

Interoperabilidad posicional: elemento clave para las Infraestructuras de datos espaciales

Francisco Javier Ariza López

Departamento de Ingeniería Cartográfica, Geodésica y Fotogrametría

Universidad de Jaén

Resumen

En este trabajo se presenta el concepto de interoperabilidad posicional entre conjunto de datos geográficos. En la línea de ISO 19113: Información Geográfica, Principios de la Calidad, se propone crear un nuevo elemento de la calidad de los datos espaciales denominado interoperabilidad, para el que se desarrollan algunos subelementos (interoperabilidad posicional planimétrica e interoperabilidad posicional altimétrica), y algunas medidas. Se propone una metodología general de evaluación de esta componente de la calidad de la información geográfica y se aplica para el caso de combinaciones entre conjuntos de datos vectoriales, de imágenes y modelos de elevaciones del terreno.

Palabras clave

Interoperabilidad, interoperabilidad posicional, calidad de la información geográfica.

Introducción

Las Infraestructuras de Datos Espaciales (IDE) están alcanzando una gran expansión dado el notable conjunto de beneficios que potencialmente son capaces de generar a la sociedad, tal como indican algunos estudios (EU-JRC 2006), y como nosotros mismos podemos observar, en el día a día, por nuestras propias experiencias y las de los que nos rodean. En muchos casos este desarrollo viene favorecido por marcos legales, como es el caso de la Directiva INSPIRE europea (Directiva 2007/2/CE; EU 2007), y las iniciativas administrativas y legislativas equivalentes en otros ámbitos como EEUU (FGDC 1994) y Australia, Canadá, etc.

Un concepto clave para el desarrollo de las IDE es la interoperabilidad. De una forma general, la interoperatividad se puede definir como la condición por la cual sistemas heterogéneos pueden interactuar intercambiando flujos de energía, información y/o procesos. Esta amplia definición cubre campos como los transportes, el eléctrico, etc. Desde un punto de vista más geomático, y por ello más cercano a lo informático, interoperabilidad se puede definir como la capacidad de un sistema o producto para trabajar con otros sistemas o productos sin un esfuerzo especial por parte del cliente.

De manera general, la interoperabilidad se apoya en normas o estándares, los cuáles, en el caso de la Información Geográfica (IG), han sido desarrollados por el *Open Geospatial Consortium* y por el Comité Técnico 211 de ISO, donde se destaca la familia de normas ISO 19100.

El aspecto posicional de la IG es definitorio, siempre ha sido de gran importancia en cartografía, y por ello es el elemento de la calidad de los datos espaciales más extendido y evaluado por las organizaciones productoras de cartografía (Jakobsson y Vauglin 2002). Se necesita una buena

calidad posicional cuando se van a utilizar de manera conjunta dos Conjuntos de Datos Geográficos (CDG) distintos, circunstancia que ocurre constantemente en el uso de las IDE. Comportamientos posicionales diferentes de los CDG significan la existencia de una distorsión posicional entre ellos, y una barrera a la interoperabilidad efectiva de los mismos (Church y col. 1998). Pero lo más importante es que esta barrera no es sólo posicional, y por ello geométrica y topológica, también lo es temática. Como indican numerosos estudios (p.e. McGwire 1996, Carmel y col. 2001 y 2006), el aspecto temático entre distintos CDG se ve muy influido por desajustes posicionales.

Tanto las IDE ya comentadas, como la democratización de los Sistemas de Posicionamiento Global (GPS), tanto a niveles profesionales como domésticos, han propiciado una renovada mayor atención a esta componente de la IG. Consecuencia de lo anterior es que distintos organismos productores de IG hayan o estén desarrollando proyectos de mejora de la exactitud posicional de sus productos (Rönsdorf 2004), ejemplos de ello son el *Ordnance Survey* del Reino Unido, el Instituto Geográfico Nacional de Francia y otras instituciones europeas (EUROSDR 2004).

Desde esta perspectiva de que la componente posicional es clave para la interoperabilidad, autores como Giordano y Veregin (Giordano y Veregin 1994) propusieron, con el nombre de confrontabilidad, una nueva componente indicativa de la calidad de la IG en esta línea. Se trata de una componente que, a pesar de su interés práctico y dependiente de aplicación, no ha sido estudiada, ni formalizada, ni aplicada, quizá por su complejidad o por existir otras prioridades. El concepto de confrontabilidad, tal como se recoge en la obra de dichos autores, es relativamente vago y no está plenamente formalizado. Por confrontabilidad se entiende el nivel al que es posible fundir diversos CDG de una misma área geográfica, pero provenientes de sistemas de producción distintos, o que poseen modelos y escalas/resoluciones distintas, o que tienen fechas distintas. Esta confrontabilidad debe hacer referencia a las características posicionales, temáticas, etc. La idea es obtener una medida del grado de superposición coherente entre los CDG intervenientes, en particular para el caso de dos conjuntos que se combinan y dan un tercero formado por el total o partes de los dos primeros, ya sea de carácter permanente o temporal: $CDG_1 \times CDG_2 \rightarrow CGD'$. Como es lógico pensar, no puede existir un índice único que evalúe la confrontabilidad de manera total y universal, dado que para cada par de conjuntos de datos que se proponga utilizar de manera conjunta podrán derivarse índices y valores distintos. Es más, debido a la modelización de la realidad en objetos y temas, la confrontabilidad podría establecerse por tipologías de objetos (puntuales, lineales, superficiales), por clases de objetos, por temas, etc., y evaluarse y presentarse de manera particular para cada uno de ellos o de forma conjunta o global para cada par de productos.

La confrontabilidad, o integrabilidad o interoperabilidad, es una propiedad deseable de los productos (Levinshohn 2001, Ariza 2006) puesto que facilita el uso conjunto de informaciones específicas procedentes de fuentes o procesos distintos, por ello atiende tanto a aspectos geométricos o espaciales (Church y col. 1998, Curtin y col. 1998) como temáticos u ontológicos (Bishr 1998, Fonseca y col. 1999 y 2002). En el caso de las IDE es una necesidad evidente al permitir éstas poner en común datos de distintos productores. La confrontabilidad entre productos de distintas fuentes productoras es de gran valor y no está asegurada debido a diferencias entre procesos productivos, modelos, etc. y a que no existen normas generales de aplicación. Pero si esta interoperabilidad entre productos de distintos productores es interesante, el conseguirlo entre distintos productos generados por una misma organización (p.e. versiones distintas de un mismo producto, productos distintos como bases vectoriales y ortofotografías), puede entenderse como un aspecto aún más importante, y como una

exigencia de coherencia interna de la producción de la propia organización. Este problema nos ha sido apuntado numerosas veces por los usuarios de IDE.

Los orígenes de la falta de interoperabilidad posicional pueden ser muchos y diversos, entre otros:

- ¬ Problemas del marco geométrico/geodésico que origine desplazamientos, escalados y giros entre los conjuntos de datos.
- ¬ Uso de distintas fuentes y fechas (p.e. vuelos distintos para una restitución).
- ¬ Uso de criterios de captura distintos.
- ¬ Uso de modelos distintos (clasificaciones, escalas de medida, atributos, relaciones, etc.).
- ¬ Intervención de procesos cartográficos distintos (p.e. la generalización).

Por ello, según Giordano y Veregin (1994), entre los factores a considerar para evaluar el grado de confrontabilidad estarían:

- ¬ Exactitud posicional.
- ¬ Exactitud temática.
- ¬ Niveles de resolución posicional, temática y temporal de los datos.
- ¬ Grado de generalización empleado para construir la CDG.
- ¬ Escala cartográfica.
- ¬ Proyección utilizada.
- ¬ Modelo, estructura y formato de los datos.
- ¬ Escala de medida utilizada para los datos temáticos (nominal, ordinal, intervalo/relación).

Hasta ahora hemos presentado la confrontabilidad como el aspecto relativo al grado o evaluación de la interoperabilidad posicional, pero ligado a este concepto y su problemática también se debe hablar de sus soluciones. En este caso el concepto clave es el de conflación. Con él nos referimos a los procesos de integración y combinación de conjuntos de datos. Definiciones como las de Edwards y col. (2002) o Cobb y col. (1998) se refieren al mismo como la acción de unificar conjuntos de datos en otro que resulte “mejor” que los iniciales, lo cual podría deberse simplemente a la agregación de datos. Sin embargo, hay otras definiciones que sí vienen a indicar ya la necesidad de actuar sobre los elementos para conseguir la interoperabilidad (Casado 2006), y cuyos pasos metodológicos fundamentales serían, primeramente un emparejamiento de elementos entre los CDG, y, posteriormente, la aplicación de unas transformaciones para llevar un conjunto al otro minimizando las discrepancias de forma, tamaño e imposición espacial, y permitan hacer perfectamente “superponibles” los CDG. Los procesos de conflación se clasifican según afecten a los aspectos geométricos, topológicos o semánticos, según la relación entre los conjuntos de datos (horizontal, vertical, temporal), etc. (ver Lendínez y Ariza 2008 para más detalle sobre estos procesos).

En este análisis se atenderá sólo y exclusivamente al aspecto posicional pues, aún siendo un aspecto parcial de la interoperabilidad, es un factor clave y evidente para conseguirla (Church y col. 1998). Por ello, atendiendo al problema de interoperabilidad posicional expuesto, y sobre la convicción de que el desplazamiento relativo es el que origina la mayor parte de los problemas en la integración entre CDG, este documento se plantea como objetivo formalizar dicho problema con ejemplos concretos y proponer metodologías de evaluación de la interoperabilidad posicional. Se trata pues de desarrollar métodos de evaluación o cuantificación que podrán ser utilizados, si se han establecido especificaciones, para decidir sobre la conveniencia o no del uso conjunto de dos o más CDG, pero que también pueden servir para valorar el “coste” de la conflación geométrica o para guiar las estrategias de conflación, la selección y/o parametrización de los algoritmos a aplicar. Es decir, si

aplicáramos un modelo equivalente al Modelo Conceptual Comprensivo desarrollado por McMaster y Shea (1992) para la generalización, estaríamos hablando de la componente de evaluación cartométrica propiamente dicha y previa a la resolución del problema.

Junto a esta introducción, el presente documento se articula entorno a tres apartados más. El primero de ellos establece los aspectos formales, para lo que se sigue la estructuración de ISO 19113 en elementos y subelementos de la calidad. Posteriormente se presenta una metodología general de evaluación y su concreción para ciertas combinaciones de tipologías distintas de CDG. Finalmente, se incluyen unas conclusiones del trabajo.

Formalización

En este apartado se pretende realizar la formalización de esta componente de la calidad de la IG. Para ello, en primer lugar, se procederá a concretar el objeto y alcance que nos proponemos en este trabajo, luego se procederá, en el sentido de ISO 19113, a la definición de los elementos/subelementos que la formalizan y, finalmente, se propondrán algunas medidas para evaluarla.

Como ya se ha indicado, este trabajo se centra en la presentación del “problema geométrico de interoperabilidad” que pueden presentar dos CDG al intentar ponerlos a trabajar de manera conjunta, y en proponer herramientas o métodos de análisis. No se atenderán aspecto topológicos, semánticos o de otra índole distinta a la puramente posicional. En todo caso, se trata de un paso previo a la propia confluencia y que se orienta a la evaluación del alcance del problema, para lo cual se proponen medidas cuantitativas y metodologías de análisis.

El problema geométrico que se evalúa en este trabajo es del tipo “vertical”, es decir, el que ocurre en la superposición de conjuntos de datos. Éste posiblemente es el más común dado que es muy usual intentar interoperar con diversos productos de una misma zona geográfica. El caso horizontal, p.e. al intentar interoperar con dos CDGs de dos regiones o países vecinos, no será objeto de análisis.

En el proceso de confluencia es importante distinguir que existen elementos encajantes (p.e. caminos comunes) y encajados (p.e. construcciones, parcelario, etc., según el caso). Así, es común que en el proceso $CDG_1 \times CDG_2 \rightarrow CGD'$ se consideren subconjuntos de elementos comunes como elementos encajantes y que los elementos encajados del CDG_2 se tomen para enriquecer al CDG_1 generando un nuevo conjunto CGD' . Por ello, los elementos encajantes son aquellos que se encuentran en los dos CDG y que se utilizan para definir la transformación de un espacio geométrico a otro. Éstos permiten evaluar a priori la interoperabilidad y efectuar la confluencia. Los elementos encajados son los que no tienen homólogos y son transformados a partir de la transformación determinada para los elementos encajantes, por ello son los que pueden suministrarse, a posteriori tras la confluencia, las medidas más representativas del grado de interoperabilidad real alcanzado.

Dado que se trata de un problema de tipo geométrico debido a desfases relativos en la posición, se podría pensar en abordar el asunto a partir de algunas de las metodologías de control posicional ya existentes (ver Ariza y col. 2006, Ariza y Atkinson 2008 para una revisión reciente sobre algunos de estos métodos). Las metodologías de control posicional pueden ser aplicadas como herramientas de cuantificación del aspecto posicional pero no son suficientes, para este caso deben acompañarse de otras medidas que den una idea mejor del problema que se encara. Por ejemplo, saber que hay 5m de discrepancia promedio entre dos

conjuntos de datos viene a ser igual a saber que hay 500m, pero lo que es realmente crítico para entender el problema de confluencia es que esos metros de discrepancia relativa, bien 5m ó 500m, supongan además un comportamiento muy neto (Figura 1.a), o un comportamiento más aleatorizado y localizado por zonas (Figura 1.b). Las transformaciones geométricas posteriores serán tanto más sencillas y exitosas si existe un desplazamiento o sesgo claro y cuantificable, que si existe una situación que se pueda considerar como un campo aleatorio, o de ruido, y donde, según el método aplicado, no será posible una solución que genere resultados adecuados, o resultará más costoso conseguirla. De ésta forma, las técnicas como las de ajuste global polinomial por mínimos cuadrados sobre el conjunto de datos homólogo, que son una opción clásica, deben ir dejando paso a otras soluciones más complejas, basadas en trasformaciones más locales, para conseguir mejores ajustes, como son las de elementos finitos.

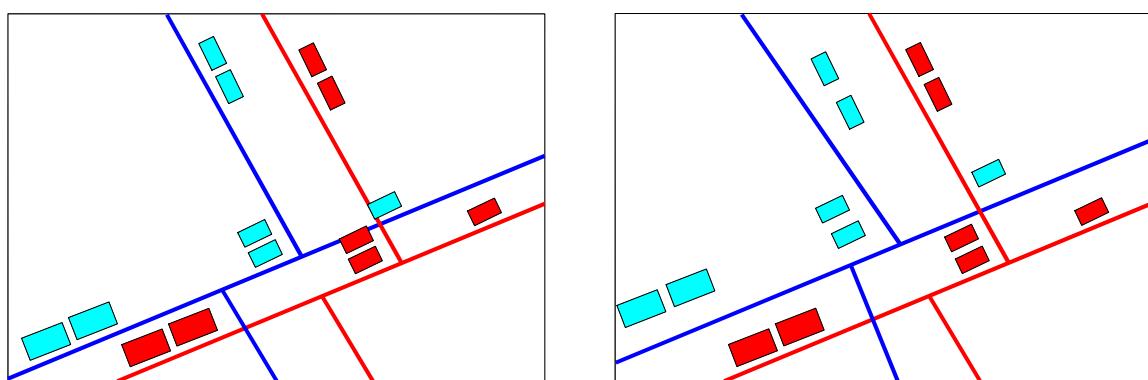


Figura 1.- Ejemplo de dos comportamientos posicionales distintos elementos encajanates de dos conjuntos de datos geográficos: a) desplazamiento neto, b) desplazamiento aleatorizado de sus elementos

El ejemplo anterior apunta sobre el carácter concreto y utilitario que necesariamente debe tener la evaluación de la interoperabilidad posicional. Por un lado debe estar necesariamente ligada a dos CDG (concreto), y por otro, debería servir para alumbrar sobre la estrategia de confluencia entre ambos conjuntos (útil). Este doble carácter es el que debe guiar la formalización de esta componente en cuanto a la definición como elemento/subelemento, en el sentido de ISO 19113, y la proposición de medidas evaluadoras.

Pasando ya a formalizar la interoperabilidad posicional, y asumiendo el marco que establece la norma ISO 19113: Información Geográfica, Principios de la calidad, lo primero que debemos hacer es decidir si este aspecto debe considerarse como un elemento nuevo e independiente, o como un subelemento de los ya existentes (Tabla 1), en cuyo caso se debería decidir de cuál de ellos depende.

Tabla 1.- Elementos y subelementos propuestos en ISO 19113	
Elemento	Subelementos
Compleción	Omisión Comisión
Consistencia lógica	Conceptual Dominio Formato Topológica
Exactitud posicional	Absoluta Relativa Malla

Exactitud temporal	Exactitud medida tiempo Consistencia temporal Validez temporal
Exactitud temática	Corrección de la clasificación Corrección de los atributos cualitativos Exactitud de los atributos cuantitativos

Dado su carácter posicional, parece lógico y directo hacer una primera vinculación con el elemento exactitud posicional. Sin embargo, y a modo de ejemplo, también se puede pensar en que una falta de interoperabilidad posicional genera problemas de consistencia lógica en el conjunto resultante de la integración.

Centrándonos en lo posicional, por parecer más evidente, la justificación para desarrollarla como un subelemento dentro de la exactitud posicional se motiva en su clara vinculación con esa componente de la calidad de la IG: la interoperabilidad posicional es una exactitud posicional relativa entre productos. Sin embargo nuestra propuesta es considerarla como un nuevo elemento con posibles subelementos. Lo proponemos por tres motivos fundamentales:

- Significado distinto. La exactitud posicional absoluta se refiere a un CDG y un marco, y la exactitud posicional relativa a relaciones entre los elementos de ese mismo CDG dentro de ese mismo marco. La interoperabilidad posicional necesita de ese marco pero las relaciones no son entre elementos del mismo conjunto de datos, necesariamente debe haber dos o más CDG para considerar la interoperabilidad posicional. Todo ello no implica que no puedan existir relaciones funcionales de paso entre interoperabilidad posicional y exactitud posicional absoluta/relativa entre dos CDG (Figura 2), pero este campo no ha sido estudiado.

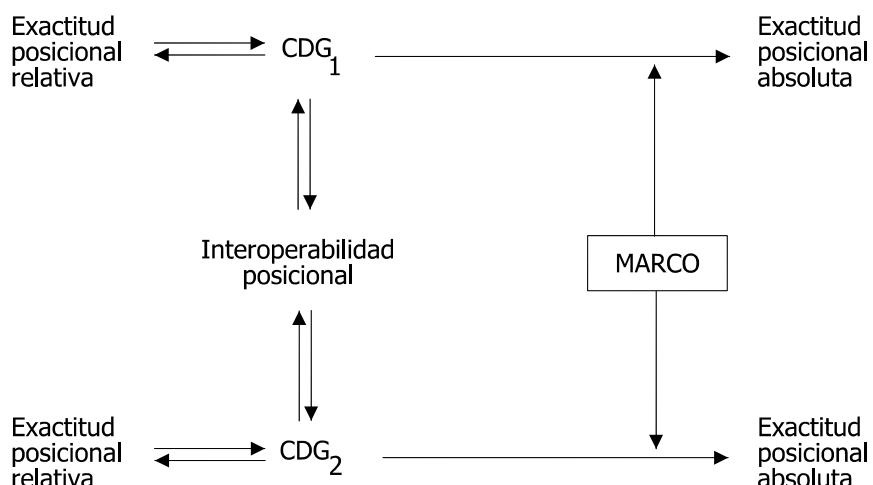


Figura 2.- Relación entre la interoperabilidad posicional y la exactitud posicional

- Objeto distinto: La exactitud posicional absoluta viene a indicar claramente la interoperabilidad posicional de un CDG respecto a un marco general o absoluto que, en teoría, se supone válido para otros CDG. La exactitud posicional relativa evalúa la interoperabilidad posicional entre los elementos propios de un único CDG. Por el contrario, en el caso de la interoperabilidad posicional uno de los conjuntos funciona como marco y no importa si realmente su comportamiento posicional absoluto ni relativo.
- Clarifica conceptualmente la situación: Al separar propiedades intrínsecas del conjunto de datos, con propiedades relacionadas con comportamientos entre conjuntos de datos se

clarifica la existencia de un problema distinto. Esta separación además permite proponer un ámbito nuevo de elementos orientados a la interoperabilidad entre conjuntos de datos, aspecto que consideramos de gran relevancia conceptual y aplicada.

Por ello, con esta pretensión de una mayor claridad conceptual se propone que la interoperabilidad sea considerada como un nuevo elemento de la calidad de la IG. Este elemento quedaría definido de la siguiente manera:

- ¬ Interoperabilidad: Elemento que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos.

Que para nuestro caso de interés quedaría desplegada por un subelemento de base posicional definido como:

- ¬ Interoperabilidad posicional: Subelemento de la interoperabilidad que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos debido a sus distintos comportamientos posicionales relativos.

Pudiéndose desarrollar, a partir de éste último, otros subelementos más específicos para casos de interés concretos, como por ejemplo son:

- ¬ Interoperabilidad posicional planimétrica: Subelemento de la interoperabilidad que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos debido a sus distintos comportamientos posicionales planimétricos relativos.
- ¬ Interoperabilidad posicional altimétrica: Subelemento de la interoperabilidad que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos debido a sus distintos comportamientos posicionales altimétricos relativos.

Como ya se ha indicado, la interoperabilidad posicional tiene implicaciones sobre otros aspectos de la IG, por lo que se podrían definir nuevos subelementos que las tengan en consideración, como son, por ejemplo:

- ¬ Interoperabilidad posicional topológica. Subelemento de la interoperabilidad que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos debido a problemas topológicos surgidos por la combinación de los CDG (p.e. número de polígonos astilla).
- ¬ Interoperabilidad posicional conceptual. Subelemento de la interoperabilidad que cuantifica y describe la mayor o menor facilidad de interoperar que tienen dos conjuntos de datos geográficos debido a los problemas que surgen en el modelo conceptual por la combinación de los CDG (p.e. superposiciones no válidas, etc.).

Otros aspectos importantes de la formalización que pretendemos realizar atienden a los métodos de evaluación y a las medidas cuantificadoras que se pueden aplicar.

Según ISO 19114, para la evaluación de la calidad se pueden aplicar métodos directos e indirectos. Los métodos directos son aquellos que se basan en la comparación o medida y se dividen en internos y externos. En el primer caso sólo se utilizan datos del propio CDG, es el caso de comprobaciones como la de consistencia lógica de carácter topológico. Los métodos externos necesitan fuentes externas al CDG, como por ejemplo, para el caso de un topónimo, la comprobación otra fuente más exacta o la salida a campo para consultar a los lugareños. Dentro de esta alternativa de métodos directos existen procesos plenamente automatizables, que permiten lo que se denomina una inspección o control al 100%. No obstante son muchos los elementos de la calidad que requieren procesos de control con un desarrollo manual, lo

cual lleva al uso de técnicas de muestreo o inspección como las apuntadas en las normas ISO 2859-1 e ISO 3159-1.

Los métodos indirectos se basan en estimaciones e informaciones relacionadas con el CDG pero con fuentes distintas a los propios datos del CDG bajo consideración. En este caso no se realiza medición o cuantificación alguna. Los métodos indirectos pueden estar al alcance de los usuarios avezados, dado que en muchos casos se trata de interpretaciones basadas en la experiencia. Si los metadatos son adecuados darán buena información. En este sentido, el linaje y otras informaciones que presenten ejemplos de uso, o los productos derivados del producto que se referencia, serán datos de gran valor, y de ello se encarga la norma ISO 19115.

Como puede deducirse de lo indicado, y siguiendo la clasificación propuesta por ISO 19114, los métodos que se deben aplicar para evaluar la interoperabilidad posicional son del tipo directo y externo; si bien este caso tiene matices que lo distinguen de los procesos de evaluación de la calidad propiamente dichos.

El siguiente paso consiste en definir y proponer un conjunto de medidas que permitan cuantificar el grado de interoperabilidad posicional existente. Estas medidas deben seleccionarse con criterio para que resulten significativas de la magnitud del problema y de ayuda a los algoritmos de confluencia.

La norma ISO 19138 se dedica a formalizar cómo se deben establecer medidas de la calidad y, además, concreta un conjunto de 73 medidas para la evaluación de los distintos subelementos propuestos en la norma ISO 19113. Por ello consideramos que debe ser el primer punto de partida. Sin embargo, hay muchas más medidas propuestas en la literatura y que podrían ser utilizadas según su conveniencia. A modo de ejemplo, se puede destacar el Diccionario de Distancias (Deza y Deza 2006) que es una copiosa recopilación en la que se pueden encontrar numerosas definiciones de distancias utilizables como medidas evaluadoras en el ámbito de la IG.

Un aspecto importante de esta norma es que define lo que denomina “medidas básicas” de la calidad del dato. Éstas se introducen para evitar la repetición en la definición de conceptos y evitar así problemas y ambigüedades. De esta forma sirven como base para el desarrollo de todas las medidas concretas posibles que se pueden utilizar. Se relacionan con dos formas de trabajo muy propias del control de calidad como son:

- ¬ **Conteo de errores:** Se expresan mediante medidas de conteo. Son aquellas que cuentan el número de errores (defectos) o casos correctos. Son adecuadas para los aspectos de la calidad en los que la medida es el conteo de ocurrencias de una circunstancia (error o caso correcto).
- ¬ **Estimación de la incertidumbre:** Se expresa mediante medidas de incertidumbre. Son aquellas que se basan en modelos estadísticos sobre la incertidumbre en las medidas y son adecuadas para los aspectos medibles como el posicional, abarcando casuísticas 1D, 2D y 3D, así como diversos niveles de confianza.

De esta forma, y en relación a la interoperabilidad posicional entre CDG se pueden proponer múltiples alternativas de conteo como el número de elementos que presentan problemas de violación (p.e. superposición, polígonos astilla, etc.); y que todas ellas pertenecerían a la tipología de conteo de errores. Pero también se pueden proponer medidas de incertidumbre, tradicionalmente más ligadas a la componente posicional.

Por ello, desde un punto de vista aplicado al caso y a orientar las estrategias de confluencia, estas dos opciones pueden considerarse como:

- Medidas de esfuerzo. Son medidas que dan idea del esfuerzo que conllevaría la integración de los dos CDG. Este esfuerzo podría venir dado por el número de elementos que deberían ser modificados (geometría, topología, etc.). De aquí se podrían derivar unas estimaciones de los tiempos (costes) computacional y manual necesarios arreglarlos y así lograr la confluencia efectiva. En algún caso estas medidas pueden desarrollarse a partir de metodologías de análisis interpretativo puesto que el la combinación de dos CDG pueden darse situaciones en las que sea difícil derivar una cuantificación clara.
- Medidas de posición: Suponen una perspectiva más ligada a la posición, planteamiento que enraíza y se justifica plenamente en el hecho de que, con independencia de lo que ocurra en lo topológico, temático, etc., el comportamiento posicional de un conjunto de datos respecto a otro va a generar problemas debido a desplazamientos relativos. Para este caso el número de medidas que existen actualmente propuestas en la literatura, no necesariamente para este problema, es innumerable. Estas medidas estarán basadas en la “distancia” o discrepancia posicional entre elementos homólogos de los CDG que se analizan. En cuanto tales, son medidas que resultan más familiares que las basadas en el esfuerzo, aunque, tal vez, sean menos intuitivas respecto a lo que significa arreglar el problema.

Un aspecto crítico de las evaluaciones de la calidad de la IG es que, por lo general, se trabaja con valores globales para un CDG. La IG tiene, por definición, variabilidad espacial y por ello la visión global no es algunas veces del todo adecuada para transmitir la complejidad del problema. En el caso que nos ocupa esta situación es ciertamente una limitación. Lo más adecuado sería espacializar el comportamiento de cada una de las cualidades de un CDG representándolas gráficamente. Existen propuestas al respecto (Mackaness y col. 1994, Beard y col. 1994, Heuvelink 1998, Matos 1998 y 2001), pero las normas ISO no recogen estas circunstancias de ninguna manera.

Campos de error (Figura 3) son la mejor forma de mostrar los resultados de las evaluaciones con vistas a guiar la estrategia de confluencia. Campos de error que muestren comportamientos tendenciales claros sugieren operaciones de ajuste geométrico más elementales. Si los campos visualizan situaciones mucho más complejas se debe pensar en operaciones con capacidad de adaptación local (Ariza 2002), como las técnicas basadas en elementos finitos. Dado que el uso de estas representaciones es complicado, en este trabajo como primera aproximación se propone el uso de medidas de variabilidad de las medidas base, con el propósito de indicar la presencia de comportamientos dispares. Sin embargo, es el análisis sobre esas representaciones en forma de campo el que debe guiar la decisión final sobre los procesos de confluencia. A modo de ejemplo, la Figura 3.a presenta las discrepancias entre dos redes viales que actúan como elementos enjacantes, y la Figura 3.b el campo de errores. En esta última se observa claramente la existencia de comportamientos muy diferenciados espacialmente (p.e. banda izquierda, centro, esquina inferior derecha, etc.). Todo ello sugiere que ajustes básicos del tipo traslación, escalado y rotación no serán buenas soluciones.

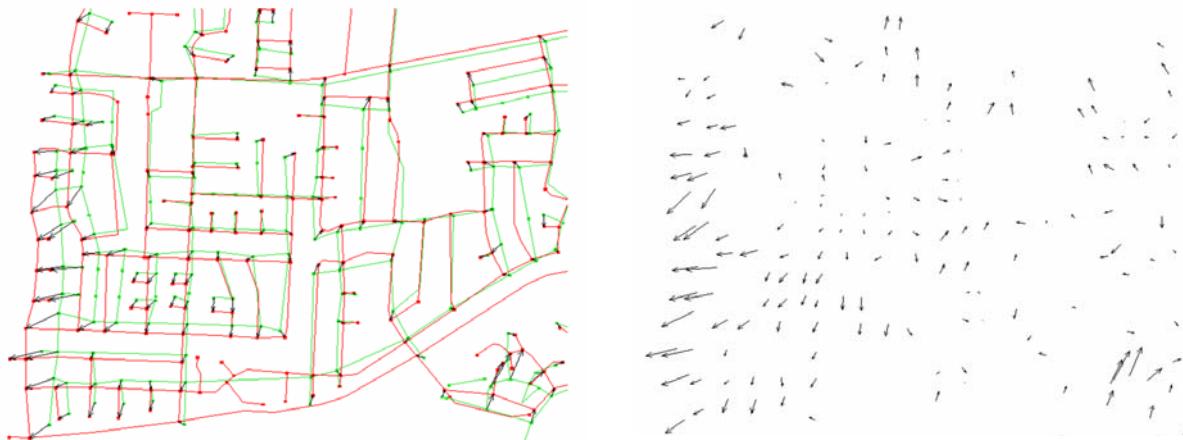


Figura 3.- Ejemplo de conjunto de elementos encajantes (callejero) y de campo de error asociado

Metodología

En este apartado se va a exponer una metodología general para la evaluación, a priori, de la interoperabilidad y, específicamente, en lo relativo a los subelementos interoperabilidad posicional horizontal e interoperabilidad posicional vertical. La metodología se basan en un proceso general que es el siguiente:

- F1: Superposición de los conjuntos de datos. Donde se supone que se habrán realizado las tareas previas que fueran necesarias, como cambios de sistema de referencia y proyección, selección de los elementos de la zona común de interés, etc.
- F2: Determinación de elementos homólogos encajantes. Esta fase de casado, o *matching*, es crítica dado que en algunos casos no podrá ser automática.
- F3: Cómputo de medidas de evaluación seleccionadas.
- F4: Cálculo de estadísticas y análisis de los resultados.

El proceso se va a aplicar a distintas casuísticas de combinación de productos, que son las siguientes:

- Producto vectorial con producto vectorial (VxV).
- Producto imagen con producto imagen (IxI).
- Producto vectorial con producto imagen (VxI).
- Producto modelo de elevaciones con modelo de elevaciones (MDExMDE).
- Producto imagen con modelo de elevaciones (IxMDE).

En todas ellas los dos productos involucrados son de gran importancia para las organizaciones productoras y para los usuarios. Cada una de ellas requerirá una cierta especificación del método general. A continuación se presentarán dichas metodologías particulares con algunos comentarios y propuestas de medidas más adecuadas.

Combinación 1^a: vectorial con vectorial

Los productos vectoriales son los que pueden presentar mayor complejidad de combinación debido a los distintos esquemas topológicos, esquemas de aplicación, etc., que pueden suponer, lo que genera una gran heterogeneidad. Por contrapartida, la perspectiva de orientación a objeto que los fundamenta es un elemento clave para facilitar el caso entre

elementos homólogos encajantes, pero también es un problema en cuanto a los interregnos que existen entre las posiciones de los elementos encajantes y donde se sitúan los elementos encajados.

La combinación entre productos se realiza con vistas a obtener un producto final enriquecido respecto a los dos primeros por medio de la agregación selectiva de elementos de uno y otro. Un ejemplo del caso que nos ocupa, y de gran interés en España, puede ser la combinación de un conjunto de datos de una Base Topográfica (BT), por ejemplo a escala E10k como los que realizan los gobiernos regionales, con otro conjunto de datos procedentes de una Base Catastral (BC), por ejemplo a escala E5k, como la que realiza el Estado.

El empleo de ambos conjuntos de datos de forma conjunta es de utilidad tanto a uno como a otro. Las ventajas para la BC residen en recibir la aportación de toda la información que un mapa topográfico puede suministrar, desde altimetría hasta vías de comunicación pasando por la hidrografía, límites territoriales, etc. Para la BT también supone una gran ventaja en cuanto le permite disponer de un parcelario real y actualizado con todas las construcciones, lo que posibilita aplicaciones temáticas como la de determinación de los posibles usos del suelo, ubicación de edificaciones en zonas rurales, etc.

Si suponemos que tanto la BT como la BC proceden de la cartografía oficial, lo lógico es que estén referidos a un mismo marco y posean un mismo sistema de proyección cartográfica, por lo que la fase F1 de la metodología general puede ser obviada.

A pesar de las diferencias de los esquema de aplicación entre ambos conjunto de datos, el que los dos se encuentren en formato vectorial puede facilitar la determinación de los elementos homólogos encajantes (Fase F2). Para la realización de un caso automático se necesita conocer el esquema de clases, atributos y relaciones (ontologías) de uno y otro CDG para establecer la correspondencia. Si esa correspondencia se puede establecer de manera directa, por ejemplo, entre ejes de caminos, el caso podrá ser automático. Puede ocurrir que la correspondencia suponga un cambio en la tipología de elemento (p.e. colapso de superficial a lineal o puntual) o que, incluso, existan otras circunstancias que no la hagan aplicable. En función de la casuística se podrán utilizar métodos plenamente automatizados, métodos semiautomatizados o métodos totalmente manuales. Esto último también condicionará que la evaluación se realice sobre el 100% de los elementos o que se proceda a un muestreo, que deberá ser significativo.

En relación a la F3 general lo más importante no es el propio cálculo de las medidas, que se realizará de manera automatizada por el sistema, sino la determinación de las propias medidas de interés que vienen condicionadas por el aspecto y subelemento que se deseé evaluar. Según lo indicado anteriormente podríamos utilizar dos tipologías de medidas. Las primeras relativas a desplazamientos entre elementos (p.e. puntuales y lineales), y las segundas tales que dieran idea del principal problema que se encuentra en el uso conjunto de estas dos conjuntos de datos. Este problema consiste fundamentalmente en la falta de ajuste del parcelario catastral al viario de la base topográfica (Figura 4). Por ello se propone contabilizar el número de parcelas afectadas por las vías de comunicación realizando un buffer con diferentes distancias de orlado alrededor de los ejes procedentes de la base topográfica y extrayendo el número de parcelas afectadas en función de la distancia al eje.



Figura 4.- Ejemplos de desajuste posicional entre dos conjuntos de datos vectoriales (una base topográfica y un parcelario catastral: a) encaje general, b) detalle entre un eje de camino (rojo) en la base topográfica y el camino y parcelas (negro) en la base catastral

Según lo anterior, las medidas propuestas para evaluar la interoperabilidad posicional planimétrica en este caso son:

- ¬ Desplazamiento horizontal promedio: Es una estimación de la discrepancia posicional horizontal y relativa entre ambos productos. Se materializa por medio de un valor medio de las discrepancias posicionales medidas entre elementos homólogos.
- ¬ Incertidumbre del desplazamiento horizontal promedio: Es la estimación de la variabilidad correspondiente a la medida anterior.
- ¬ Número de superposiciones incorrectas: Estimación del número de elementos que presentan problema de superposición al combinar los dos CDG.
- ¬ Coherencia en la superposición horizontal: Estimación global del grado de coherencia alcanzado en la superposición.

Las dos medidas primeras son clásicas y no necesitan mayor explicación. La tercera y cuarta establecen criterios de cantidad y coherencia o severidad del problema. Como ya se ha indicado, el interés de la tercera medida propuesta es ligar un desplazamiento posicional relativo con el número de elementos de interés afectados (parcelas) por lo que da una doble idea: magnitud del problema posicional y magnitud de la afectación sobre los elementos de interés, lo cual también supone una cierta información sobre el coste de los trabajos de “arreglo” para conseguir la integración de ambos CDG. Estos trabajos de arreglo pueden desarrollarse bien mediante operaciones de ajuste geométrico (*rubber sheeting*), que demandarían una gran cantidad de información sobre elementos homólogos, o por un proceso más convencional de intersección de coberturas con eliminación de polígonos astilla, modificación geométrica y unión de polígonos, etc. En todo caso, consideramos que el índice propuesto es adecuado para marcar el grado del problema con independencia de que el proceso se realice manual o automatizadamente. La última de las medidas supone usar una escala del tipo bajo, medio, alto para indicar el “grado” de incoherencia general que se aprecia en la superposición, en función de algún conjunto de criterios predeterminado y que aplique el operario según la tipología y apreciación de los problemas encontrados.

A modo de ejemplo, la Figura 5 presenta gráficamente los valores correspondientes al número de parcelas afectadas por las discrepancias relativas entre los ejes de un viario en una BT y las parcelas catastrales de una BC. Para el caso que nos ocupa el valor realmente significativo es el que se corresponde con un desplazamiento nulo, es decir, la línea en su posición actual. Bajo esta condición, el número medio de elementos de la BC que se cortan por cada kilómetro de eje de camino o carreteras de la BT es de 11,4 y 5,2, respectivamente, con un valor

promedio de 10,6 elementos por kilómetro. Esto significa que para integrar adecuadamente ambos CDG, si el proceso de confluencia no es adecuado, se necesitará una actividad de edición manual relativamente intensa para corregir circunstancias similares a las que se muestran en la Figura 4.b anterior.

En el caso de trabajar con dos productos vectoriales la correspondencia entre elementos homólogos se suele hacer manualmente, si bien cada vez son más numerosos los algoritmos existentes para automatizar este proceso, y cuyo grado de complejidad va aumentando. Por ejemplo, Mustière (2006) diseña un proceso geométrico, topológico y semántico para obtener el caso automático entre redes de carreteras, incluso representadas a escalas distintas. Cuando se realiza manualmente el número de elementos emparejados no suele ser muy elevado lo que limita los resultados del método y de la posterior confluencia.

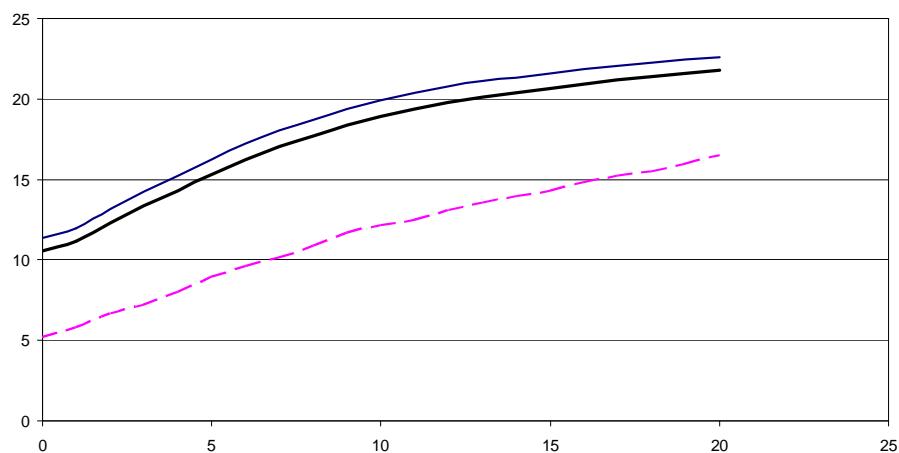


Figura 5.- Ejemplo del número de parcelas catastrales afectadas por kilómetro de longitud según la anchura de orlado para caminos (azul), carreteras (magenta) y de forma conjunta (negro)

Combinación 2^a: imagen con imagen

Supuestas ciertas condiciones previas, que usualmente son satisfechas, estos productos no presentan gran complejidad de superposición debido a la simplicidad conceptual del modelo de representación. Por contrapartida, su carácter de imagen, en las que no existen más objetos que las celdas o píxeles, que deben ser interpretadas por el usuario o un sistema informático, lleva a que la detección de elementos homólogos sea una labor crítica.

La combinación entre productos se realiza con vistas a obtener un producto final enriquecido respecto a los dos primeros por medio de la agregación de las bandas, lo que puede resultar en una base de datos con mayor información y mayores posibilidades de explotación (p.e. interpretación, clasificación, etc.). Se trata pues de un proceso de fusión de datos muy usual en el campo de la teledetección, y de interés en estudios multitemporales y multiespectrales.

Si suponemos que tanto la imagen I_1 como la imagen I_2 proceden de la cartografía oficial, lo lógico es que estén referidas a un mismo marco y posean un mismo sistema de proyección cartográfica, por lo que la fase F1 de la metodología general puede ser obviada. Respecto al tamaño de píxel, si no son iguales se puede proceder al remuestreo o incluso a trabajar con los tamaños originales en esta primera fase.

La determinación de elementos homólogos es una fase crítica que puede desarrollarse con distintos grados de automatización. La automatización total estará basada en la detección de patrones comunes en ambas imágenes, que puede basarse en métodos puramente estadísticos (p.e. correlación) o mucho más complejos. Estas técnicas, al uso en los procesos de fotogrametría digital, permiten detectar y usar en el casado tanto elementos puntuales, lineales como superficiales. Otra opción es la manual, lo que supone una labor de interpretación y digitalización sobre las imágenes con las que se trabaja de un conjunto de elementos homólogos (p.e. puntos, vías de comunicación, etc.). Si se asume la opción manual, se debe tener en consideración si se va a realizar un trabajo que incluya el 100% de los elementos, o si se va a realizar por muestreo. En este último caso se deberá diseñar de tal forma que sea representativo.

Entrando ya en la F3, tal y como se ha indicado en el caso anterior, lo importante es determinar las medidas de interés. Éstas podrán aplicarse sobre elementos puntuales, lineales o superficiales con vistas a determinar los desplazamientos relativos existentes entre I_1 e I_2 . Las medidas propuestas para evaluar la interoperabilidad posicional planimétrica son:

- ¬ Desplazamiento horizontal promedio: Es una estimación de discrepancia posicional horizontal relativa entre ambos productos. Se materializa por medio de un valor medio de las discrepancias posicionales medidas entre elementos homólogos.
- ¬ Incertidumbre del desplazamiento horizontal promedio: Es la estimación de la variabilidad correspondiente a la medida anterior.

Como ya se ha indicado anteriormente, la mejor manera de analizar los resultados es disponiendo del campo de error. En el caso de imágenes éste puede obtenerse, al igual que el caso anterior, manual o automáticamente. Al trabajar con imágenes el proceso automático es bastante más elemental dado que puede realizarse por correlación entre las imágenes, lo que permite disponer de un modelo de error con bastante detalle y que suele garantizar, si se aplican las técnicas adecuadas de transformación, una confluencia adecuada entre los dos productos.

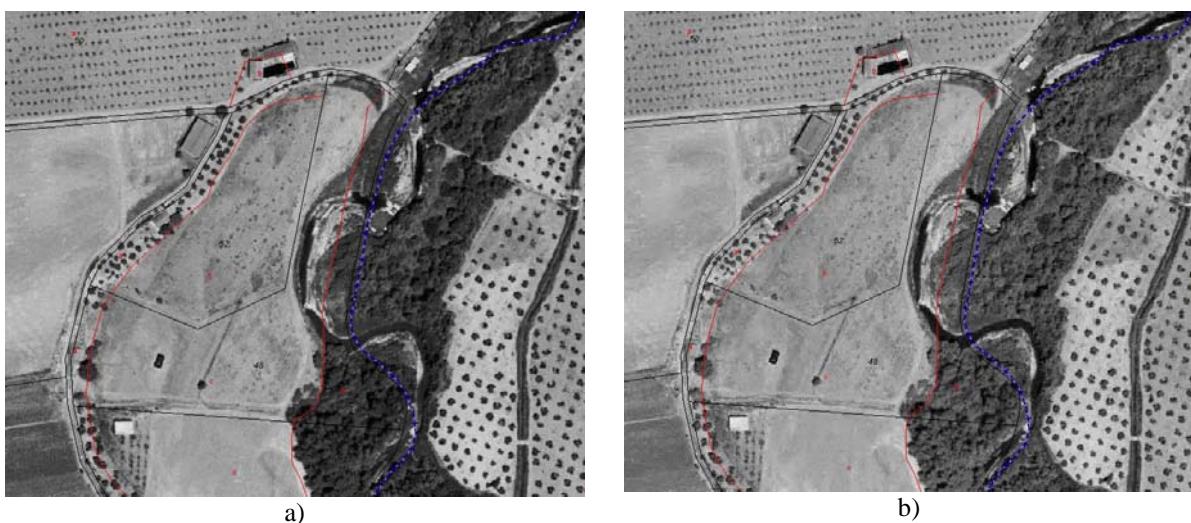


Figura 6.- Ejemplo de una misma zona de interés en dos ortofotos distintas: a) PNOA; b) SIGPAC

Combinación 3^a: vectorial con imagen

Ésta es una opción muy común y de gran interés dado que las dos tipologías de los productos son complementarias y permiten obtener un resultado que integra la estructuración y

modelización de los productos vectoriales con la componente semántica e interpretativa de las imágenes.

Supuestas las condiciones previas que permiten considerar satisfecha la F1, la F2 es la fase que presenta mayor complejidad. En general se dispondrá de un conjunto de elementos (puntos, líneas o superficies) presentes en V, a los cuáles se les deberá determinar sus homólogos en I. La determinación de estos elementos homólogos podrá realizarse manualmente, por digitalización, o de forma automatizada. En el caso de vías de comunicación son numerosos los trabajos que indican, frente al uso de puntos, una mayor viabilidad de la extracción automatizada de estos elementos a partir de imágenes (PE&RS, 2004; Mayer y col. 2006), lo que ofrece además una mayor fiabilidad de la estimación al contar con mayor número de medidas y una implantación espacial que puede llegar a ser relativamente densa. La Figura 6 presenta el ejemplo de diversos ejes de caminos procedentes de una base topográfica y sus homólogos digitalizados sobre unas ortofotos. El disponer de los elementos en V permite además simplificar las búsquedas en I y la asignación de los códigos de homólogos. En la combinación de ciertos productos los elementos homólogos son más complejos de determinar debido a la existencia de límites que han de interpretarse. Un ejemplo de esto último es la integración de un parcelario catastral con una ortofotografía. La evaluación puede realizarse por estimaciones posicionales utilizando puntos o elementos lineales como vías de comunicación, pero un aspecto base de la calidad de la integración de ambos CDG es que los límites de las parcelas catastrales sean verosímiles con los límites que se intuyen en la ortofoto (Figura 7), este es un caso que deberá resolverse manualmente. Al igual que en se ha apuntado en los casos anteriores, si se asume la opción manual se debe tener en consideración si se va a realizar un trabajo que incluya el 100% de los elementos o si se va a realizar por muestreo. En este último caso se deberá diseñar de tal forma que sea representativo.

Según lo indicado, las medidas que se proponen para estimar la interoperabilidad posicional planimétrica son:

- ¬ Desplazamiento horizontal promedio: Es una estimación de la discrepancia posicional horizontal y relativa entre ambos productos. Se materializa por medio de un valor medio de las discrepancias posicionales medidas entre elementos homólogos. Puede estimarse a partir de elementos puntuales, lineales o superficiales.
- ¬ Incertidumbre del desplazamiento horizontal promedio: Es la estimación de la variabilidad correspondiente a la medida anterior, con independencia del tipo de geometría base.
- ¬ Número de superposiciones incoherentes horizontales: Estimación del número de elementos que presentan problema de superposición al combinar los dos CDG. Por ejemplo, considerar los casos en los que los límites de las parcelas no coinciden en uno y otro modelo o sufren un desplazamiento excesivo.
- ¬ Coherencia en la superposición horizontal: Estimación global del grado de coherencia alcanzado en la superposición.

Las dos medidas primeras son clásicas y no necesitan mayor explicación. Como ya se ha indicado, el interés de la tercera medida propuesta es estimar la coherencia espacial de la propia superposición. Conviene indicar que este análisis sólo puede desarrollarse donde existe cambio de cultivo o un ribazo, u otro elemento que materialice, de forma clara, la existencia de las lindes de las parcelas. En los casos en los que parcelas contiguas poseen el mismo cultivo también se pueden identificar lindes de parcelas si existen circunstancias (densidad de cultivo, desarrollo fenológico, laboreo del suelo, etc.) que permitan detectar en la imagen un comportamiento distinto entre las zonas bajo consideración (Figura 8). Se trata de una

evaluación que deberá ser realizada por un operario capacitado. En este caso se pueden establecer criterios de cantidad y/o severidad. El primero de ellos se corresponde con la tercera de las medidas propuestas, mientras que el segundo con la cuarta. Por ejemplo, se pueden contabilizar los casos (cantidad), pero también se puede usar una escala del tipo bajo, medio, alto para indicar el “grado” de incoherencia general que se aprecia en la superposición, en función de un conjunto de criterios predeterminado y que aplique el operario según la tipología y apreciación de los problemas encontrados.

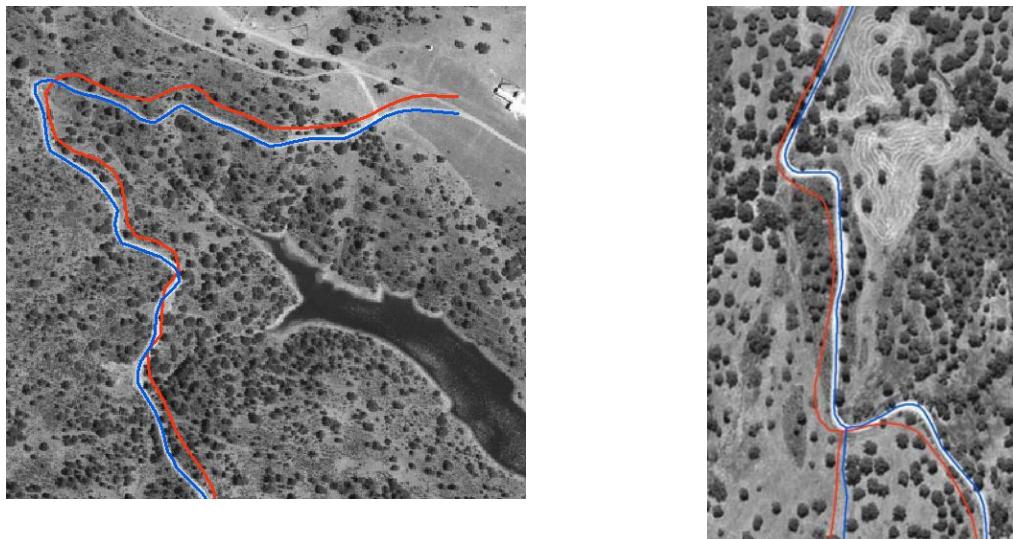


Figura 7.- Ejemplo de elementos homólogos lineales (caminos) procedentes de una base topográfica vectorial (rojo) y una ortofoto (azul)

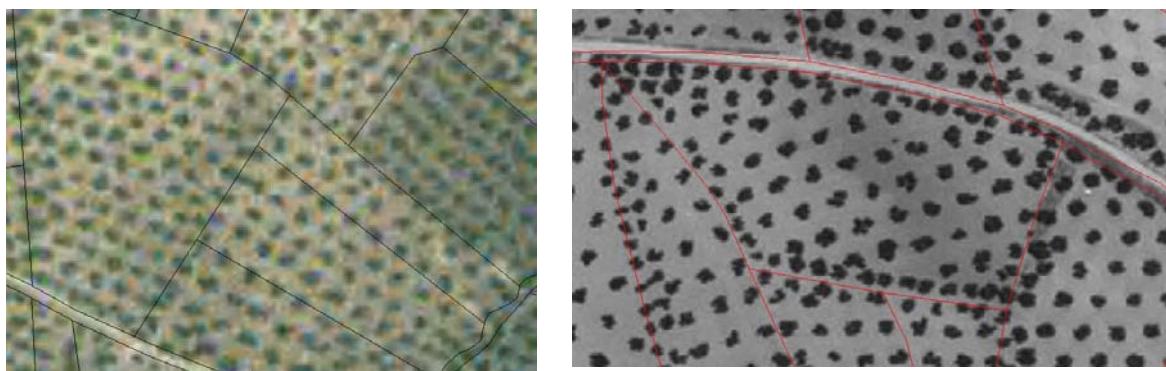


Figura 8.- Ejemplo de superposición entre una base catastral parcelaria y una ortofotografía. En algunos caso se aprecia la coherencia/falta de coherencia entre los límites parcelarios vectoriales y los cambios que se observan en el cultivo del olivar

Combinación 4^a: vectorial con MDE

Se supone que V dispone de altimetría en forma de curvas de nivel. Para el caso del MDE se asume una estructura de malla o TIN.

El beneficio de esta integración provendría, en caso de ser suficientemente interoperables, de poder sustituir el modelo de curvas de nivel del producto V por el modelo M, más funcional de cara a análisis en un sistema de información geográfica, que uno basado en curvas de nivel. Este integración permitiría integrar análisis propios de una base topográfica (p.e. selecciones

temáticas, zonas de influencia, distancias, etc.) con las de un MDE (p.e. pendientes, insolación, etc.) de forma coherente.

Como en los casos anteriores, se supone que se dan las condiciones que permiten considerar que la F1 se satisface. En este caso se incluye dentro de esta fase la conversión, por interpolación, desde los modelos de rejilla o TIN al modelo de curva de nivel con vistas a disponer de la información altimétrica en un mismo modelo. También se podría convertir la información altimétrica en forma de curvas de nivel en una rejilla o en un TIN. En estos comentarios se presentará sólo el primero de los casos indicados.

En la obtención de las curvas homólogas por interpolación se debe cuidar que el proceso de interpolación no genere sobre las mismas resultados degradados o con artefactos (lazos, creación de curvas inexistentes, etc.). Además de los problemas ya indicados, otro de los aspectos que interesa tener en cuenta es que debido al modelo y resolución de los MDE de tipo malla se generan curvados con un comportamiento en zig-zag. Estos son aspectos limitantes, si bien pueden resolverse por selección de los mejores elementos, o por algún tipo de postprocesado. En cualquier caso, esta solución permitirá automatizar en gran medida el proceso y llevarlo, si se desea, a una inspección 100%.

Con respecto a la F2, la determinación de elementos homólogos queda solucionada por los códigos de altura que portan las curvas de nivel y su implantación espacial que limita el espacio de búsqueda.

Las medidas que se proponen para evaluar la interoperabilidad posicional altimétrica son las siguientes:

- Desplazamiento horizontal promedio: Es una estimación de la discrepancia posicional horizontal y relativa entre ambos productos. Se materializa por medio de un valor medio de las discrepancias posicionales medidas entre las trazas horizontales de curvas de nivel homólogas.
- Incertidumbre del desplazamiento horizontal promedio: Es la estimación de la variabilidad correspondiente a la medida anterior, con independencia del tipo de geometría base.
- Desplazamiento vertical promedio: Es una estimación de la discrepancia posicional vertical relativa entre ambos productos. Se materializa por medio de un valor medio de las discrepancias posicionales medidas entre las trazas verticales de curvas de nivel homólogas.
- Incertidumbre del desplazamiento vertical promedio: Es la estimación de la variabilidad correspondiente a la medida anterior, con independencia del tipo de geometría base.

Las medidas propuestas requieren cierta explicación para entender su funcionamiento, para lo que se han graficado las discrepancias horizontales y verticales entre dos curvas de nivel “homólogas”. La Figura 9.a muestra un ejemplo del concepto que se pretende cuantificar con el desplazamiento horizontal promedio. Se observan dos curvas de nivel, la procedente de V y la extraída por interpolación del MDE y como no coinciden en su traza planimétrica. La Figura 9.b presenta la evolución de la distancia planimétrica entre ambas cuando se rectifican ambos elementos lineales haciéndolos coincidir con el eje de abscisas, obteniendo así una representación más cómoda de interpretación y que recuerda a las series temporales. La representación planimétrica de estas discrepancias permite obtener un campo de error en cierta forma equivalente a los ya presentados.

Por otra parte, la representación horizontal de cada una las curvas de nivel del producto V permite extraer un perfil altimétrico en el MDE. Este perfil indicará una altitud variable de

dichas curvas de nivel en el MDE si bien, por definición, deberían ser de cota constante. Por tanto, este perfil también permite apreciar cuánto se desvía un producto respecto a otro. Esta circunstancia es la que muestra la Figura 10 donde se presenta el caso de una curva de nivel de cota 400 m, en esta figura el eje horizontal representa desplazamiento (longitud) a lo largo de la curva, es decir, la rectificación de la curva, mientras que el eje vertical presenta las variaciones respecto al nivel de cota de referencia. Las variaciones más acusadas están ligadas a los comportamientos cóncavos/convexos más acusados de las curvas de nivel, pues es donde pequeñas variaciones posicionales planimétricas pueden generar mayor discrepancia altimétrica. La aplicación de este método sobre todos los elementos de interés permitiría obtener una superficie de discrepancias altimétricas entre ambos modelos.

Como se ha indicado al inicio de este epígrafe, la idea es sustituir una información por otra, por lo que en este caso los campos de error no se plantean con el objeto de una utilización directa como base de un proceso de confluencia, sino con un carácter meramente de evaluación de la compatibilidad entre las informaciones de cada CDG.

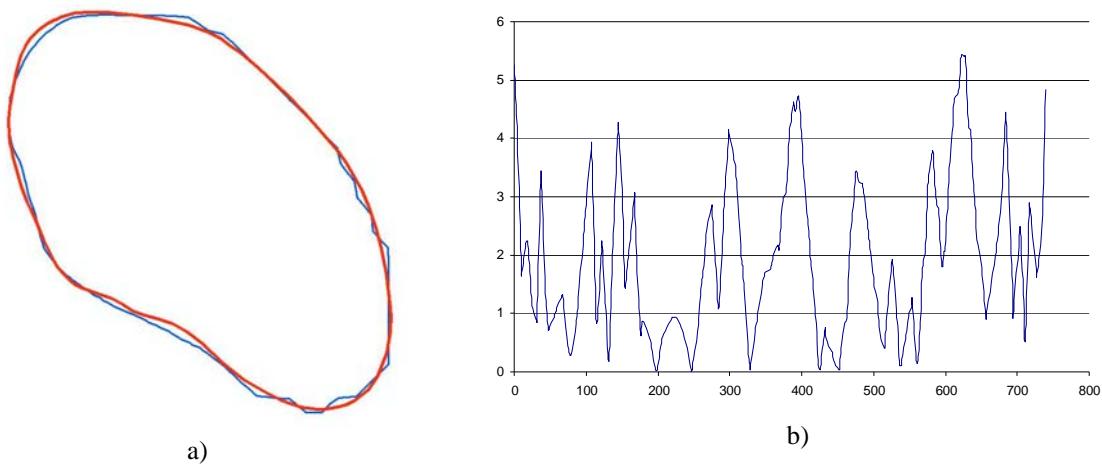


Figura 9.- Evolución de la discrepancia planimétrica entre curvas de nivel de un modelo de curvas de nivel y su homóloga derivada de un modelo de elevaciones tipo malla: a) aspecto planimétrico, b) aspecto tras la rectificación de las discrepancias planimétricas

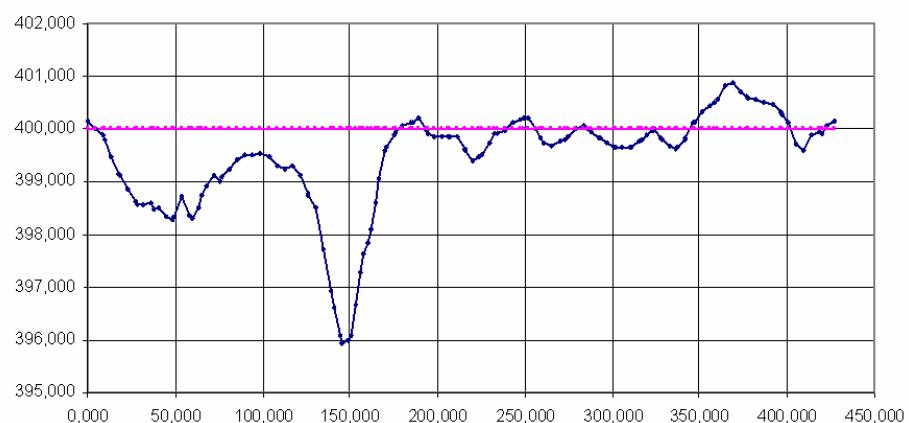


Figura 10.- Comparación del perfil altimétrico de una curva de nivel de un modelo de curvas de nivel y las cotas extraídas para esa curva de nivel en un modelo digital de elevaciones

Combinación 5^a: imagen con MDT

Se supone que I es una ortofotografía y que el MDE dispone de una estructura base de tipo malla, TIN o curva de nivel.

La interoperabilidad entre conjuntos de datos de este tipo es de gran interés en aplicaciones como simulaciones de vuelos y en otras oportunidades que brinda la geomática. Sin embargo, este análisis de confrontabilidad es, sin duda, el caso más difícil de determinar de los presentados hasta ahora. La dificultad proviene de comparar un producto que sólo incluye información altimétrica (MDE) con otros que sólo tienen información interpretativa o fotográfica (I).

Como en los casos anteriores, se supone que se dan las condiciones que permiten considerar que la F1 se satisface. En este caso, dentro de esta fase, se incluyen la conversión, por interpolación, desde los modelos de rejilla o TIN a modelo de curva de nivel con vistas a disponer de la información altimétrica en un mismo modelo.

Si bien los productos de tipo I son imágenes planas (ortofotos), el relieve se encuentra implícito por lo que puede ser interpretado en gran medida. De esta forma, se puede considerar la posibilidad de una interoperabilidad posicional altimétrica entre productos I y MDE. Indudablemente, esta interoperabilidad posicional altimétrica estará estrechamente ligada con la planimétrica pero permite presentar de una manera específica la situación desde esta importante perspectiva de la IG.

Dado que el producto de tipo I no incluye altimetría, la evaluación de este subelemento no puede ser directa, pero sí podrá derivarse a partir de elementos de la planimetría (p.e. desplazamientos en el sistema hidrográfico, etc.). Nuestra propuesta va en dos líneas:

- Superponer la I sobre superficies 3D del MDE, en el caso de los modelos malla y TIN.
- Superponer las curvas de nivel sobre la I, en el caso de modelo basados en curvas de nivel (p.e. altimetría de productos V), o de derivar estos elementos a partir de los modelos malla y TIN.

En cualquiera de los casos el análisis consistirá en la inspección y evaluación visual de la superposición entre el producto I y del MDE. Para el primero de los casos se superpondrá la imagen a la superficie generada por la malla o el TIN y se atenderá al comportamiento de los elementos estructurantes del territorio. Se analiza de esta forma la coherencia visual 3d entre los dos CDG a partir del encaje de la red de drenaje, de los embalses, de los desmontes y terraplenes, de las vías de comunicación, de la morfología del terreno (picos, laderas, vaguadas, etc.), etc. En el segundo de los casos se superpondrán las curvas de nivel a la imagen, y se atenderá básicamente al mismo conjunto de elementos y criterios. Estas evaluaciones deben ser siempre complementarias a la horizontal y aportan la perspectiva específica de la funcionalidad del comportamiento altimétrico.

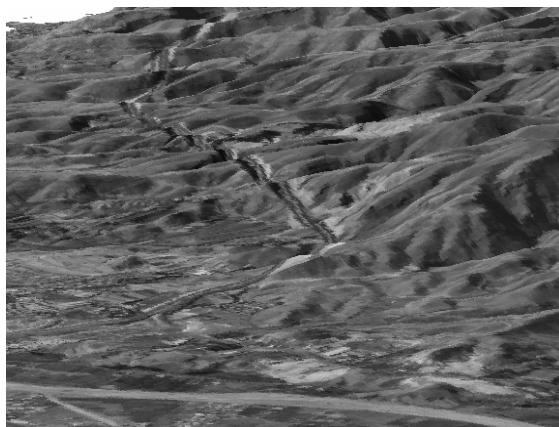
Según la metodología general propuesta la fase F2 de determinación de elementos homólogos deberá ser realizada por un operario. El operario deberá estar entrenado en este tipo de evaluación y disponer, conocer y aplicar un conjunto de guías de evaluación que sean públicas para que todo el mundo pueda entender el método aplicado y el significado claro de sus resultados.

Por todo lo anterior, las medidas propuestas para evaluar la interoperabilidad posicional altimétrica son:

- ¬ Número de superposiciones incoherentes horizontales: Estimación del número de elementos que presentan problema de superposición al combinar los dos CDG.
- ¬ Coherencia en la superposición 3D: Estimación global del grado de coherencia alcanzado en la superposición de la imagen sobre una superficie generada a partir del MDE.

Para la primera medida, se recogerá, de forma cuantitativa, el número de casos encontrados en cuanto a los desplazamientos, apareciendo para cada uno de los elementos estudiados (vaguadas, divisorias, caminos y curvas de nivel). Este conteo se puede estratificar atendiendo al grado del desplazamiento entre dichos elementos, por ejemplo, alto (cuando el desplazamiento muy llamativo), medio (cuando el desplazamiento se aprecia pero es leve) o bajo (cuando el desplazamiento es inapreciable).

Advirtiendo que en este caso el criterio es el de “encaje” y no el de desplazamiento. El segundo criterio utilizado es el de grado, que hace referencia a la bondad del encaje global, denotándose como alto, cuando los elementos encajan muy bien, medio, cuando en el encaje se observa un cierto desplazamiento y bajo, cuando los elementos no encajan. Se trataría pues de una estimación más subjetiva, derivada bien de la experiencia del operario o a partir de la cuantificación de la medida anterior. Como ejemplo de lo que se quiere indicar, la Figura 11 presenta una perspectiva en la que se ha superpuesto la I al MDE. Mediante vuelos virtuales el operario irá visitando la zona de interés y analizando la consistencia de la superposición de la imagen sobre la superficie altimétrica atendiendo a taludes, laderas, divisorias, etc. En planimetría se puede realizar algo parecido, y también de carácter estimativo, como se muestra en la Figura 12 que presenta las curvas de nivel superpuestas a la imagen, y donde se puede apreciar una falta de encaje en el arroyo que atraviesa la zona.



a)



b)

Figura 11.- Ejemplos de elementos considerados en la evaluación de la coherencia del encaje entre una ortofotografía y la superficie 3D y un modelo de elevaciones del terreno de tipo malla: a) taludes y desmontes, b) laderas y divisorias

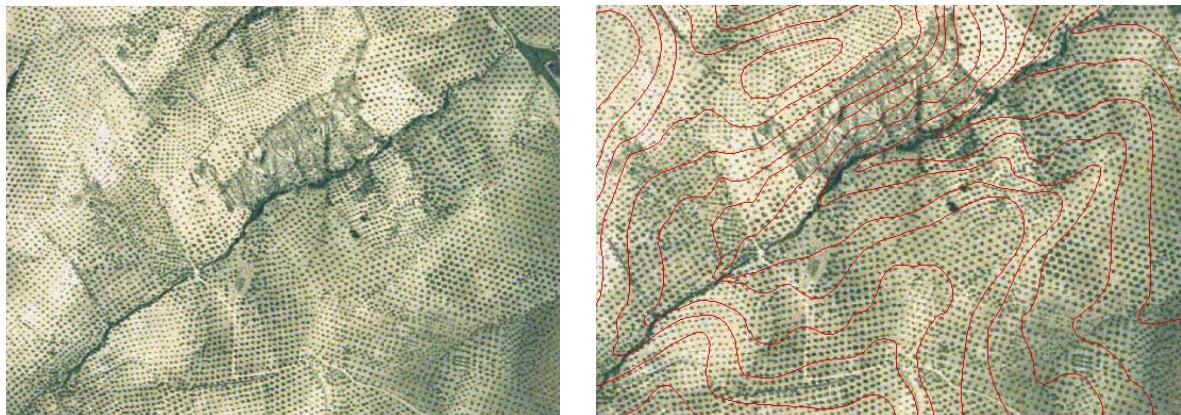


Figura 12.- Ejemplo de análisis del encaje entre una ortofotografía con un modelo de elevaciones representado en forma de curvas de nivel

Conclusiones

En este trabajo se ha realizado una justificación de la importancia del concepto de interoperabilidad posicional dentro del marco de las IDE, dado que existe un problema real de falta o pérdida de interoperabilidad debida a las discrepancias posicionales entre conjuntos de datos geográficos.

La interoperabilidad posicional se ha formalizado a partir de las definiciones ya existentes de confrontabilidad e integrabilidad. Esta formalización se ha materializado estableciendo el objetivo y alcance, indicando posibles elementos y subelementos de la calidad y proponiendo medidas y métodos de evaluación.

Una conclusión importante es la necesidad de incorporar la interoperabilidad en el conjunto de elementos definitorios de la calidad de un conjunto de datos. Si bien en este caso se trata de un elemento relativo a dos conjuntos de datos geográficos.

Además, se han propuesto subelementos, medidas y métodos de evaluación específicos que permiten, en parte, evaluar el grado de interoperabilidad posicional existente entre dos conjuntos de datos geográficos. Estos subelementos y las medidas han de permitir a los usuarios tener una visión más clara del alcance del problema y guiar posibles estrategias de confluencia.

Todo lo anterior se ha realizado para cinco combinaciones usuales de productos: vector-vector, imagen-imagen, vector-imagen, vector-modelo de elevaciones, e imagen-modelo de elevaciones. Ello ha permitido concretar una metodología general y algunos aspectos específicos para cada una de esas combinaciones, estableciendo una propuesta de medidas iniciales a aplicar a procesos que pueden ser totalmente automatizables o de interpretación visual por parte de un operario. Los primeros se nos muestran como los más adecuados por su mayor robustez, objetividad, y posibilidad de ligarlos a procesos automatizados y a muestreos 100% sobre grandes poblaciones. Los métodos interpretativos presentan una mayor problemática para obtener un resultado de valor, sin embargo, en algún caso pueden ser la única opción viable. Por ello, si bien conviene reducir su uso, en caso de necesidad se hace indispensable determinar muy bien los criterios de su desarrollo y basarlas las estimaciones finales en un diseño muestral adecuado. Cada una de estas combinaciones de conjuntos de

datos geográficos se ha mostrado con ejemplos gráficos reales que muestran de una manera evidente el alcance del problema.

En relación a las instituciones cartográficas y los usuarios, se evidencia la necesidad de evaluar esta nueva componente e incluir los resultados en los metadatos como forma de informar, de manera objetiva y significativa, de la posibilidad de interoperar entre conjuntos de datos.

Sin embargo, por nuestra experiencia en España en la evaluación de esta nueva componente de la calidad de la IG, consideramos que la principal conclusión tras las diversas evaluaciones realizadas es la necesidad de establecer y ejecutar planes de mejora de la interoperabilidad posicional. Estos planes deben atender tanto a las líneas de producción de una misma organización cartográfica, de cara a obtener una producción espacialmente más integrada, coherente y de calidad, como a varios productores. Las acciones a emprender en este caso pasan por incluir en las especificaciones de los productos exigencias de interoperabilidad con otros productos (propios o de terceros), asegurando con ello métodos, marcos geodésicos, fuentes de información, etc., que sean lo más comunes posible.

Agradecimientos

Al Gobierno regional de la Comunidad Autónoma de Andalucía por las ayudas que viene aportando al Grupo de Investigación en Ingeniería Cartográfica, Geodésica y Fotogrametría (TEP-164) dentro del Plan Andaluz de Investigación e Innovación.

Referencias

- Ariza, F.J. (2002). *Control de Calidad en la Producción Cartográfica*. Ed. Ra-Ma. Madrid.
- Ariza, F.J. (2006). *Factores determinantes de la calidad de los productos/servicios cartográficos*. Mapping, nº 112, pp. 30-39.
- Ariza, F.J.; Atkinson, A.D. (2008). *Analysis of Some Positional Accuracy Assessment Methodologies*. Journal of Surveying Engineering, ASCE, Vol 134, nº 2, pp. 45-54.
- Ariza, F.J.; Atkinson, A.D.; Nero, M. (2006). *Análisis de algunas metodologías de evaluación de la componente posicional*. Topografía y Cartografía, vol. 24, nº 140, pp. 32-45.
- Beard, M.K.; Buttenfield, B.; Mackaness, W. (1994). *Research Initiative 7: Visualization of the Quality of Spatial Information*. Closing Report. NCGIA. Santa Bárbara, California.
- Bishr, Y., (1998). *Overcoming the semantic and other barriers to GIS interoperability*. En International Journal Geographical information science, 1998, vol. 12, no. 4, pp. 299-314.
- Carmel, Y.; Dean, D.; Flather, C. (2001). *Combining Location and Classification Error Sources for Estimating Multi-Temporal Data Base Accuracy*. En PE&RS, vol. 67, nº 7.
- Carmel, Y.; Flather, C.; and Dean, D. (2006). *A methodology for translating positional error into measures of attribute error, and combining the two error sources*. Proceedings of Accuracy 2006. 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Caetano, M and Painho, M. (eds.), Lisbon, 3-17.

- Casado, M. (2006). *Some Basic mathematical constrains for the geometric conflation problem*. En 7th International Symposium on Spatial Accuracy in Natural Resources and Environmental Sciences.
- Cobb, M.; Chung, M.; Foley, H.; Petry, F.; Shaw, K. (1998). *A Rule-Based Approach for the conflation of attributed Vector Data*. Geoinformatica, 2(1), 7-35.
- Curtin, K., Funk, C., Goodchild, M., Montello, D., (1998). *Formulation and test of a model of a positional distortion fields*. En www.ncgia.ucsb.edu/vital/research/reports.html
- Church, R., Curtin, K., Fohl, P., Funk, C., Goodchild, M., (1998). *Positional distortion in geographic data sets as a barrier to interoperation*. En www.geog.ucsb.edu/~good/
- Deza, E.; Deza, M.M. (2006). Dictionary of Distances. Elsevier.
- Directiva 2007/2/CE del Parlamento Europeo y del Consejo, de 14 de marzo de 2007, por la que se establece una infraestructura de información espacial en la Comunidad Europea (Inspire)
- Edwards, D. Simpson, J. (2002). *Integration and access of multisource vector data. Geospatial Theory, Processings and Applications*. ISPRS Commission IV. Ottawa, Canada.
- EU-JRC (2006). *Report of the international Workshop on Spatial Data Infrastructure Return on investment: Assessing the Impact of SDI*. Ispra, Italy.
- EUROSDR (2004). *Positional Accuracy Improvement: Impacts of improving the positional accuracy of GI databases*. En http://www.eurosdr.net/km_pub/no48/html/positional/positional_index.htm.
- FGDC (1994). *The 1994 plan for the National Spatial Data Infrastructure. Building the Foundation of an Information Bases Society*.
- Fonseca, F., Egenhofer, M., (1999). *Ontology-Driven Geographic Information System*. 7th ACM Symposium on Advances in Geographic Information System, Kansas City, MO.
- Fonseca, F., Egenhofer, M., (2002). *Using ontologies for integrated geographic information systems*. En F. Fonseca, M., Egenhofer, P., Agouris and C. Câmara. Trasactions in GIS 6(3).
- Heuvelink, G. (1998). *Error propagation in Environmental Modelling*. Taylor & Francis. Londres.
- ISO 19113:2002. *Geographic information - Quality principles*.
- ISO 19114:2003. *Geographic information - Quality evaluation procedures*.
- ISO 19115:2003. *Geographic information – Metadata*.
- ISO 19138:2005. *Geographic Information - Quality Measures*.
- ISO 2859-1:1999. *Sampling procedures for inspection by attributes - Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection*.
- ISO 3951-1:2005. *Sampling procedures for inspection by variables - Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL*
- Jakobsson, A.; and Vauglin, F. (2002). *Report of a questionnaire on data quality in National Mapping Agencies*. CERCO Working Group on Quality. Marne-la-Vallée.
- Lendínez, J.J.; Ariza, F.J. (2008). *Conflación de Bases de Datos Geográficos: clasificación y algoritmos*. En Prensa.
- Levinshohn, A., (2001). *La interoperabilidad Geospatial: El santo grial del campo SIG*. En Mapping (Abril 2001).

- Mackaness, W.; Buttenfield, B.; Beard, M.K.; (1994). *Selected Annotated Bibliography on Visualization of the Quality of Spatial Information*. Research Initiative 7. NCGIA. Santa Bárbara, California.
- Matos, J.L. (2001). *Fundamentos de Informação Geográfica*. Lidel, Lisboa.
- Matos, J.L.; Goncalves, A. (1998). *Measurement and analysis of position error*. En *Proceedings of the 8th International Symposium on Spatial Data Handling*. Ed. IGU.
- Mayer, H.; Hinz, S.; Bacher, U.; Baltsavias, E. (2006). *A test of automatic road extraction approaches*. En Photogrammetric Computer Vision, ISPRS Symposium of Comision III, Bonn. Pp. 1-6.
- McGwire, C., (1996). Cross Validated Assessment of geometric accuracy. En *PE&RS*, vol. 62, nº 10.
- McMaster, R. y Shea, K. (1992). *Generalization in digital cartography*. AAG. Washington, D.C.
- PE&RS (2004). *Special Issue: Automatic Road Extraction*. PE&RS, vol. 70, nº 6.
- Rönsdorf, C. (2004). *Positional Integration of Geodata. Positional Accuracy Improvement: Impacts of improving the positional accuracy of GI databases*. EuroSDR Publication N°. 48.
- Veregin, H.; Giordano, A.; (1994). *Il controllo di qualità nei sistemi informativi territoriali*. Il Cardo, Venezia.
- Mustière, S. (2006). *Results of experiments on automated matching of networks at different scales*. En *Proceedings of the ISPRS WorkShop on Multiple Representation and Interoperability of Spatial Data*, Hanover, pp. 92-100.