relación directa con los de esa misma variable en otros puntos cercanos. Por ejemplo, en el caso de medirse la temperatura, los puntos cerca de un foco de calor tendrán una temperatura mayor que aquellos cerca de focos fríos. Si estudiamos la distribución de una enfermedad infecciosa, es más probable que los casos se encuentren agrupados, de forma que la presencia de un alto número de casos implique también una alta incidencia en poblaciones cercanas.

Otra forma de expresar la autocorrelación espacial es mediante la conocida como **Primera Ley Geográfica de Tobler**, que establece que «todo está relacionado con todo, pero las cosas próximas entre sí están más relacionadas que las distantes».

La autocorrelación espacial, tal y como se ha descrito antes, es **positiva**. Puede, no obstante, existir una *autocorrelación espacial negativa*, si los valores altos se rodean de valores bajos y viceversa.

En caso de no existir ningún tipo de autocorrelación espacial, se tiene que los datos recogidos en una serie de puntos

son **independientes entre sí** y no se afectan mutuamente, sin que tenga influencia de la distancia.

La figura 8.3 muestra unas sencillas capas ráster en las que se presentan los tres tipos de autocorrelación espacial anteriores.

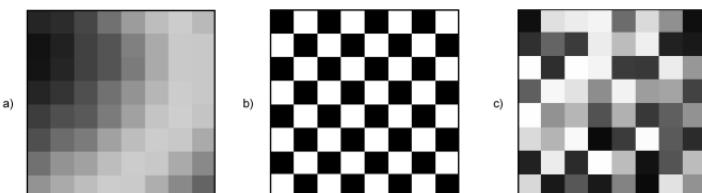


Figura 8.3: a) Autocorrelación espacial positiva. b) Autocorrelación espacial negativa. c) Ausencia de autocorrelación espacial (independencia)

Las consecuencias de la existencia de autocorrelación espacial son numerosas y de gran importancia.

Por una parte, muchos de los análisis estadísticos suponen la **independencia de la variable**. Puesto que existe una dependencia de la componente espacial, será necesario para obtener resultados correctos introducir dicha componente espacial como una variable más.

Algo similar sucede cuando los datos presentan alguna **tendencia espacial** (los valores de una variable están relacionados con sus propias coordenadas geográficas), ya que esto también invalida el supuesto de la independencia de los datos.

Existiendo autocorrelación espacial, y siendo esta positiva, la **inferencia estadística es menos eficaz** que si se cuenta con un número igual de observaciones de una variable independiente.

La autocorrelación espacial no es, no obstante, un elemento que siempre tenga consecuencias negativas. Puesto que los puntos cercanos a uno dado guardan relación con este, la autocorrelación puede aprovecharse para **estimar valores** en un punto cualquiera si conocemos los valores en

puntos cercanos. Este es el fundamento de los **métodos de interpolación**

Existencia de estructura

Tanto la disposición de los datos como las propiedades de la variable estudiada (por ejemplo, la propia autocorrelación espacial como propiedad intrínseca) exhiben una estructura determinada. Esta estructura puede condicionar los resultados del análisis y tener influencia sobre estos.

Los dos principales conceptos estadísticos que definen la estructura espacial de los datos son la **estacionaridad** y la **isotropía**. La estacionaridad indica que el proceso es **invariante a la traslación**. Es decir, que las propiedades son constantes en el espacio y no existe tendencia alguna. La isotropía indica que el proceso es **invariante a la rotación** y tiene lugar del mismo modo en todas direcciones.

Efectos de borde

Las zonas que estudiamos dentro de todo análisis espacial **tienen unos límites establecidos**. Estos límites vienen definidos de forma artificial —el límite de la fotografía aérea de la que disponemos, por ejemplo— o bien de forma natural —si estudiamos un bosque junto a un pantano, el bosque encuentra su límite al borde de este último—. La presencia de estos bordes **distorsiona el resultado de los análisis**, en especial para aquellos parámetros no puntuales que requieren la definición de un área de estudio (densidades, etc., como ya vimos para el caso del PUAM)

La figura 8.4 muestra un ejemplo de efecto de borde.

En algunos casos, el efecto de borde no se manifiesta únicamente para puntos cercanos a dicho borde, sino para **todos aquellos relacionados o conectados** con él según un determinado criterio, con independencia de su distancia a este.

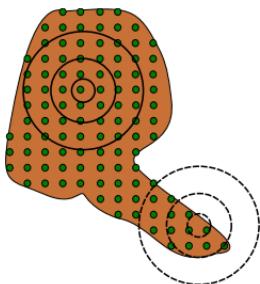


Figura 8.4: Representación del efecto borde y cómo este afecta en mayor o menor medida en función de la escala de análisis. Las zonas en trazo continuo no se ven afectadas. Las zonas en trazo punteado están afectadas de efecto de borde en diferente grado.