

# CALIDAD DE DATOS GEOESPACIALES BÁSICOS

## QUALITY OF BASIC SPATIAL DATA

Fecha de recepción: 29 de julio de 2009 / Fecha de aprobación: 15 de enero de 2011

Edilberto Niño Niño

### Resumen

La valoración de la calidad de los datos geográficos persigue suministrar tanto al productor como al consumidor, la aptitud del producto para sus usos potenciales. No pretende definir un nivel mínimo aceptable de calidad para los datos geográficos.

La calidad de un conjunto de datos geoespaciales se examina bajo los lineamientos indicados por cinco (5) elementos de calidad con sus subelementos.

Cada país ha desarrollado estrategias, normas y documentación a fin de crear conciencia respecto a medición de la calidad de los datos geoespaciales, en el caso colombiano, se creó la Norma Técnica Colombiana NTC 5043, la cual fue desarrollada por entidades que tienen que ver con la producción o manejo de datos geoespaciales, con el liderazgo del Instituto Geográfico Agustín Codazzi.

Dicha norma propone examinar la calidad de los datos teniendo en cuenta los elementos de calidad los cuales son:

Grado de totalidad, Consistencia lógica, Exactitud de posición, Exactitud temporal, Exactitud temática. Se espera que examinando estos aspectos a unos datos geoespaciales se garantice la valoración de los datos en todos los aspectos.

**Palabras Clave:** Georreferenciados, inherentes, topología, restitución, conformidad, calidad

### Summary

The assessment of the quality of geographic data aims to provide both the producer and the consumer, the product's suitability for its potential uses. No attempt to define a minimum acceptable level of quality for geographic data.

The quality of a geospatial data set is determined for guidelines given by five quality elements and sub-elements.

Each country has developed strategies, policies and documentation to create awareness of measuring the quality of geospatial data, in the case of Colombia, was created a Colombian Technical Standard NTC 5043 developed by different organizations that are related with the production and management of geospatial data, under the leadership of the Geographic Institute Agustín Codazzi.

The standard NTC 5043 proposes to inspect the quality of the data taking into account the quality elements which are: completeness, logical consistency, positional accuracy, temporal accuracy and thematic accuracy. It is expected that checking these elements, we guarantee the whole aspects of the quality of spatial data.

**Keywords:** Geo-referencing, inherent, topology, photogrammetric restitution, compliance, quality.

## 1. Introducción

El término “calidad” parece expresar un modelo de perfección en la producción de bienes y servicios; la calidad debería implicar eficiencia y eficacia, lo que significa reducir costos y aumentar rentabilidad; sin embargo lo cierto es que la calidad de los bienes y servicios es y será una característica que implica un valor estratégico y una invaluable ventaja frente a la competencia.

Las sociedades gastan gran cantidad de esfuerzos y de recursos en verificar si lo que se hizo está bien hecho, y en un índice más alto, si lo que se dice es verdad; esto probablemente se deba a que los sistemas de producción están siendo orientados a obtener la mayor ganancia sin haber involucrado la calidad en las líneas de producción como una característica implícita de los productos.

En relación con los datos espaciales básicos, conocer su calidad implícita es fundamental ya que de ésta depende la calidad de todos los trabajos que con ella se realicen. Adicionalmente, una vez los datos cartográficos salgan de las áreas de producción es muy difícil medir la calidad de éstos. En el presente artículo se presentan de manera detallada los aspectos que se deben examinar para determinar la calidad de una base de datos cartográfica.

## 2. Qué se entiende por mapa

De forma general, un mapa básico es una representación gráfica convencional de una porción de la superficie terrestre con todo lo que en ella existe, o lo que el hombre percibe que existe. Los elementos que debe contener un mapa son definidos en las especificaciones de él, mediante un catálogo de objetos y un modelo de datos.

### 2.1 Un modelo de la realidad

Un mapa básico es una representación gráfica convencional de los rasgos naturales y culturales existentes en el terreno. Debido a que los objetos topográficos con sus características, contenidos en el mapa, están georreferenciados<sup>1</sup> con al menos cuatro coordenadas, es lícito llamarlos datos geoespaciales. Los objetos que debe contener un mapa

1 Que están unidos a un sistema de coordenadas sobre la superficie terrestre.

son definidos en sus especificaciones, mediante un catálogo de objetos<sup>2</sup> y un modelo de datos.<sup>3</sup>

## 3. Concepto general de calidad

El término “calidad” lo utiliza el ser humano en su lenguaje diario; es entendido y aceptado por todos. Se habla de buena y mala calidad dependiendo del consumidor, en este sentido el término calidad se vuelve relativo, es decir, depende del usuario. La calidad en sentido general se define como la propiedad o conjunto de propiedades inherentes<sup>4</sup> a un producto, que permite apreciarlo como mejor, igual o peor que los demás de su especie.

Teniendo en cuenta la definición anterior, la calidad no se limita a una única propiedad del objeto o producto, sino que involucra varios factores, los cuales se deben determinar en el momento de calificar la calidad de un producto.

Por otra parte, si se toma en cuenta que los sistemas de información geográfica (SIG), los grandes avances en la geomática, la teledetección y la fotogrametría digital impulsan, facilitan y permiten el consumo masivo de los datos geográficos; y a éste sumados la relativa facilidad de adquirir y modificar de forma rápida los datos geoespaciales (en algunos casos el hecho de ser producidos por técnicos no especialistas en la materia), todo trae como consecuencia la obtención de productos cartográficos con un gran número de inconsistencias, las cuales se hacen manifiestas cuando el producto se usa, o cuando se somete a una revisión o proceso de control de calidad.

### 3.1. Medición de la calidad de datos geoespaciales

La valoración de la calidad de los datos geográficos persigue suministrar tanto al productor como al consumidor la aptitud del producto para sus usos potenciales. Los consumidores o usuarios de cartografía generalmente esperan que un mapa sea exacto, lógico, coherente.

La medición de la calidad de datos geoespacia-

2 Contiene una descripción de los objetos topográficos, así como la definición, clasificación, codificación, geometría y características de ellos.

3 Permite representar la información geográfica y topográfica, lo mismo que las relaciones de los elementos del paisaje.

4 Que está de tal manera unido a algo, que no se puede separar.

les se realiza con base en lineamientos de normas técnicas elaboradas para tal fin. En Colombia existe la Norma Técnica Colombiana (NTC) 5043,<sup>5</sup> la cual fue desarrollada por entidades que tienen que ver con la producción o manejo de datos geoespaciales, con el liderazgo del Instituto Geográfico Agustín Codazzi.

### 3.1.1 Norma Técnica Colombiana (NTC) 5043

Principales características:

- El objetivo general es proporcionar los conceptos básicos que permiten describir la calidad de los datos geográficos, disponibles en formato digital y análogo.
- Presentar un modelo conceptual que facilite el manejo de la información sobre la calidad de datos geográficos.
- Define objetos de calidad que los productores pueden aplicar para describir y evaluar si un conjunto de datos cumple su función de representar un universo abstracto (el terreno), de conformidad con las especificaciones del producto y que los usuarios pueden usar, para establecer si un conjunto de datos cumple con la calidad para una aplicación específica.
- No pretende definir un nivel mínimo aceptable de calidad para los datos geográficos. (Tomado de la norma técnica mencionada).

La calidad de un conjunto de datos geoespaciales se examina bajo los lineamientos indicados por cinco elementos de calidad con sus subelementos los cuales se explicitan a continuación:

## 3.2 Elementos de Calidad

Los elementos de calidad son: Grado de totalidad, Consistencia lógica, Exactitud de posición, Exactitud temporal y Exactitud temática. A su vez, cada uno de estos elementos cuenta con subelementos de calidad como se ilustra en el cuadro 01.

Elementos de calidad	Subelementos de calidad
A- Grado de totalidad	a1- Comisión
	a2- Omisión
B- Consistencia lógica	b1- Consistencia de dominio
	b2- Consistencia topológica
	b3- Consistencia de formato
C- Exactitud de posición	c1- Exactitud absoluta o externa
	c2- Exactitud relativa o interna
	c3- Exactitud de posición de datos de celdas
D- Exactitud temporal	d1- Exactitud en la medición del tiempo
	d2- Consistencia temporal
	d3- Validez temporal
E- Exactitud temática	e1- Exactitud de clasificación
	e2- Exactitud de un atributo cualitativo
	e3- Exactitud del valor dado a un atributo cuantitativo
Cuadro 1. Elementos y subelementos de calidad	

En la práctica, la aplicación de los lineamientos indicados en los subelementos de calidad es lo que permite medir y obtener el estado de la calidad de los datos. A continuación se describen los elementos y subelementos de calidad.

### a. El grado de totalidad

Este elemento de calidad exige que la totalidad de los objetos que se deben capturar según lo definido en las especificaciones o en el modelo de datos, deben estar presentes en la base de datos, es decir, el mapa debe estar completo. Así por ejemplo, para el plano cartográfico básico a escala 1:2000 de la ciudad Gótica se debieron capturar todas las vías; para medir el grado de totalidad del plano en este aspecto se determina el número total de objetos tipo vía existentes en el terreno y se cuentan los que están presentes en la base de datos, la diferencia nos dará el grado de precisión del plano. Al realizar dicha comparación se puede encontrar que faltan o sobran vías en el plano. Si sobran vías se cuantifica como un error de comisión (exceso de objetos en la base de datos). Si faltan vías se cuantifica como un error de omisión (falta de objetos en la base de datos).

<sup>5</sup> Editada por el Instituto Colombiano de Normas Técnicas y Certificaciones (Icontec) el 14 de mayo del 2002

B. *Consistencia lógica*: permite determinar la calidad de los datos en los siguientes aspectos: estructura interna, reglas de topología,<sup>6</sup> atributos y relaciones. Los objetos de calidad que se tienen en cuenta para medir la calidad en dicho aspecto son:

b1. *Consistencia de dominio*: se verifica que tanto los datos como sus atributos estén dentro de los rangos estipulados en las especificaciones del producto.

b2. *Consistencia topológica*: se verifica que los datos cumplan con las reglas topológicas establecidas en las especificaciones iniciales del producto. Los objetos geográficos digitales adquieren topología cuando son sometidos al proceso de estructuración, mediante el cual la información geográfica se recrea con todas sus características internas y externas en un computador; es decir, después del proceso de estructuración los objetos adquieren las características geométricas y atributos físicos que tienen en el terreno y estas características pueden ser conocidas por los usuarios mediante consultas a la información almacenada en las computadoras. Para la evaluación de este aspecto se tienen en cuenta las reglas topológicas, que son, entre otras, las siguientes: conectividad, adyacencia, continuidad, dirección, geometría, etcétera.

b3. *Consistencia de formato*: se revisa el formato digital en el que están almacenados los datos. Los formatos digitales permiten almacenar datos geográficos digitales tipo vector o *raster*, entre ellos se tienen los formatos DXF, DGN, DWG y otros que son nativos de cada programa. Para el almacenamiento y el manejo de información estructurada se cuenta con los formatos E00 para Arc\_info, Shape para Arc\_view. La información debe estar en el formato indicado en las especificaciones iniciales del producto.

C. *Exactitud de posición*: aplica a las posiciones verticales y horizontales y están en función de la escala para cual fueron capturados los vectores, los puntos y los polígonos.

Todos los objetos presentes sobre la superficie terrestre tienen una posición única con relación a los demás.

Los objetos se pueden ubicar, o georreferenciar; es decir, determinar su posición respecto de un sistema de referencia previamente asumido. (En el caso de que la posición de los objetos se dé en coordenadas geográficas latitud y longitud, el punto de partida, o en términos técnicos el marco de referencia, es el ecuador terrestre y el meridiano de Greenwich; en el caso de coordenadas planas de Gauss y concretamente en el caso de datos para Colombia (en Colombia se cuenta con cinco orígenes de coordenadas que abarcan franjas de 3 grados en sentido este-oeste, para cartografía a escalas pequeñas. Para cartografía a escala grande se utilizan proyecciones cartesianas locales).

c1. *Exactitud absoluta o externa*: se define como la exactitud en posición del conjunto de objetos del mapa respecto a su posición real en el terreno; es decir, es la diferencia existente entre las coordenadas que tienen los objetos en los archivos digitales o en el papel comparando con las coordenadas que estos mismos objetos tienen en el terreno.

c2. *Exactitud relativa o interna*: se busca determinar la consistencia interna del mapa, o sea que, que si en el terreno se mide una distancia X, dicho valor o uno muy cercano sea el que tiene la medida en el mapa.

c3. *Exactitud de posición de datos de celdas*: la precisión de los datos en formato *raster* está sujeta a la resolución espacial de la imagen fuente o tamaño de píxel en unidades métricas. En los programas utilizados para el manejo de imágenes, el píxel es referenciado por el valor de intersección entre una fila y una columna, pero la posición relativa del píxel está referida al centro de éste.

D. *Exactitud temporal*: la exactitud temporal hace referencia al período o períodos de tiempo por los cuales el conjunto de datos se corresponde con el terreno.

La información sobre la temporalidad de los datos es importante para todos los usuarios y la mayoría de ellos están interesados en la vigencia

<sup>6</sup> Rama de las matemáticas que estudia ciertas propiedades de las figuras geométricas.

de los datos en relación con las condiciones del terreno, o sea, cuándo o en qué fecha el mundo real lucía tal como es descrito en el juego de datos.

En la mayoría de casos se conoce la fecha en que la información fue recogida o publicada; sin embargo, para la exactitud temporal se requiere que el productor identifique las fechas de período asociado al contenido o a la condición del terreno; es decir, se debe citar las fechas en que se llevaron a cabo cada una de las actividades en terreno para la obtención del producto, como recolección de datos, toma de imágenes o aerofotografías, procesos de control terrestre y clasificación de campo, entre otros.

El hecho de conocer la exactitud temporal con alta confiabilidad tiene grandes ventajas, por ejemplo, en cartografía social, en mapeo para estudios arqueológicos, y en general para estudios ambientales, ya que estos estudios requieren detectar cambios de la naturaleza en el tiempo.

Los subelementos que permiten medir la calidad de los datos respecto a la temporalidad son:

- d1. Exactitud en el tiempo de medición:* este aspecto tiene que ver con la precisión en la medición o toma del dato de la fecha en el que ocurrió un evento, tiene su mayor aplicación en el registro de fenómenos meteorológicos tales como datos pluviométricos, gravimétricos, de temperatura, etcétera.
- d2. Consistencia temporal:* se refiere a la consistencia en las fechas de ocurrencia de los eventos o relación de los datos, en el caso de que éstos obedezcan a una secuencia u orden de ejecución. Por ejemplo, en la elaboración de un mapa la fecha de edición debe ser posterior a la de restitución.<sup>7</sup>
- d3. Validez temporal:* como su nombre lo indica, los datos deben ser válidos con respecto al tiempo. Es decir, el dato debe ser reportado sin ningún sesgo respecto a la fecha de captura o de actualización.

*E. Exactitud temática:* describe el grado de fidelidad de la clasificación correcta de los objetos, sus relaciones y atributos con respecto de su verdadera

<sup>7</sup> Se denomina restitución al proceso mediante el cual se capturan con altas precisiones los elementos presentes en una fotografía aérea a un archivo digital o a un papel.

característica presentada en el terreno.

Los subelementos que permiten realizar la medición en este aspecto son:

*e1. Exactitud de clasificación:* se debe determinar si, los elementos presentes en la base de datos fueron clasificados tal cual están en el terreno. Cuando se hace este análisis a una base de datos no es raro encontrar, senderos clasificados como drenajes sencillos, o polígonos que en el terreno son bosques y en la base de datos están codificados como humanales o pantanos.

*e2. Exactitud de un atributo cualitativo:* se trata de determinar si los atributos que tienen los objetos en la base de datos corresponden a los que éstos tienen en el terreno; ejemplo, una vía que en terreno está pavimentada en la base de datos, debe aparecer con un atributo que indique que es pavimentada.

#### 4. Exactitud del valor dado a un atributo cuantitativo.

Se trata de determinar si los atributos que tienen los objetos en la base de datos corresponden a los que éstos tienen en el terreno en valores cuantificables y su precisión; ejemplo: 100 cotas fotogramétricas presentan una diferencia en la cota de más de 10 cm.

*Presentación de los resultados de la calidad de un conjunto de datos*

Los resultados obtenidos de la medición de la calidad generalmente se presentan en porcentajes (%) de aceptación o de rechazo (conformidad o no conformidad); dichos porcentajes se hallan utilizando estadígrafos como el error medio cuadrático o simplemente obteniendo el cociente entre el número de objetos que cumplen (NOC) sobre el número de objetos total examinados (NOE).  

$$\frac{NOC}{NOE} = \% (d \text{ conformidad})$$
 con base en este porcentaje y el estándar definido en las especificaciones del producto, se define si el producto cumple o no.

Se tienen en cuenta los descriptores de un subelemento de calidad, que son:

- 1- Nivel de medición, 2- Indicador de la calidad, 3- Método de evaluación, 4- Resultado de la evaluación, 5- Tipo de valor de la evaluación, 6- Fecha de evaluación y 7- Reporte de calidad.

### 1. Nivel de medición de la calidad de los datos

El nivel de medición debe ser definido para cada subelemento de calidad aplicable. Teniendo en cuenta que la calidad puede variar dentro de un conjunto de datos, se deben considerar los siguientes aspectos:

**Tamaño de muestra:** se especifica si se examinó todo el conjunto de datos, o se examina una muestra, en este caso se debe especificar el porcentaje correspondiente.

**Tipo de objetos:** se deben explicitar los ítems u objetos a ser revisados (listas de tipo de objetos, atributos y relaciones entre objetos) que se revisarán.

**La extensión geográfica:** se consignarán las coordenadas extremas inferior izquierda y superior derecha del proyecto.

Se cita la extensión de tiempo (fecha de inicio y terminación de la evaluación)

### 2. Indicador de calidad de datos

Un indicador de calidad debe ser proporcionado para cada nivel de medición. Dicho indicador debe nombrar y describir brevemente el tipo de prueba que se aplica a los datos en un nivel de medición e incluir los parámetros límite o de frontera. Ejemplo: el umbral superior de error permitido es del 10% de incertidumbre para el caso del análisis por matriz de incertidumbre y menor a 1 metro para el caso del valor del error medio cuadrático.

### 3. Método de la evaluación de calidad

El método de evaluación de la calidad debe ser proporcionado para cada indicador y debe describir o referenciar la documentación correspondiente del procedimiento utilizado para obtener el indicador de calidad

### 4. Resultado de la evaluación de calidad

Un resultado de evaluación de la calidad debe ser proporcionado para cada indicador de ésta. Dicho resultado puede ser, por ejemplo, el valor o conjunto de valores obtenidos al medir un indicador de calidad. Se debe establecer si el dato está conforme o no con la especificación después de comparar el valor o conjunto de valores obtenidos al medir un indicador de calidad contra un valor mínimo aceptable.

### 5. Tipo de valor de calidad

Un tipo de valor de la calidad debe ser proporcionado para cada resultado de calidad de datos. El tipo de valor para un resultado ha de ser: cumple / no cumple.

### 6. Fecha de evaluación de la calidad de datos

Debe proporcionarse la fecha de evaluación de la calidad por cada indicador examinado. Esa fecha puede incluir un rango de tiempo si ello se considera relevante.

### 7. El reporte de calidad

El reporte de calidad debe ser generado por el productor de los datos y ser explícito, claro; el reporte requiere estar a disposición de todos los usuarios con el fin de que sirva como herramienta o guía para la utilización de los datos en aplicaciones específicas y se genera teniendo en cuenta básicamente los descriptores de los subelementos de calidad.

## Elementos generales de calidad

Los elementos generales de calidad son historia, propósito y uso de los datos; éstos describen aspectos cualitativos de la información.

### Propósito

El propósito de los datos debe ser registrado por el productor y esbozar claramente para qué fueron creados.

### Uso

El productor deberá hacer una actualización dinámica en el tiempo, respecto de todos los usos que los usuarios en general hayan dado, o estén dando a los datos.

### Historia

El productor deberá tener disponible un recuento del ciclo de vida de los datos, desde la colección, la edición y transformación, hasta la puesta a punto o forma actual. Se debe registrar con claridad la procedencia u origen de los datos, así como las diferentes etapas de transformación que tengan los datos durante su vida útil.

## Bibliografía

- [1] Francisco Javier Ariza López. *Calidad en la producción cartográfica*. Madrid, Ra-ma, 2002.
- [2] Disponible en: [http://www.mappinginteractivo.com/plantilla-ante.asp?id\\_articulo=642](http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=642)
- [3] Norma Técnica Colombiana (NTC) 5043.
- 

## EDILBERTO NIÑO NIÑO

Ingeniero catastral y geodesta de la Universidad Distrital Francisco José de Caldas, Bogotá. Especialista en Astronomía, de la Universidad Nacional de Colombia, Bogotá. Especialista en Sistemas de Información Geográfica (SIG), de la Universidad Distrital Francisco José de Caldas, Bogotá. Docente de tiempo completo en la Universidad Distrital Francisco José de Caldas, Bogotá. [edilbertonino@gmail.com](mailto:edilbertonino@gmail.com)