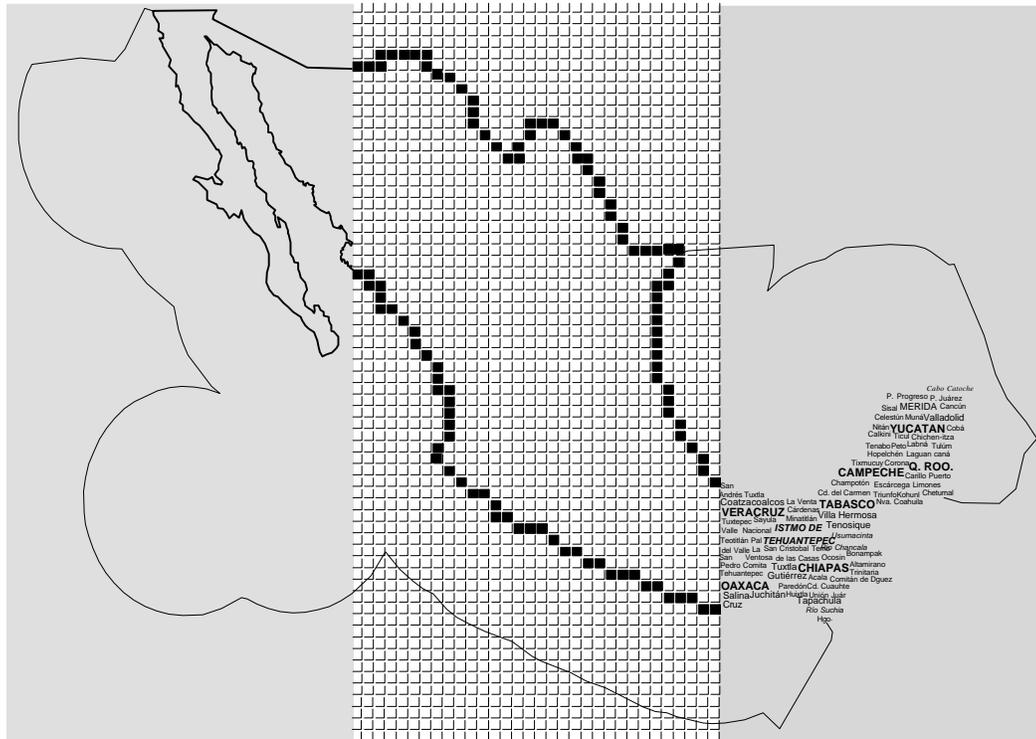


BASE DE DATOS GEOGRÁFICOS



MODELO DE DATOS VECTORIALES

INTRODUCCIÓN GENERAL

MODERNIZACIÓN DE LA ACTIVIDAD GEOGRÁFICA DEL INSTITUTO NACIONAL DE ESTADÍSTICA, GEOGRAFÍA E INFORMÁTICA (INEGI).

El INEGI tiene, entre otros objetivos, el de producir la información geográfica que el país requiere para planificar y conducir el aprovechamiento óptimo de sus recursos naturales y culturales. Para ello, la Dirección General de Geografía (DGG) del INEGI, ha realizado diversos grados de cubrimiento cartográfico del territorio nacional, en diversas escalas de temas tales como Topografía, Geología, Hidrología, Edafología, Uso del Suelo, etc. Esta cartografía ha sido utilizada de muy diversas maneras para el estudio de las obras o eventos de mayor relevancia que han acontecido en el Territorio Nacional y la Zona Económica Exclusiva durante el último cuarto de siglo.

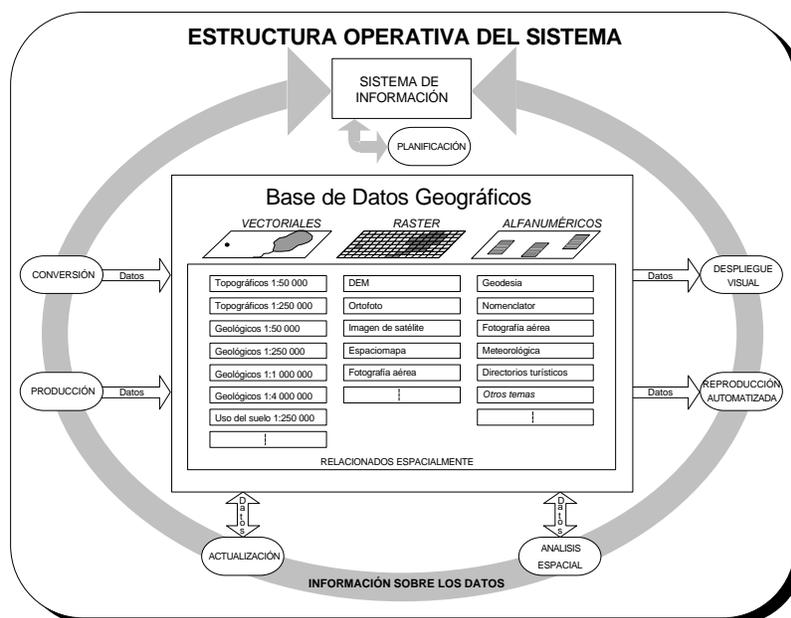
Sin embargo, el cambio tanto cualitativo como cuantitativo en las necesidades de información de la sociedad, y la relativa lentitud de los procesos manuales (analógicos) tradicionales de producción de información para responder con eficiencia a estas demandas, creó la necesidad de buscar alternativas para agilizar de manera sustancial el proceso de generación de información geográfica. Como respuesta a esta necesidad se emprende el Proyecto de Modernización de la Actividad Geográfica del INEGI, aprovechando el desarrollo de nuevas tecnologías y recursos de procesamiento digital de información geográfica, que ofrecen mayor agilidad y/o precisión que sus predecesores analógicos.

Este proyecto sienta las bases generales de carácter normativo conforme a las cuales se establecerán las acciones de integración de la información geográfica producida en el Instituto y otras dependencias y entidades de la Administración Pública Federal en un Sistema Nacional de Información Geográfica (SNIG).

Para cumplir los objetivos del proyecto se establece una estructura operativa con un conjunto de capacidades o funciones que se indican a continuación, y que permiten:

- capturar los datos que en forma de documentos, mapas e imágenes, existen;
- producir y actualizar la información de regiones no cubiertas o desactualizadas;
- almacenar y recuperar la información, de manera selectiva, confiable y eficiente;
- procesar los datos usando herramientas estadísticas y de análisis espacial;
- desplegar en pantalla y generar resultados con calidad de edición cartográfica;
- y supervisar y controlar la operación en general

De las funciones anteriores, en la tercera, denominada Base de Datos Geográficos, (BDG) residirá la representación digital de los datos que conforman la totalidad de información geográfica producida por la Institución. A continuación se presenta un diagrama general con los componentes de la estructura.



DISEÑO DE UNA BASE DE DATOS

El desarrollo de cualquier tipo de Base de Datos (BD), comprende tres etapas típicas:

- El diseño conceptual, en el cual se identifican tanto los requerimientos de información, como los datos disponibles en cada una de las áreas de la Institución, generando el **modelo conceptual** y las descripciones específicas de los datos producidos en cada una de ellas. La descripción de cada uno de los datos de una área en particular, conforma el **Diccionario de Datos** del área o tema.
 - La segunda etapa se denomina diseño lógico, y consiste en la integración de los modelos particulares en un modelo global de información. Este modelo global se analiza con el fin de eliminar redundancias, y se realiza una optimización general, para construir el esquema formal (modelo lógico) de la BD, que define su estructura en un lenguaje establecido.
 - La última etapa, denominada implementación física o modelo físico, consiste en la creación de la representación computacional (diseño y creación de registros, archivos, métodos de acceso, restricciones de seguridad, etc.) de la BD. Es hasta esta etapa en que la BD se materializa en un equipo de computación, utilizando las capacidades que el equipo seleccionado presenta en particular. Ello también significa que un modelo lógico o conceptual determinado puede tener diferentes implementaciones físicas, dependiendo del equipo computacional o de los equipos en que se decida implementarlo.
- Las tres etapas aquí descritas son dinámicas, pues los modelos generados en cada una requieren de revisión y actualización constantes, debido, entre otros factores, a la evolución de las necesidades internas y externas de información, al desarrollo de más detalladas especificaciones de los datos, o a la disponibilidad de nuevos equipos y herramientas computacionales.

EL DISEÑO DE LA BDG.

En el caso particular de la BDG, se inicia el diseño conceptual con la revisión del contenido de cada uno de los productos tradicionales de información (cartas, reportes, estudios). Esta revisión brinda la oportunidad de analizar en toda su extensión, los contenidos, detectando elementos de información que podrían ser complementarios, o posibles duplicaciones e inconsistencias de información.

Convertir los productos de información geográfica de forma analógica a forma digital, impone la necesidad de considerar que los mecanismos de percepción y análisis de información digital difieren de los tradicionales. Los productos convertidos serán procesados por computadoras, y aunque pueden ser visualizados en monitores gráficos, su análisis se realiza fundamentalmente por la combinación de métodos de análisis geométrico, métodos estadísticos, y consultas de bases de datos (sistemas de información geográfica).

Los datos que constituyen esta información se clasifican, de acuerdo a su naturaleza, en tres tipos: vectorial, raster y alfanuméricos. El tipo vectorial contiene los datos provenientes de las cartas que a diferentes escalas y temas se han producido; el tipo raster contiene la información de tipo imagen, de rejilla o teselar, tal como las imágenes satelitarias y los modelos digitales de elevación. El tipo alfanumérico comprende los datos tabulares y textuales, tales como los reportes de campo, o los resultados de análisis de muestras en laboratorio.

Para cada uno de estos tipos de información, es necesario contar con la definición explícita de su contenido, estructura, relaciones y normas que los rigen. Estas características constituyen el **Modelo de Datos** específico para cada tipo de dato. De esta manera se conforman el Modelo de Datos Vectoriales, el Modelo de Datos Raster y el Modelo de Datos Alfanuméricos.

USTED ESTÁ AQUÍ



LOS MODELOS DE DATOS

- I Modelo de datos vectoriales
- II Modelo de datos alfanuméricos
- III Modelo de datos raster



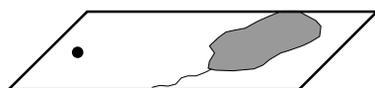
ACERCA DE LOS DICCIONARIOS DE DATOS

- I Vectoriales
- II Alfanuméricos
- III Raster

TEMA

- | | | |
|--|--|---|
| <input type="checkbox"/> Aeronáutica | <input type="checkbox"/> Fisiografía | <input type="checkbox"/> Topografía |
| <input type="checkbox"/> Batimetría | <input type="checkbox"/> Geología | <input type="checkbox"/> Recursos Turísticos |
| <input type="checkbox"/> Climatología | <input type="checkbox"/> Hidrología de Aguas Subt. | <input type="checkbox"/> Uso del Suelo y Vegetación |
| <input type="checkbox"/> Edafología | <input type="checkbox"/> Hidrología de Aguas Sup. | <input type="checkbox"/> Uso Potencial Agricultura |
| <input type="checkbox"/> Efect. Clim. Reg. May.-Oct. | <input type="checkbox"/> Humedad del Suelo | <input type="checkbox"/> Uso Potencial Forestería |
| <input type="checkbox"/> Efect. Clim. Reg. Nov.-Abr. | <input type="checkbox"/> Precip. Total Anual | <input type="checkbox"/> Uso Potencial Ganadería |
| <input type="checkbox"/> Evapotrans. y Déf. de Agua | <input type="checkbox"/> Temp. Medias Anuales | <input type="checkbox"/> Otro |

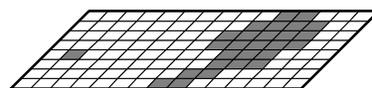
TIPO DE DATOS



Vectorial



Alfanumérico



Raster

ESCALA



1:50 000



1:250 000



1:1 000 000



1:4 000 000

TABLA DE CONTENIDO

MODELO DE DATOS VECTORIALES.....	1
1. Introducción	1
2. Conjunto de datos	1
3. Características de los datos	1
3.1 Entidad.....	1
3.2 Representación geométrica.....	2
3.2.1 Punto	2
3.2.2 Línea	2
3.2.3 Área.....	3
4. Estructura espacial de los datos	4
4.1 Relaciones	4
4.1.1 Conectar	4
4.1.2 Compartir.....	6
4.2 Integración de conjuntos de datos	8
5. Aspectos generales	9
5.1 Precisión	9
5.2 Resolución	10
5.3 Tolerancias	10
5.4 Sistema de referencia, proyección cartográfica y sistemas de coordenadas	10
5.4.1. Sistema geodésico de referencia	10
5.4.2. Sistema de alturas y datum vertical	10
5.4.3. Sistema de coordenadas horizontales y proyección cartográfica	10
5.5 Límites de conjuntos de datos	11

MODELO DE DATOS VECTORIALES

1. Introducción

Para propósitos del modelo vectorial, los objetos reales se suponen inscritos en un sistema coordinado, el cual permite determinar con precisión, su posición, sus dimensiones y el número de éstas. Así, en el modelo vectorial, los objetos se representan mediante la especificación de los valores de estas magnitudes.

Se ha incorporado el tipo de representación vectorial de datos a la BDG, debido a que la información cartográfica producida por la DGG, concuerda con este principio.

Las cartas consideradas en este modelo son: la topográfica, geológica, edafológica, uso del suelo, climática, humedad del suelo, evapotranspiración, hidrológicas, fisiográfica, uso potencial, turística y batimétrica, las cuales existen en una o más de las siguientes escalas 1:50 000, 1:250 000, 1:1 000 000 y 1:4 000 000

El modelo consta de dos secciones: una general, relacionada con el modelo conceptual de datos vectoriales, en la que se describen la estructura, características, convenciones, y reglas de integridad que rigen este tipo de información, y una particular, que contiene los diccionarios de datos que definen la manera en que son representados los elementos geográficos para cada escala, dentro de cada tema cartográfico considerado.

Nota: Las figuras que ilustran este documento pueden provenir de diversos temas y están representadas a diferentes escalas.

2. Conjunto de datos

En el "Sistema Nacional de Información Geográfica", la unidad de información está determinada de acuerdo con la división tradicional en hojas de cada una de las cartas en las escalas: 1:50 000, 1:250 000, 1:1 000 000 y 1:4 000 000.

Un **conjunto de datos** vectoriales consiste de toda la información existente en una hoja para cualquiera de dichas escalas. Más información se encuentra en la sección 5.5 de este documento. El Sistema toma como referencia el ITRF92 (Marco de Referencia Terrestre Internacional de 1992).

Los **datos auxiliares** de un conjunto de datos ofrecen una información relacionada con aspectos tales como método de compilación, proyección, etc. (ver: "Acerca de los diccionarios de datos vectoriales", en la sección de datos auxiliares).

3. Características de los datos

Para representar un rasgo (o fenómeno) geográfico particular en la BDG, se usarán dos componentes:

- Entidad (componente descriptivo no gráfico)
- Representación geométrica (componente gráfico)

3.1 Entidad

En general, una entidad es una cosa (objeto, persona, evento, concepto, etc.) distinguible de lo que le rodea, acerca de la cual se requiere información.

Para propósitos de la BDG, una **entidad** es la representación digital del componente descriptivo de un rasgo geográfico. Se le asocia un nombre con el fin de distinguirla de otras entidades (ejemplos: *carretera*, *presa*, *línea de transmisión*, *eje estructural*, *área agrícola*, etc.).

Las entidades están descritas en los diccionarios de datos, de acuerdo con el tema y la escala.

Un **atributo** es una característica que califica y describe un aspecto de una entidad. El número de atributos asociados con cada entidad es variable. Los diccionarios relacionan los atributos de cada entidad. Por

ejemplo, los atributos para la entidad topográfica *carretera*, serían: revestimiento, número de carriles, derecho de tránsito, jurisdicción y condición. Para la entidad geológica *eje estructural*, los atributos serían: *tipo, inclinación y sentido*.

Para representar un rasgo geográfico, se especifica un valor para cada atributo de la entidad. Un **valor de atributo** es una cantidad o cualidad específica.

El conjunto de los valores que describen un rasgo geográfico particular, constituye una **ocurrencia de la entidad**. Por ejemplo, para una ocurrencia de la entidad topográfica *carretera*, los valores de los atributos son: *pavimentada, dos carriles, libre, federal y en operación*. Para una ocurrencia de la entidad geológica *eje estructural*, los valores de los atributos son: *anticlinal, recostado y derecha*.

Con el propósito de distinguir las diferentes ocurrencias de una entidad en un conjunto de datos, se usa el atributo llamado **identificador**. El valor del identificador está dado por un número secuencial que se incrementa con cada ocurrencia.

3.2 Representación geométrica

Constituye la representación digital del componente espacial de un rasgo geográfico. Los conceptos definidos para entidad: atributo, ocurrencia e identificador, se aplican también para la representación geométrica. La BDG sustenta tres tipos diferentes de representación geométrica: punto, línea y área. Cada entidad puede estar asociada con distintos tipos de representación geométrica. Por ejemplo, una *subestación eléctrica* puede estar representada ya sea como punto, o como área, dependiendo de sus dimensiones y de la escala. Cada tipo de representación geométrica (punto, línea y área) estará definida en dos dimensiones (X, Y) (Ver sección 5.4.3). La definición de la altimetría se da por la asociación de una elevación como atributo, sólo para entidades específicas, por ejemplo: curva de nivel.

3.2.1 Punto

Es la representación geométrica constituida sólo por un par de coordenadas (X, Y) (Ver figura 1). Un punto se usa para describir geoméricamente un rasgo geográfico considerado como puntual para propósitos de la BDG. Junto con sus coordenadas, el punto puede requerir una dirección, sin embargo, esta dirección estaría considerada en la entidad como un atributo.



Figura 1: Punto

Un punto posee entre otros atributos, un calificador de posición. Los valores permitidos para los calificadores de posición de puntos son (ver: "Acerca de los diccionarios de datos vectoriales", en la sección de representación geométrica):

- Definida
- Aproximada

3.2.2 Línea

Es la representación geométrica constituida por una serie de dos o más pares distintos de coordenadas (vértices) ligados secuencialmente (ver figura 2). Una línea se usa para describir total o parcialmente la geometría de un rasgo geográfico, considerado como línea, para propósitos de la BDG. También se usa la línea para delimitar áreas. Junto con sus coordenadas, la línea puede requerir un sentido al cual está asociada una característica del rasgo geográfico, sin embargo, este sentido estaría considerado en la entidad como un atributo.

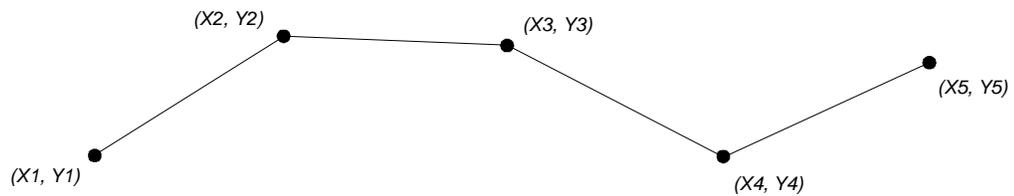


Figura 2: Línea

Una línea posee entre otros atributos, un calificador de posición. Los valores permitidos para los calificadores de posición de líneas son (ver: "Acerca de los diccionarios de datos vectoriales", en la sección de representación geométrica):

- Definida
- Aproximada
- Virtual

3.2.3 Área

Es la representación geométrica delimitada por una línea cerrada o serie de líneas que cierran. Un área se usa para describir geoméricamente un rasgo geográfico considerado como una extensión o superficie. Un área puede ser simple o compleja (ver figura 3). Un área compleja está constituida por líneas inclusivas y exclusivas.

Los calificadores de posición no se aplican al área, pero sí a las líneas asociadas con ellas. De esta manera, un cuerpo de agua puede estar delimitado por líneas definidas e indefinidas.

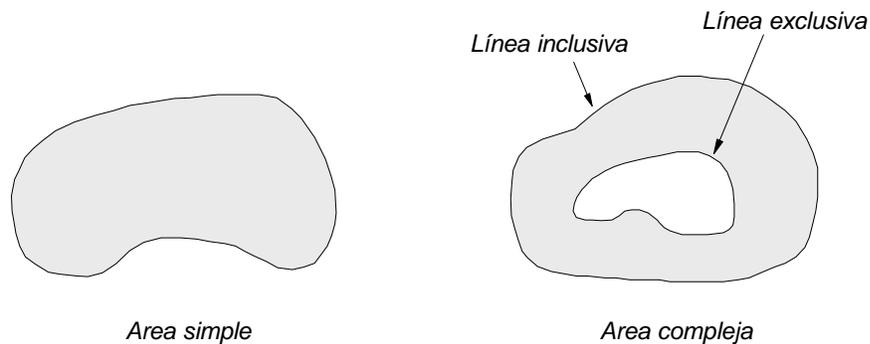


Figura 3: Tipos de área

Una área puede ser adyacente o estar sobrepuesta a otras. La figura 4 muestra la adyacencia entre áreas que tienen o no una línea común compartida, indicada ésta por el símbolo ©(ver sección 4.1.2). La figura 5 ilustra áreas sobrepuestas que no tienen relación de compartir.

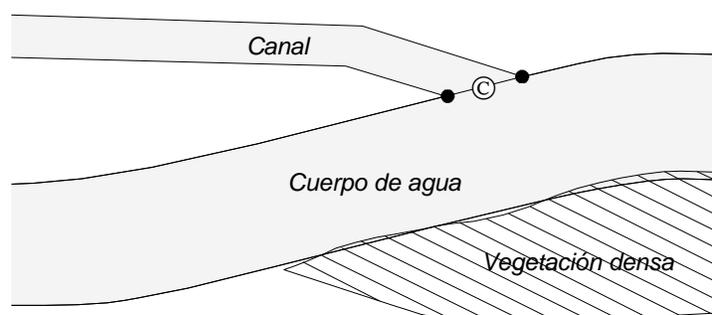


Figura 4: Adyacencia de áreas con y sin relación de compartir

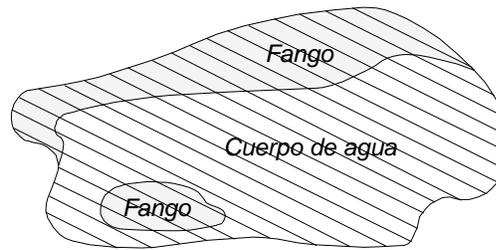


Figura 5: Sobreposición de áreas sin relación de compartir

Se indica con una nota en los diccionarios de datos para aquellas entidades que pueden estar sobrepuestas a otras. De otra manera las entidades de área serán adyacentes entre ellas

4. Estructura espacial de los datos

4.1 Relaciones

Una **relación** es una asociación entre entidades.

Los datos vectoriales deben estar libres de cualquier inconsistencia espacial, tales como:

- Excesos o defectos en las uniones de líneas o de puntos con líneas
- Contornos de áreas no cerrados.

Para garantizar lo anterior, se definen dos tipos de relación: **conectar** y **compartir**.

4.1.1 Conectar

Se da una relación de **conectar** entre ocurrencias de entidades, cuando y sólo cuando se satisfacen las dos condiciones siguientes:

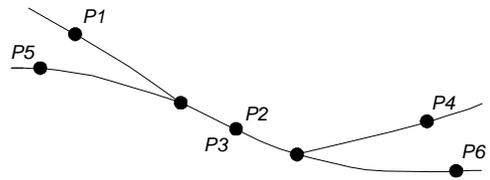
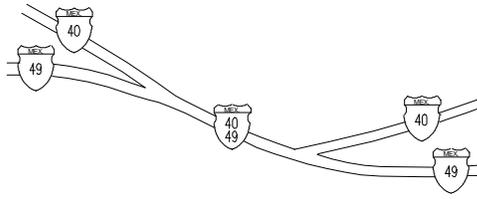
- Que exista una intersección planimétrica entre los rasgos geográficos involucrados.
- Que la BDG sustente esta relación.

Esta relación requiere que las ocurrencias de las entidades involucradas en la relación de conectar, tengan las mismas coordenadas de sus ocurrencias de representación geométrica en el punto de conexión.

Una relación de conectar implica una terminación de las ocurrencias de representación geométrica de todas las ocurrencias de entidades presentes en el punto de conexión (ver figuras 6a a la 6f).

Todas las relaciones de conectar entre entidades se encuentran indicadas en los diccionarios de datos.

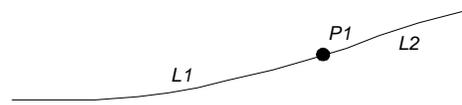
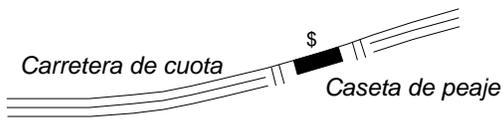
Datos en la BDG:



P1 a P6: Números de carretera
(P2 y P3 conectados)

Figura 6 a): Conectar (punto y punto)

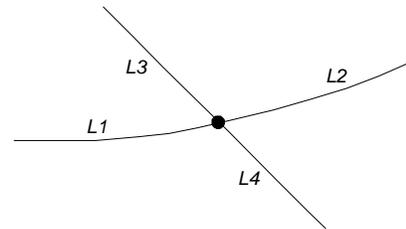
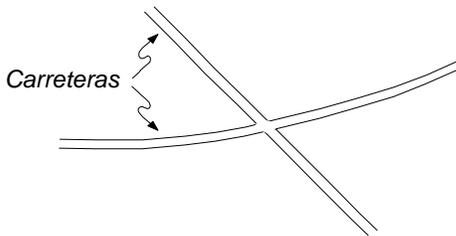
Datos en la BDG:



L1, L2: Carretera
P1: Caseta de peaje

Figura 6 b): Conectar (punto y línea)

Datos en la BDG:



L1, L2, L3, L4: Carretera

Figura 6 c): Conectar (línea y línea)

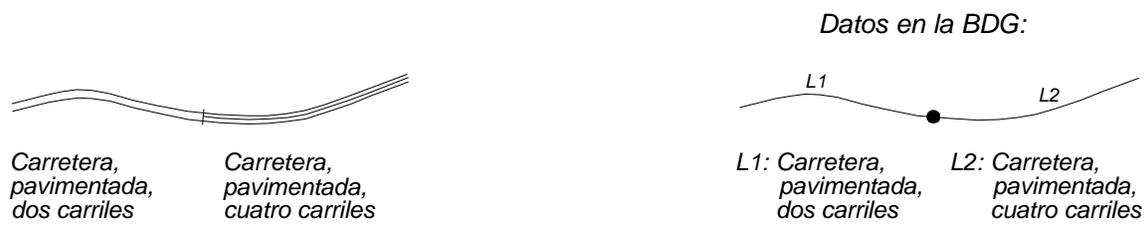


Figura 6 d): Conectar (línea y línea)



Figura 6 e): Conectar (línea y línea)

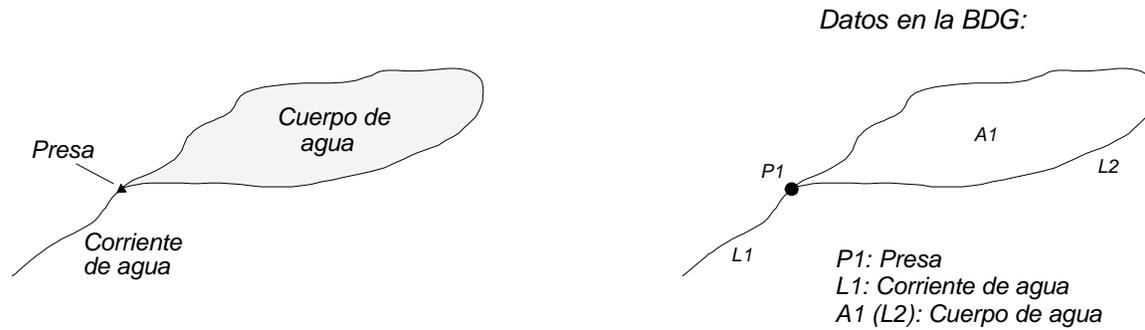


Figura 6 f): Conectar (punto, línea y área)

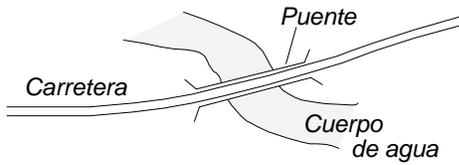
4.1.2 Compartir

Se da una relación de **compartir** entre ocurrencias de entidades cuando y sólo cuando se satisfacen las dos condiciones siguientes:

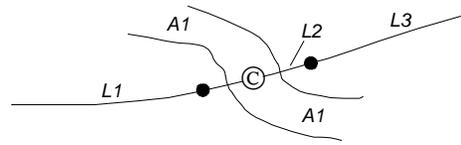
- Que los rasgos geográficos percibidos como líneas o áreas sean parcial o totalmente contiguos o coincidentes.
- Que la BDG sustente esta relación.

Esta relación requiere que las ocurrencias de las entidades involucradas en la relación de compartir, tengan las mismas coordenadas en sus ocurrencias de representación geométrica lineal.

Una relación de compartir implica conexión para todas las líneas que terminen en los extremos de la línea compartida (ver figuras 7a a la 7e).

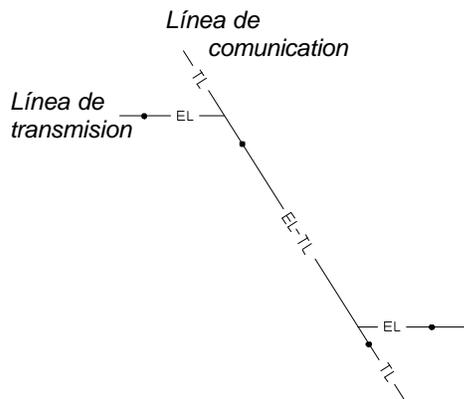


Datos en la BDG:

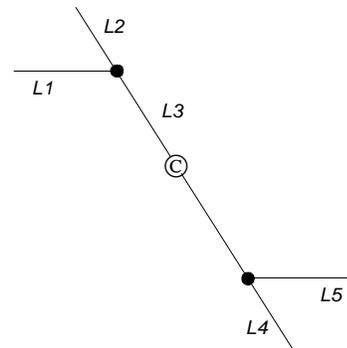


L1, L2, L3: Carretera
L2: Puente
A1: Cuerpo de agua

Figura 7 a): Compartir (línea y línea)

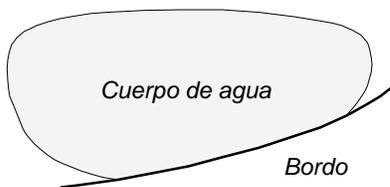


Datos en la BDG:

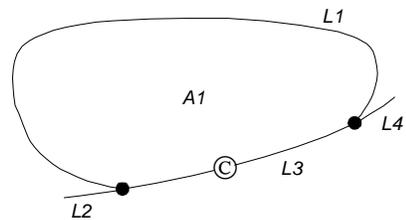


L1, L3, L5: Línea de transmisión
L2, L3, L4: Línea de comunicación

Figura 7 b): Compartir (línea y línea)



Datos en la BDG:



L2, L3, L4: Bordo
A1 (L1, L3): Cuerpo de agua

Figura 7 c): Compartir (línea y área)

Datos en la BDG:

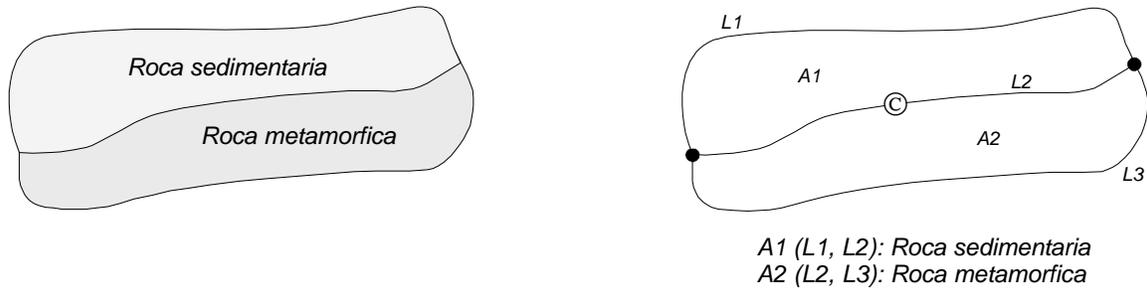


Figura 7 d): Compartir (área y área)

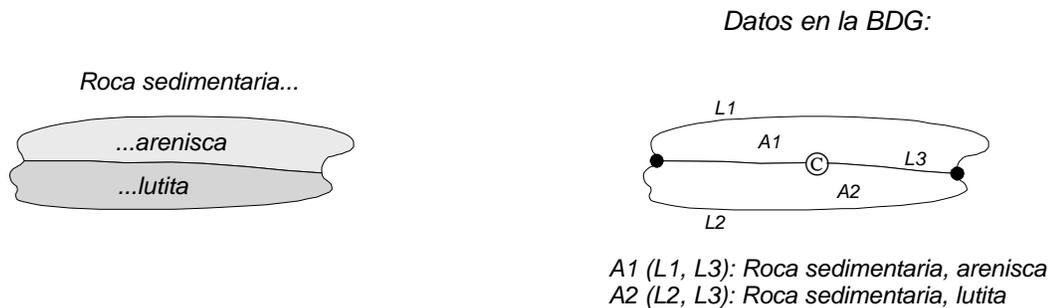


Figura 7 e): Compartir (área y área)

Las palabras "contiguo" y "coincidente" se usan en un sentido amplio: se permite una cierta tolerancia. Por lo tanto, los rasgos geográficos localizados aparentemente en la misma posición planimétrica, sin exceder una tolerancia, y para los cuales la BDG sustenta una relación de compartir, deben ser parte de tal relación para cada una de las entidades representadas.

Como se mencionó en el punto 3.2.3, en general las áreas son adyacentes. Para saber si las áreas involucradas están sobrepuestas, referirse a las respectivas entidades, en los diccionarios de datos.

Todas las relaciones de compartir entre entidades se encuentran indicadas en los diccionarios de datos.

4.2 Integración de conjuntos de datos

En el límite común (límite teórico) de conjuntos de datos de precisión equivalente, la continuidad espacial de rasgos geográficos se asegura cuando y sólo cuando se satisfacen las siguientes dos condiciones:

- Las ocurrencias de las entidades que representan el mismo rasgo geográfico, se encuentran en cada uno de los conjuntos de datos adyacentes.
- Las ocurrencias de línea involucradas (una en cada uno de los conjuntos de datos adyacentes) están localizadas dentro de un intervalo de tolerancia predeterminado a lo largo del límite teórico de los conjuntos de datos.

Para una entidad que reúna las condiciones arriba mencionadas, las ocurrencias de representación geométrica lineal deben segmentarse en el punto donde coinciden con el límite teórico de los conjuntos de datos. Las coordenadas del punto de continuidad en el límite teórico deben duplicarse en cada ocurrencia de línea involucrada (un par de coordenadas por cada conjunto de datos).

Cualquier área que representa una entidad que está sobrepuesta al límite teórico de conjuntos de datos adyacentes, debe estar matemáticamente cerrada y separada por una ocurrencia de línea en el límite. En este caso, la ocurrencia de línea se duplica en cada conjunto de datos adyacente y debe tener un "calificador de representación geométrica" (calificador de línea), "virtual" (ver Línea en "Acerca de los diccionarios de datos vectoriales", en la sección de representación geométrica).

Cada una de estas líneas está asociada con un área que representa una entidad. La figura 12 muestra un rasgo geográfico de área (entidad superficial), que se extiende a través de más de un conjunto de datos.

En todos los casos, los valores de las coordenadas en el límite teórico de los conjuntos de datos se redondean al valor más próximo de acuerdo con la resolución establecida (ver sección 5.2).

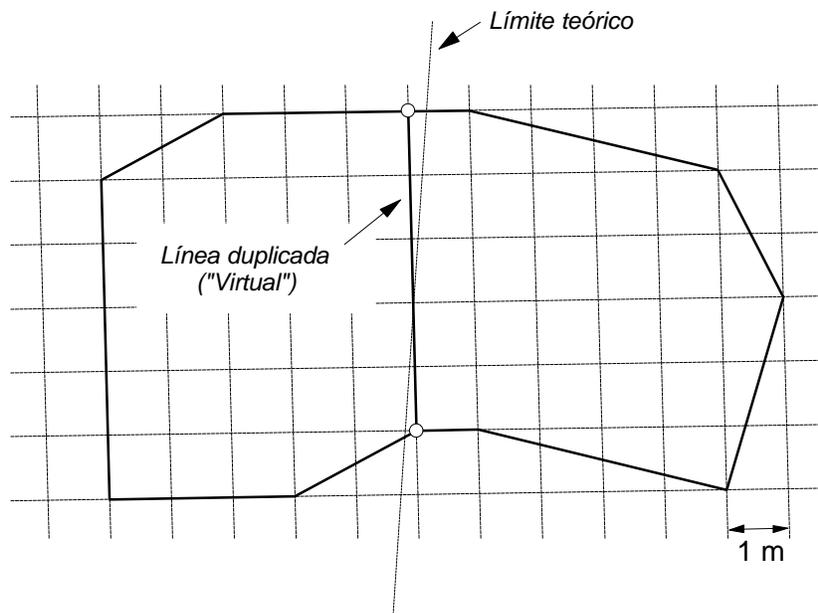


Figura 12: Cierre de áreas en el límite de conjunto de datos

5. Aspectos generales

Esta sección trata con varios aspectos relacionados con los datos: precisión, resolución, tolerancias, sistema de referencia, proyecciones cartográficas, sistemas de coordenadas y límites de los conjuntos de datos.

5.1 Precisión

La **precisión** de los datos de representación geométrica está dada por la diferencia entre la posición de la representación geométrica asociada con una entidad, y la posición real del rasgo geográfico correspondiente, medido con respecto a la red geodésica.

En los datos auxiliares de los conjuntos de datos, se indica la precisión del método de compilación original de los datos topográficos de la BDG. Asimismo, se mantendrá un registro de los métodos y materiales empleadas en los diferentes procesos lo cual permitirá tener una idea de la calidad de los datos.

5.2 Resolución

La **resolución** especifica la unidad de medida más pequeña que se adopta para registrar datos. Para los datos vectoriales de la BDG, se establece en un metro para los ejes X e Y.

Para representaciones geométricas lineales, la densidad de coordenadas debe ser suficiente para permitir curvas suaves a la escala de representación (1:50 000, 1:250 000, 1:1 000 000, etc.), mientras se respete la precisión y se evite una sobreabundancia de coordenadas. Las especificaciones y estándares de adquisición de datos definen los parámetros mínimos con respecto a la regla arriba mencionada.

5.3 Tolerancias

Dos rasgos geográficos deben ocupar la misma posición planimétrica para tener una relación de compartir (ver sección 4.1.2), la contigüidad o coincidencia entre dos rasgos deber ser exacta o estar dentro de una tolerancia predeterminada (de acuerdo con las especificaciones de adquisición de datos). Esta tolerancia permite considerar que rasgos geográficos que coinciden aparentemente en el material fuente, compartan exactamente las mismas coordenadas. Esta consideración cartográfica depende (cuando se aplique), de la resolución y escala.

Además, en el límite común de conjuntos de datos de precisión equivalente, la continuidad espacial de rasgos geográficos debe asegurarse con satisfacción de las condiciones de prerrequisito (ver sección 4.3). Una de ellas involucra una tolerancia predeterminada que no puede ser excedida a lo largo del límite teórico. En todos los casos estas tolerancias están consideradas en las especificaciones para adquisición de datos.

5.4 Sistema de referencia, proyección cartográfica y sistemas de coordenadas

5.4.1. Sistema geodésico de referencia

Se usa como sistema de referencia para las coordenadas planimétricas, el sistema ITRF92, época 1988.0, definido por el Servicio Internacional de Rotación de la Tierra.

5.4.2. Sistema de alturas y datum vertical

Las elevaciones son ortométricas y están referidas al Nivel Medio del Mar. Se basan en el Datum Vertical para Norteamérica de 1929.

5.4.3. Sistema de coordenadas horizontales y proyección cartográfica

Las coordenadas X (este) y Y (norte) para puntos y líneas están expresadas de acuerdo con la proyección Universal Transversa de Mercator (UTM), en las escalas 1:50 000 y 1:250 000, y con la proyección Cónica Conforme de Lambert con paralelos base 17°30' y 29°30', para las escalas 1:1 000 000 y 1:4 000 000. Para esta proyección se considera como falso origen el punto de coordenadas 12° de latitud norte y 102° de longitud oeste y con coordenadas de 2 500 000 m en X y 0 m en Y.

5.5 Límites de conjuntos de datos

Los límites de conjuntos de datos coinciden en general con la división en hojas del Sistema Nacional de Información Geográfica, la cual se basa en las subdivisiones del sistema de numeración de la hoja del Mapa Mundial Internacional (MMI), a escala 1:1 000 000. Sin embargo, para la BDG, los conjuntos de datos que corresponden a hojas en escalas 1:50 000, 1:250 000 y 1:1 000 000 que exceden los límites teóricos, se separan en conjuntos regulares de datos, cortados a lo largo del límite teórico. Los límites internacionales también constituyen límites de conjuntos de datos y la BDG no contiene datos fuera de estos límites.

La extensión de los conjuntos de datos se muestra en el cuadro 1.

Escala	Dimensiones (latitud por longitud)
1:50 000	Quince por veinte minutos
1:250 000	Uno por dos grados
1:1 000 000	Cuatro por seis grados
1:4 000 000	Un conjunto de datos cubre todo el país

Cuadro 1. Extensión de los conjuntos de datos

Los límites de los conjuntos de datos, llamados también "límites teóricos", se determinan de la siguiente manera:

Para cartas a las escalas 1:50 000 y 1:250 000:

LIMITES ESTE Y OESTE: Segmentos rectos que unen puntos de igual longitud a cada 5 minutos de latitud entre las esquinas Norte y Sur.

LIMITES NORTE Y SUR: Segmentos rectos que unen puntos de igual latitud a cada 5 minutos de longitud entre las esquinas Este y Oeste

NOTA: Puede aparecer una discontinuidad de un metro cuando el corte de vectores ocurre en conjuntos de datos sobre zonas UTM adyacentes.

Para cartas a la escala 1:1 000 000 y 1:4 000 000:

Los límites se determinan con los meridianos y paralelos que los delimitan; para los paralelos, se usan segmentos rectos a cada 1° para la escala 1:1 000 000 y cada 4° para la escala 1:4 000 000.

Estos límites son teóricos o imaginarios, y por lo tanto no existen en la BDG, excepto para las líneas que ostentan el calificador de representación geométrica "virtual". Esta última está localizada en el límite teórico a la resolución más cercana. Todas las ocurrencias de representación geométrica lineal que se extienden a través de los límites de un conjunto de datos deben segmentarse y deben terminar en el límite teórico, en el punto más cercano de acuerdo con la resolución (ver secciones 4.3 y 5.2).