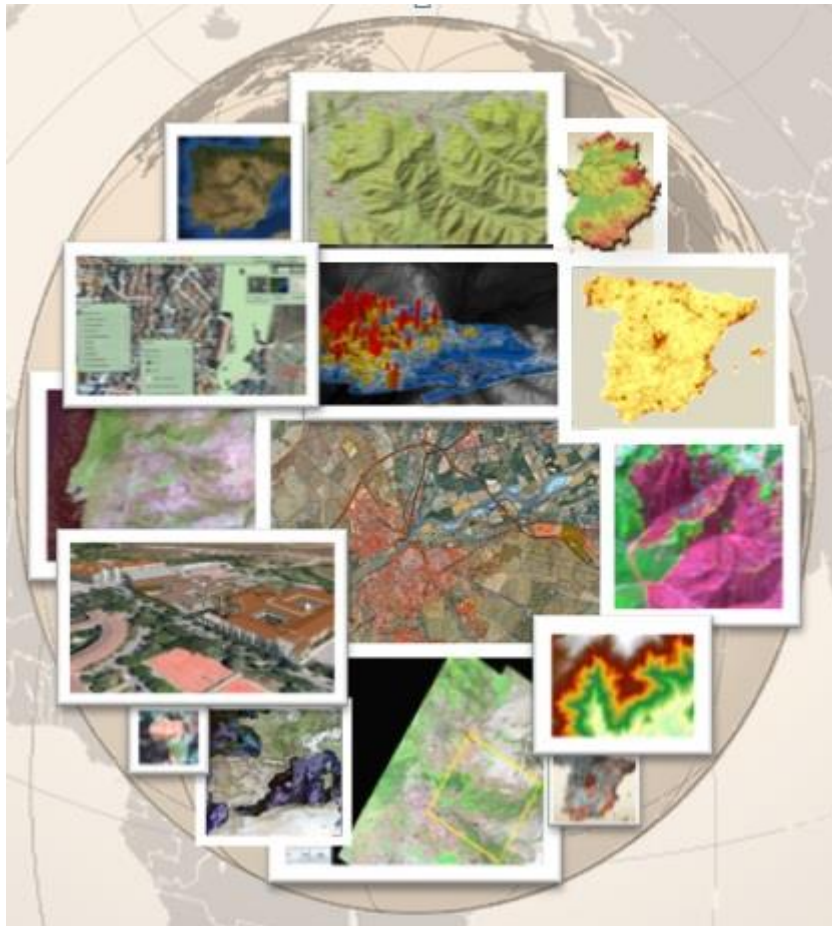


# Máster en Tecnologías de la Información Geográfica: SIG y Teledetección



## GEOESTADÍSTICA Y CALIDAD DE LA INFORMACIÓN

### Exactitud posicional

1.	CALIDAD DE LA INFORMACIÓN GEOGRÁFICA.....	2
2.	EXACTITUD POSICIONAL.....	3
3.	TEST PARA EL CONTROL DE LA EXACTITUD POSICIONAL .....	3
4.	LA NATURALEZA VECTORIAL DEL ERROR DE POSICIÓN.....	5
5.	BIBLIOGRAFÍA .....	6

## 1. Calidad de la Información Geográfica

El uso de los datos espaciales por parte de diferentes usuarios ha crecido de forma importante en las últimas décadas, tanto desde un punto de vista del volumen como de la solicitud de la exactitud requerida. Las innovaciones tecnológicas, el desarrollo de internet y las necesidades de empresas, organismos y de la sociedad en general en los llamados geoservicios, ha contribuido a ello. Además, los datos espaciales son usados por diferentes profesionales para fines diversos como planificación territorial, gestión de la agricultura o seguimiento y prevención de catástrofes, entre otros.

Se hace necesario, por tanto, conocer la calidad de los datos con los que estamos trabajando, aunque en general los metadatos sobre la calidad suelen ser escasos. Definimos un **elemento de calidad** como un aspecto de la calidad de los datos que puede ser medido cuantitativamente. Su propósito es identificar las inconsistencias de un conjunto de datos con respecto a las especificaciones técnicas [1].

Para describir los aspectos cuantitativos de calidad en datos espaciales se han descrito diferentes elementos. Una síntesis básica de los mismos, aunque con alguna variación según autores, sería la siguiente [1, 2] :

- **Exactitud posicional.** Responde a la pregunta de dónde está el objeto geográfico, cuantificando la proximidad entre las coordenadas medidas y las reales (o consideradas como reales).
- **Exactitud temática.** Está relacionada con los atributos y debe responder a la pregunta de qué es el objeto geográfico.
- La **exactitud temporal** nos debe aclarar cómo varía a lo largo del tiempo el objeto y a si está actualizada su información.
- La **consistencia lógica** se refiere a las reglas topológicas que debe cumplir cualquier base de datos espaciales.
- La **compleción** mide lo completa que es la base de datos representando todos los elementos del mundo real.
- El **linaje o genealogía** especifica de dónde viene la información y qué procesos y transformaciones ha sufrido.

Básicamente hay que contestar a las siguientes preguntas cuando hablamos de elementos de calidad de datos espaciales:

### Elementos de calidad de los datos espaciales

- **Posición** ¿dónde está?
- **Atributos** ¿qué es?
- **Temporalidad** ¿cómo varía a lo largo del tiempo?
- **Consistencia lógica** ¿sigue una estructura lógica?
- **Compleción** ¿están todos los datos necesarios para representar la realidad?
- **Linaje** ¿de dónde vienen los datos?

## 2. Exactitud posicional

En los apartados que siguen y, según conocimientos adquiridos en el tema primero, nos centramos principalmente en estudiar la **exactitud posicional**, definiéndola como el grado de conformidad con que se representan las posiciones horizontales y verticales de los objetos en una base de datos espaciales [2]. Este análisis se hará desde dos puntos de vista:

- Consideración del error de posición como un vector.
- Aplicación de test de control de exactitud posicional.

## 3. Test para el control de la exactitud posicional

Existen diferentes test o pruebas para realizar un control sobre la exactitud posicional de los elementos de una base de datos. Uno de los más conocidos es el *National Standard for Spatial Data Accuracy (NSSDA)*<sup>1</sup> desarrollado por el *Federal Geographic Data Committee* de los EE.UU.

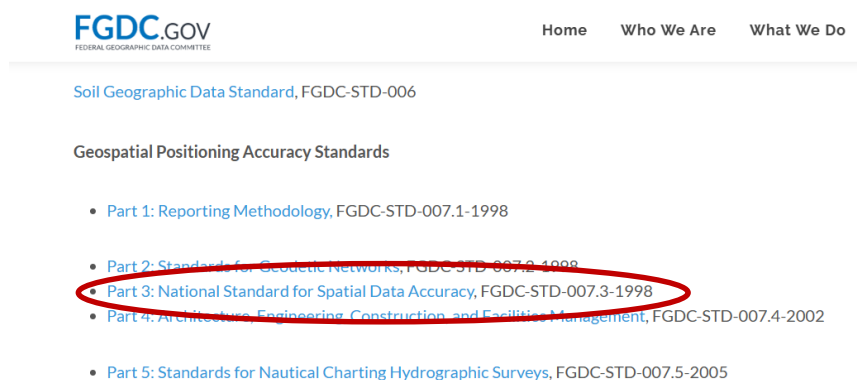


Figura 1. National Standard for Spatial Data Accuracy.

El **NSSDA** ofrece una metodología estadística para determinar la exactitud posicional de puntos de productos cartográficos y unificar criterios de exactitud proporcionando un lenguaje común para los usuarios de datos geográficos. Este test analiza la componente horizontal separada de la vertical. Para ello se selecciona una muestra de 20 puntos como mínimo, perfectamente definidos sobre el producto cartográfico y sobre una fuente de mayor exactitud. Se calculan los errores medio cuadráticos (EMC) de la muestra para las componentes X e Y y posteriormente para la Z, definiendo un coeficiente de exactitud posicional con un nivel de confianza generalmente del 95%. El EMC o RMSE (*Root Mean Squared Error*) evalúa la precisión de un modelo de regresión comparando las predicciones del modelo con los valores reales del conjunto de prueba.

Como resultado de la prueba NSSDA, junto al producto cartográfico, aparece la siguiente leyenda:

<sup>1</sup> [https://www.fgdc.gov/standards/standards\\_publications/index.html](https://www.fgdc.gov/standards/standards_publications/index.html)

**“Comprobado/compilado para X metros de exactitud horizontal/vertical al 95% de nivel de confianza”.**

Este texto indica que, con un nivel de confianza del 95%, la posición de los puntos de la base de datos espacial tendrá un error con respecto al valor convencionalmente verdadero del terreno igual o menor que el valor X indicado.

Esta norma no define un umbral de exactitud, en cada caso, según el producto a comprobar, es el usuario quien establece el límite de exactitud permitido. El proceso para aplicar el test es el siguiente:

- Se selecciona una muestra de 20 puntos como mínimo, perfectamente definidos, sobre una fuente de mayor exactitud y sobre el producto cartográfico a comprobar.
- Se restan los valores de las coordenadas X, Y y Z entre cada pareja de puntos homólogos de la fuente de mayor exactitud (valores considerados verdaderos) y la fuente a comprobar (valores medidos):

$$e_X = X_v - X_m \quad e_Y = Y_v - Y_m \quad e_Z = Z_v - Z_m$$

- Se calculan los errores medio cuadráticos (EMC) de la muestra para las componentes X e Y y posteriormente para la Z:

$$EMC_X = \sqrt{\frac{\sum_{i=1}^n e_{Xi}^2}{n}} \quad EMC_Y = \sqrt{\frac{\sum_{i=1}^n e_{Yi}^2}{n}} \quad EMC_Z = \sqrt{\frac{\sum_{i=1}^n e_{Zi}^2}{n}}$$

En el caso de la componente posicional horizontal se obtiene el valor NSSDA según si los valores de  $EMC_X$  y  $EMC_Y$  son iguales o distintos.

- Si son iguales, la exactitud posicional a un 95% de confianza es:

$$\text{Exactitud} = 1.7308 \cdot EMC_{XY}$$

Siendo,

$$EMC_{XY} = \sqrt{EMC_X^2 + EMC_Y^2}$$

- Si son distintos, la exactitud posicional a un 95% de confianza es:

$$\text{Exactitud} = 1.22385 \cdot (EMC_X + EMC_Y)$$

El valor del test para la componente vertical se obtiene por:

$$\text{Exactitud} = 1.96 \cdot EMC_Z$$

Otros test de procedimiento parecido son los test EMAS y NMAS, que han sido sustituidos por el NSSDA.

El **test EMAS** (*Engineering Map Accuracy Standard*) fue desarrollado por la *American Society of Photogrammetry and Remote Sensing* (ASPRS) junto con la *American Society of Civil Engineers* y el *American Congress on Surveying and Mapping* con la idea de proporcionar la exactitud de los mapas a gran escala.

El **test NMAS** (*Nacional Map Accuracy Standard*) fue utilizado por el USGS (*United States Geological Survey*) desde 1947 y se desarrolla comparando con fuentes de mayor exactitud. Los elementos analizados son puntos, tanto su componente horizontal como vertical; la X y la Y se analizan de forma conjunta.

Pero hay más opciones para determinar la exactitud posicional: fórmula de Koppe, normas para MDE, test ASPRS, estudios de exactitud posicional mediante bandas de error...[3].

Se han desarrollado otras técnicas para la mejora de la calidad posicional de los datos espaciales, como es la aplicación de estimadores robustos para determinar valores atípicos [4] y métodos basados en elementos lineales para realizar un control posicional en cartografía [5].

#### 4. La naturaleza vectorial del error de posición

El error de posición tradicionalmente se ha tratado como una magnitud escalar, con el cálculo del error medio cuadrático como estadístico básico. Si entendemos que el error de posición de un punto es la diferencia entre la posición real y la observada (en 2 o 3 dimensiones), tenemos un vector (en el plano o en el espacio) que no está definido completamente si solo señalamos su magnitud (el módulo del vector) siendo necesaria la indicación del acimut y el sentido del vector. En la figura se representan superpuestas dos cartografías a escala 1/25000 y 1/50000 de una misma zona. Los vectores (flechas negras) indican la diferencia entre las posiciones de algunos elementos comunes en ambas cartografías. Esta diferencia de posición implica una distancia y una dirección y sentido que habitualmente se obvia en el tratamiento del error.

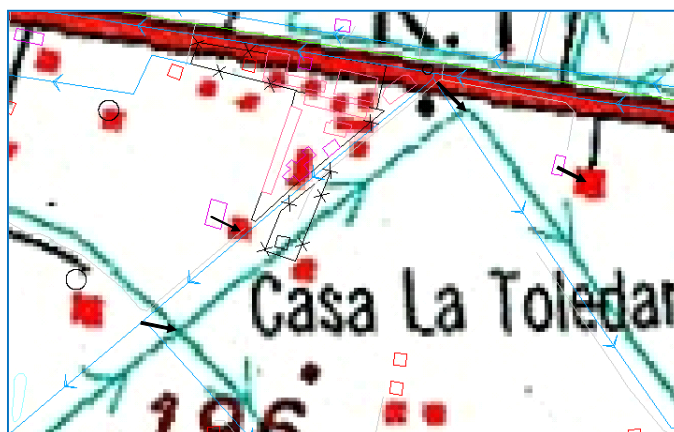


Figura 2. Ejemplo de error de posición en dos dimensiones en cartografía.

Hablamos de tuplas de coordenadas que definen la posición de un objeto en un espacio de referencia.

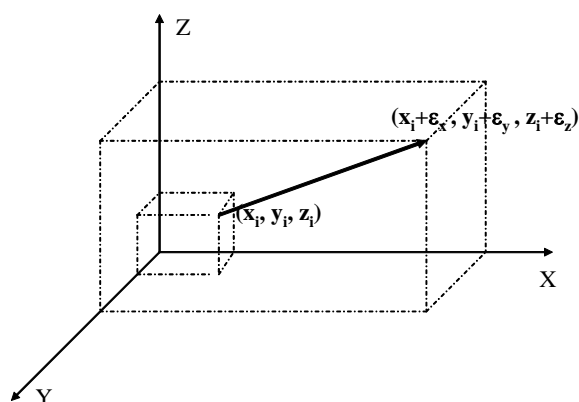


Figura 3. Naturaleza vectorial del error de posición.

Si consideramos un sistema de referencia cartesiano y un espacio tridimensional, las tuplas adoptan la forma  $(x_i, y_i, z_i)$  donde cada elemento de la tupla representa la posición respecto a uno de los ejes del sistema de referencia. Y como cada elemento de la tupla está afectado por un error,  $(x_i + e_x, y_i + e_y, z_i + e_z)$ , deberíamos conocer (o al menos acotar) los valores  $(e_x, e_y, e_z)$ .

Es decir, que una forma más completa de analizar el error de posición en cartografía es considerando su naturaleza vectorial [6] [7]. El vector de error, representado por una flecha, tiene unas propiedades como el módulo y acimut (que es un ángulo) que son propiedades métricas y el sentido que es una propiedad topológica.

Recordemos que el módulo del vector es un dato escalar y el acimut un dato circular [8] [9] y que ambos tipos de datos fueron analizados en el tema anterior obteniendo resultados como el que aparece en la Figura 4.

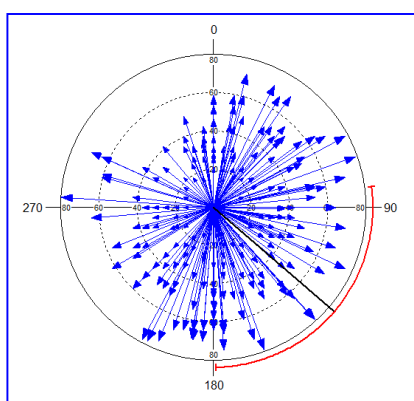


Figura 4. Gráfico que representa magnitudes vectoriales (módulo y ángulo) llevadas al mismo origen.

## 5. Bibliografía

- [1] ICONTEC, "Información Geográfica. Conceptos básicos de calidad.," ed. Santafé de Bogotá: Instituto Colombiano de Normas y Técnicas de Certificación, 2000, p. 39.
- [2] F. J. Ariza López, *Calidad en la Producción Cartográfica*. Madrid: RA-MA, 2002.
- [3] F. J. Ariza López, J. L. García Balboa, M. A. Ureña Camara, J. Rodríguez Avi, M. V. Alba Fernández, and F. J. Mesas Carrascosa, *Fundamentos de Evaluación de la Calidad de la Información Geográfica*: Publicaciones de la Universidad de Jaén, 2013.
- [4] A. D. J. Atkinson Gordo, F. J. Ariza López, and J. L. García Balboa, "Estimadores robustos: una solución en la estimación de valores atípicos para el control de calidad posicional," *Geofocus*, vol. 7, pp. 171-187, 2007.
- [5] A. T. Mozas Calvache and F. J. Ariza López, "Principales métodos de control posicional por elementos lineales. Comparación mediante su aplicación a líneas sintéticas.," *Geofocus*, vol. 8, pp. 187-204, 2008.
- [6] M. E. Polo and Á. M. Felicísimo, "Full Positional Accuracy Analysis of Spatial Data by Means of Circular Statistics," *Transactions in GIS*, vol. 14, pp. 421-434, 2010.
- [7] Á. M. Felicísimo, A. Cuartero, and M. E. Polo, "Analysis of homogeneity and isotropy of spatial uncertainty by means of GPS kinematic check lines and circular statistics," in *7th*

*International Symposium on Spatial Accuracy Assessment in Natural resources and Environmental Sciences*, Lisbon, 2006, pp. 85-90.

- [8] N. I. Fisher, *Statistical analysis of circular data*, 2nd ed. Cambridge, UK: Cambridge University Press, 1995.
- [9] N. I. Fisher, T. Lewis, and B. J. J. Embleton, *Statistical analysis of spherical data*. Cambridge: Cambridge University Press, 1987.