Especificación Técnica:

CALIDAD DE LA INFORMACIÓN GEOGRÁFICA

ideuy
Infraestructura de Datos Espaciales

Edición 01.00
2018-12

AVISO LEGAL

Este documento ha sido elaborado por la IDEuy (Infraestructura de Datos Espaciales del Uruguay).

Usted es libre de copiar, distribuir, comunicar y difundir públicamente este documento así como también de hacer obras derivadas, siempre y cuando tenga en cuenta citar la obra de forma específica y no utilizar esta obra para fines comerciales.

Contenido

Introducción 4
1 Objeto y alcance 6
2 Conformidad 6
3 Modelo de calidad y Medidas de la calidad 7
3.1 Generalidades 7
3.1.1 Visión general 7
3.1.2 Requisitos generales 8
3.2 Completitud de la IG 9
3.2.1 Comisión 9
3.2.2 Omisión 11
3.3 Consistencia de la IG: Consistencia Lógica 12
3.3.1 Consistencia conceptual 12
3.3.2 Consistencia de dominio 14
3.3.3 Consistencia de formato 15
3.3.4 Consistencia topológica 16
3.3.5 Consistencia temporal 21
3.4 Exactitud de la IG 22
3.4.1 Exactitud Posicional 22
3.4.2 Exactitud temática 47
3.4.3 Exactitud temporal 54
3.5 Usabilidad 56
Anexo A (normativo) Medidas básicas de la calidad de datos de errores 57
A.1 Generalidades 57
A.2 Variable aleatoria unidimensional, *Z* 57
A.2.1 Desviación típica conocida 58
A.2.2 Desviación típica desconocida 59
A.3 Variable aleatoria bidimensional, *X e Y* 61
A.4 Variable aleatoria tridimensional, *X, Y, Z* 62
Bibliografía 63

Introducción

La calidad es concebida a nivel general como "cumplimiento de requerimientos". De esto se deduce que, para un objeto de evaluación, tal como un conjunto de datos geográficos, **no se puede determinar "una única calidad" independiente del contexto de uso o aplicación del objeto de evaluación**, sino todo lo contrario: **la evaluación de la calidad de datos geográficos constituye una etapa fundamental para la determinación del grado de cumplimiento, por parte de los datos objeto de evaluación, de los requerimientos de los distintos actores involucrados en el ciclo de vida del dato**.

En el proceso de **producción** de un conjunto de datos geográficos, los productores realizan evaluaciones de su calidad para evaluar el cumplimiento de requerimientos establecidos en una especificación de requerimientos sobre dicho conjunto de datos. Estos requerimientos pueden ser "autoimpuestos" por el productor, o pueden ser definidos por un cliente/ usuario para satisfacer sus necesidades cuando el productor actúa como proveedor o prestador de servicios de disponibilización de estos conjuntos de datos.

Sin embargo, cada vez más los datos geográficos son compartidos, intercambiados y utilizados para propósitos diferentes de los inicialmente definidos por sus productores. En este sentido, y frente a la ingente cantidad de información geográfica, la evaluación de la calidad de los datos geográficos cumple un papel fundamental en la **valoración** de aquél conjunto de datos geográficos que mejor satisfaga las necesidades del usuario. Éste, con el fin de determinar cuál de los conjuntos de datos geográficos se adecúa mejor a sus requerimientos, evalúa la calidad de los distintos conjuntos de datos geográficos de acuerdo a su propósito de aplicación/ uso especifico, para luego comparar los resultados de las evaluaciones.

Con el objetivo de facilitar este proceso de evaluación de la calidad de datos geográficos por diferentes actores, **la presente Especificación Técnica establece no solo el Modelo de Calidad de Datos Geográficos al definir los elementos factibles de evaluación (Dimensiones, Subdimensiones y Factores), sino también un listado no exhaustivo de las medidas de la calidad más utilizadas de dichos elementos.** Estas medidas de la calidad están definidas en Tablas que incluyen su Nombre, Descripción, Función de medición y Ejemplo de aplicación nacional. Las medidas de la calidad especificadas se encuentran agrupadas en el Capítulo 3 de la presente Especificación de acuerdo a la Dimensión, Subdimensión –si existe-, y Factor que pretende evaluar.

Durante el proceso de evaluación del conjunto de datos geográficos se selecciona qué Factor se quiere evaluar de acuerdo a los requerimientos especificados (por ejemplo, Exactitud posicional), para luego seleccionar las medidas de la calidad, definidas para dicho Factor, que se desea aplicar. Dichas medidas de la calidad pueden ser seleccionadas de la presente Especificación Técnica, o ser elaboradas para los propósitos específicos de la evaluación, siempre y cuando éstas cumplan con los criterios establecidos en el Capítulo 8 *Data quality measures* de la Norma ISO 19157:2013 y pase las pruebas respecto a éstos definidas en el apartado A.5 *Test case identifier: Data quality measures*. Cumpliendo con estos criterios **se**

puede declarar conformidad con la presente Especificación, brindando mayores garantías al proceso de evaluación; para más información sobre los pasos para declarar conformidad con la esta Especificación Técnica véase el Capítulo 2 “Conformidad”.

En la elaboración de la presente Especificación Técnica se utilizó como antecedente base la Norma ISO 19157:2013 de la cual se seleccionaron aquellas medidas de la calidad que se consideraron de relevancia para la realidad nacional, modificándose también algunas de las seleccionadas y definiéndose, para cada una de ellas, uno o más ejemplos de aplicación nacional.

En este proceso participaron varias organizaciones que conformaron el Grupo de trabajo “Calidad de Información Geográfica” coordinado por la Infraestructura de Datos Espaciales del Uruguay (IDEuy); por más información sobre su integración véase el Informe al final de la presente Especificación.

1 Objeto y alcance

La presente Especificación Técnica **establece el Modelo de Calidad de Datos Geográficos** al definir las características de la calidad de los Datos Geográficos. Las mismas se clasifican en Factores, agrupados a su vez en Dimensiones y Subdimensiones (cuando aplica)[1].

Esta Especificación Técnica **también define, para cada Factor, una o más medidas de la calidad de datos geográficos** de las cuales se especifica su Nombre, Descripción, Función de medición y uno o más Ejemplos de aplicación nacional.

2 Conformidad

Esta Especificación Técnica presenta una serie de medidas de la calidad de datos geográficos a utilizar en la evaluación de los mismos. **La importancia de cada una de estas medidas de la calidad de datos geográficos varía para cada organización o proyecto particular**. Por este motivo, la implementación de esta Especificación Técnica generalmente implica la selección de los factores de la calidad y las medidas de la calidad de datos geográficos consideradas adecuadas a las necesidades y objetivos de la organización o proyecto particular. Por tanto, **para declarar conformidad con la presente Especificación Técnica deben seguirse los siguientes pasos**:

1. Especificar la Dimensión/ Subdimensión y los correspondientes Factores a evaluar de los datos geográficos, seleccionados de la presente Especificación;

2. De los Factores especificados:

a) seleccionar las medidas de la calidad a aplicar para la evaluación de los mismos, justificando su utilización, o,

b) elaborar otras medidas de la calidad que cumplan con los criterios establecidos en el Capítulo 8 *Data quality measures* de la Norma ISO 19157:2013 y pasen las pruebas respecto a éstos definidas en el apartado A.5 *Test case identifier: Data quality measures*, brindando justificación de su elaboración y uso.

3. Realizar las mediciones de la calidad acorde a las funciones de medición de las medidas de la calidad seleccionadas/ elaboradas para la evaluación de los Factores especificados.

4. Informar de los resultados de la calidad a través de los metadatos o de un informe independiente de la calidad.

[1] Por más información sobre estos conceptos véase el apartado 3.1.1 “Visión general”.

3 Modelo de calidad y Medidas de la calidad

3.1 Generalidades

3.1.1 Visión general

Todo Modelo de Calidad, cualquiera sea su objeto de definición (por ejemplo, Datos Geográficos), establece las características de la calidad a la vez que las organiza agrupándolas de acuerdo a elementos en común en: Dimensiones, Subdimensiones (cuando aplica) y Factores.

Los términos "Dimensión", "Subdimensión" y "Factor", son introducidos en esta Especificación Técnica provenientes del Modelo de Calidad de Datos de Gobierno de la Agencia para el Desarrollo del Gobierno de Gestión Electrónica y la Sociedad de la Información y del Conocimiento (AGESIC)[2]. Éstos definen los siguientes niveles de granularidad en la agrupación de características de la calidad:

- *Dimensión*: conjunto de características de calidad que refieren a una misma propiedad de alto nivel del dato. Por ejemplo, Exactitud de la Información Geográfica.
- *Subdimensión*: subconjunto de características de la calidad que pertenecen a una Dimensión determinada. Por ejemplo, la Exactitud Temática es una subdimensión de la Exactitud de la IG.

  NOTA: Solo se utiliza este nivel adicional de clasificación cuando la cantidad de factores amerita un nuevo agrupamiento.
- *Factor*: característica de calidad de una propiedad específica del dato. Por ejemplo, Corrección de la clasificación es un Factor comprendido en la Subdimensión "Exactitud Temática" de la Dimensión "Exactitud de la IG").

En el presente Capítulo se define el Modelo de Calidad Datos Geográficos Nacional (véase la **Figura** 1) junto con las medidas de la calidad para evaluarlo. Las Dimensiones definidas en el presente Modelo están alineadas con aquellas establecidas para el citado Modelo de Calidad de Datos en elaboración por AGESIC.

[2] Modelo en proceso de elaboración al momento de publicación de la presente Especificación.

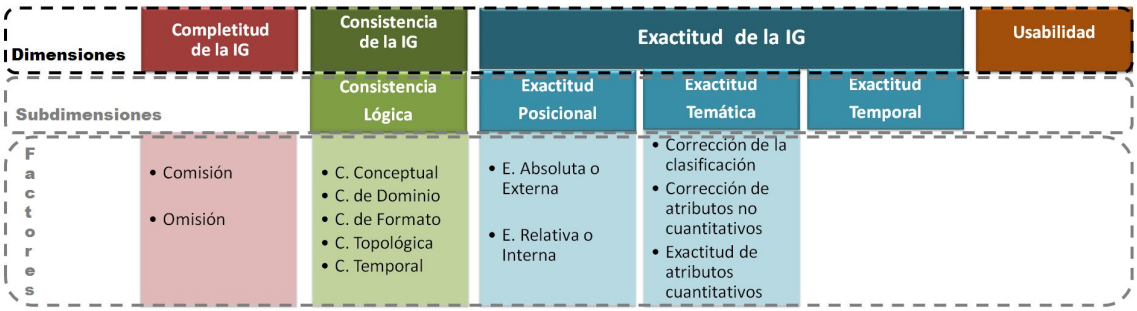


Figura *1*: Modelo de la Calidad de Datos Geográficos Nacional.

Cada una de las medidas de la calidad de datos geográficos especificadas en este Capítulo se presentan en forma de Tabla, indicando su Nombre, Descripción, Función de medición y uno o más Ejemplos de aplicación nacional. Las mismas están agrupadas por Factores que se agrupan en Subdimensiones (cuando existen) –subapartados del presente Capítulo-, las cuales a su vez se agrupan en Dimensiones (apartados del presente Capítulo). Estas medidas de la calidad constituyen un listado no exhaustivo de las más utilizadas a nivel nacional.

3.1.2 Requisitos generales

Debe explicitarse en el informe independiente de la calidad como se definen las variables, parámetros y mecanismos utilizados en cada medida de la calidad tal que el proceso sea reproducible en su totalidad

3.2 Completitud de la IG

Dimensión de la calidad de la información geográfica que define la presencia o ausencia de instancias de objetos geográficos en el conjunto de datos respecto al universo de discurso.

Los **Factores** de esta Dimensión son "Comisión" y "Omisión", y están definidos en los apartados 3.2.1 y 3.2.2 respectivamente.

3.2.1 Comisión

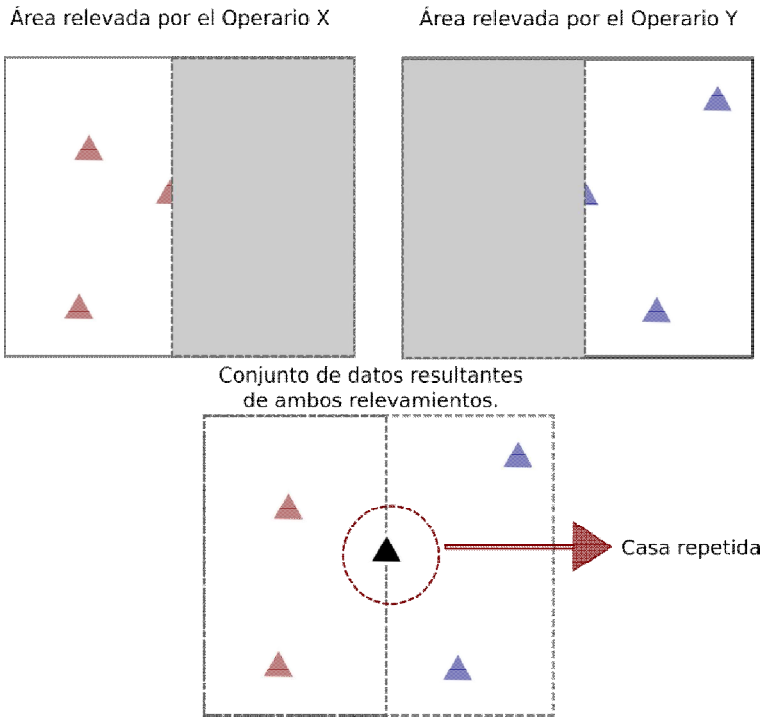
Factor de la Dimensión de la calidad de la información geográfica "Completitud de la IG" que **define** el grado en que las instancias de los objetos geográficos presentes en el conjunto de datos, considerando sus atributos y relaciones, no se encuentran en el universo de discurso.

Las medidas de la calidad utilizadas para medir este Factor están definidas en las **Tablas 1 y 2**.

Tabla 1: Índice de instancias de los objetos geográfico excedentes

Descripción	Porcentaje de instancias de los objetos geográficos presentes en el conjunto de datos sin existir en el universo de discurso, considerando sus atributos y relaciones.
Función de medición	$X = \frac{\lvert y_c - y_u \rvert}{y_u} . 100$ donde y_c es el total de instancias de los objetos geográficos presentes en el conjunto de datos; y_u es el total de instancias de los objetos geográficos presentes en el universo de discurso.
Ejemplo	Evaluando la Comisión del objeto geográfico "casa" en un conjunto de datos, se observa que, en el mismo hay 60 casas (y_c) mientras que en realidad (universo de discurso) existen 40 (y_u) lo que implica que hay un 50% más de casas en el conjunto de datos de las que existen (excedente). Por tanto, para este ejemplo, el Índice de casas excedentes es de un 50%.

Tabla 2: Número de instancias de los objetos geográficos repetidas

Descripción	Número de instancias de los objetos geográficos repetidas en un conjunto de datos, considerando sus atributos y relaciones.
Función de medición	*X* es el total de instancias de los objetos geográficos repetidas en un conjunto de datos, considerando sus atributos y relaciones.
Ejemplos	Dos operarios realizan un relevamiento de casas en una zona determinada. Éstos se dividen la zona a relevar en dos áreas obteniendo los siguientes resultados: el operario X identificó un total de 4 casas en su área asignada (rectángulo izquierdo), mientras que el operario Y identificó 4 casas en la otra área (rectángulo derecho). Al adicionar ambos conjuntos de datos se identifica un total de 8 casas cuando en la zona global relevada existen un total de 7. El número de casas excedentes en el conjunto de datos de este ejemplo, es por tanto, 1. 

3.2.2 Omisión

Factor de la Dimensión de la calidad de la información geográfica "Completitud de la IG" que **define** el grado en que las instancias de los objetos geográficos del universo de discurso no están en el conjunto de datos, considerando sus atributos y relaciones.

La medida de la calidad utilizada para medir este Factor está definida en la **Tabla 3**.

Tabla 3: Índice de instancias de los objetos geográficos omitidas

Descripción	Porcentaje de instancias de los objetos geográficos presentes en el universo de discurso no estando en el conjunto de datos, considerando sus atributos y relaciones.
Función de medición	$X = \frac{\lvert y_c - y_u \rvert}{y_u} . 100$ donde y_c es el total de instancias de los objetos geográficos presentes en el conjunto de datos; y_u es el total de instancias de los objetos geográficos presentes en el universo de discurso.
Ejemplo	En el conjunto de datos hay 80 casas (y_c) mientras que en realidad existen 110 (y_u) lo que implica que hay un 27% menos de casas en el conjunto de datos de las que existen (omisión). Por tanto, para este ejemplo, el Índice de casas omitidas es de 27 %.

3.3 Consistencia de la IG: Consistencia Lógica

La "**Consistencia de la IG"** constituye una de las **Dimensiones** de la calidad de la información geográfica que **define** el grado en que los objetos geográficos, atributos o relaciones están libres de contradicción y son coherentes en un contexto específico de uso.

En la presente Especificación técnica se presenta la **"Consistencia lógica"** como una **Subdimensión** de la Dimensión "Consistencia de la IG". La misma **se define como** el grado de incumplimiento del conjunto de datos con las reglas lógicas de las estructuras de datos, atributos y relaciones.

Los **Factores** de esta subdimensión definidos en los siguientes apartados son: **"Consistencia Conceptual" (apartado 3.3.1), "de Dominio" (apartado 3.3.2), "de Formato" (apartado 3.3.3), "Topológica" (apartado 3.3.4)** y **"Temporal" (apartado 3.3.5)**.

3.3.1 Consistencia conceptual

Factor de la Subdimensión de la calidad de la información geográfica "Consistencia Lógica" que **define** el grado de incumplimiento del conjunto de datos con las reglas del esquema conceptual correspondiente.

La medida de la calidad utilizada para medir este Factor está definida en la **Tabla** 4.

Tabla *4*: Índice de incumplimiento con las reglas del esquema conceptual

Descripción	Porcentaje de elementos del conjunto de datos que no cumplen con las reglas del esquema conceptual.
Función de medición	$X = \frac{y_e}{N} . 100$ donde y_e es el total de elementos del conjunto de datos que no cumplen con las reglas del esquema conceptual; N es el total de elementos del conjunto de datos.

NOTA: Se considera que un elemento no cumple las reglas del esquema conceptual, cuando, por ejemplo:

— existen emplazamientos no válidos de instancias de objetos geográficos a menos de una tolerancia determinada,

— existen superposiciones no válidas de instancias de objetos geográficos.

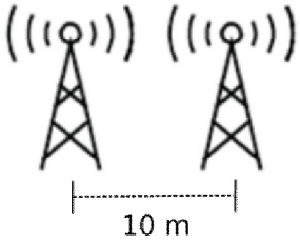
Ejemplos

Ejemplo de emplazamientos no válidos de objetos geográficos a menos de una tolerancia determinada:

Ejemplo de varias torres dentro de una tolerancia no válida

- *Regla del esquema conceptual:* Una torre no puede estar a una distancia de menos de 12 m de otra.

- *Universo de discurso*: Todas las torres existentes están a una distancia de al menos 12 m.



Representación de dos torres del conjunto de datos a una distancia de 10 m.

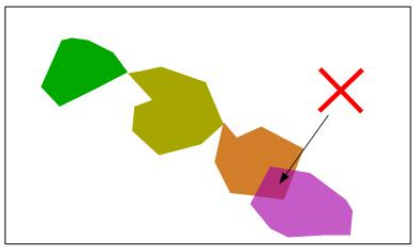
- *Conjunto de datos:* De un total de 15 torres del conjunto de datos se detectaron 3 a menos de 12 m de distancia, existiendo por tanto un 20% de incumplimiento con las reglas del esquema conceptual.

Ejemplo de superposición no válida de objetos geográficos:

Ejemplo de Planes de Uso de Suelos parcialmente solapados

- *Regla del esquema conceptual:* Un Plan de Uso del Suelo (PDU) no puede estar total o parcialmente solapado con otro.

- *Universo de discurso*: Ningún PDU está solapado con otro.



Representación de dos PDU solapados.

- *Conjunto de datos:*
De un total de 1.500 PDU se detectaron que 40 de ellos se encontraban solapados, existiendo por tanto, un 2,66% de incumplimiento con las reglas del esquema conceptual.

3.3.2 Consistencia de dominio

Factor de la Subdimensión de la calidad de la información geográfica "Consistencia Lógica" que **define** el grado en que los valores de los atributos del conjunto de datos se encuentran comprendidos dentro del dominio de valores permitidos.

La medida de la calidad utilizada para medir este Factor está definida en la **Tabla 5**.

Tabla 5: Índice de conformidad con el dominio de valores permitidos

Descripción	Porcentaje de atributos del conjunto de datos que son conformes con el dominio de valores permitido.
Función de medición	$$X = \frac{y_d}{N} . 100$$ donde y_d es el total de atributos del conjunto de datos que tienen valores comprendidos dentro del dominio de valores permitidos; N es el total de atributos del conjunto de datos.
Ejemplo	En el catálogo de objetos del conjunto de datos el objeto "Plan de Uso de Suelos" (PDU) tiene el atributo "cultivo de invierno" el cual puede tomar los siguientes valores (dominio de valores permitidos): Barbecho, Cebada, Cereales de invierno, Colza, Cultivo de cobertura, Pastura Consociada, Pastura no consociada, Pasturas, Trigo. En el conjunto de datos hay 1.500 PDU de los cuales únicamente 50 tienen asignados valores no incluidos dentro del dominio de valores del atributo "cultivo de invierno" (por ejemplo, de dichos 50, 7 tienen el valor "Soja"). Por lo tanto, para este ejemplo, el Índice de conformidad con el dominio de valores del atributo "cultivo de invierno" del objeto geográfico PDU para el conjunto de datos evaluado es de 96,66%.

3.3.3 Consistencia de formato

Factor de la Subdimensión de la calidad de la información geográfica "Consistencia Lógica" que define el grado en el que los datos se almacenan de acuerdo con la estructura física del conjunto de datos.

La medida de la calidad utilizada para medir este Factor está definida en la **Tabla *6***.

Tabla 6: Índice de conflictos en la estructura física

Descripción	Porcentaje de elementos del conjunto de datos que están almacenados en conflicto con la estructura física del conjunto de datos.
Función de medición	$$X = \frac{y_f}{N} . 100$$ donde y_f es el total de elementos del conjunto de datos que están almacenados en conflicto con la estructura física; *N* es el total de elementos del conjunto de datos.
Ejemplo	Dentro del conjunto de datos, el objeto "Plan de Uso de Suelos" debe tener un identificador único ("Id plan") con el tipo de dato "integer" (número entero). De un total de 1.500 PDU se encontraron 30 con valor "string" (valor textual) del atributo "Id plan". Por lo tanto el porcentaje de "ID plan" (atributo del objeto "PDU") almacenado en conflicto con la estructura física definida para dicho atributo es de 2%.

3.3.4 Consistencia topológica

Factor de la Subdimensión de la calidad de la información geográfica "Consistencia Lógica" que **define** el grado en que las características topológicas codificadas explícitamente son correctas.

Las medidas de la calidad utilizadas para medir este Factor están definidas en las **Tablas 7 a 12.**

Tabla 7: Índice de fallos en conexiones de nodos de enlace

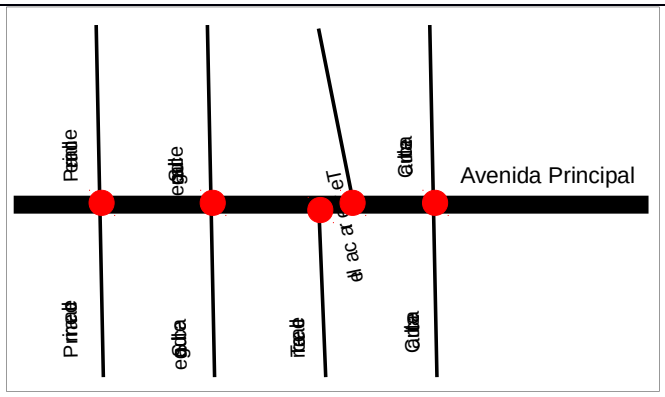
Descripción	Porcentaje de fallos en las conexiones de nodos de enlace en el total de conexiones de nodos de enlace.
Función de medición	$$X = \frac{y_{fn}}{N} . 100$$ donde y_{fn} es el total de fallos en las conexiones de nodos de enlace; N es el total de conexiones de nodos de enlace.
NOTA 1: En algunos casos los fallos en las conexiones de nodos de enlace se deben a que el sistema coloca automáticamente conexiones donde no existe justificación espacial en función de una limitación de vértices programada en el código del software. **NOTA 2**: En el Informe independiente de la calidad **debe** estar explicitado qué se considera un fallo.	
Ejemplo	 *-Conjunto de datos*: Hay un total de 2 fallos en las conexiones de nodos de enlace. *- Universo de discurso:* Existe un total de 4 conexiones de nodo de enlace (*N*). Conclusión: se presenta un 50% de fallos en las conexiones de nodo de enlace.

Tabla 8: Índice de conexiones omitidas por subtrazos

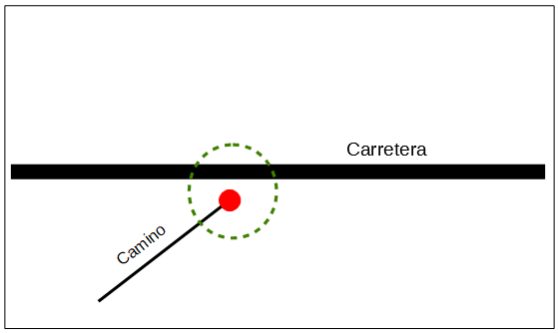
Descripción	Porcentaje de conexiones omitidas por subtrazos.
Función de medición	$$X = \frac{y_{su}}{N}.100$$ donde y_{su} es el total de conexiones omitidas por subtrazos; N es el total de conexiones en el universo de discurso.
Ejemplo	 *Representación de una conexión entre dos carreteras omitida en el conjunto de datos por un subtrazo.* - *Conjunto de datos*: Hay un total de 5 conexiones. - *Universo de discurso:* Existen 7 conexiones (*N*). - *Conclusión:* Comparando el conjunto de datos con el universo de discurso se identifica que 2 conexiones no figuran en el conjunto de datos debido a subtrazos (y_{su}). Por lo tanto el Índice de conexiones omitidas por subtrazos, para este ejemplo, es de 28,6 % (*X*).

Tabla 9: Índice de conexiones omitidas por sobretrazos

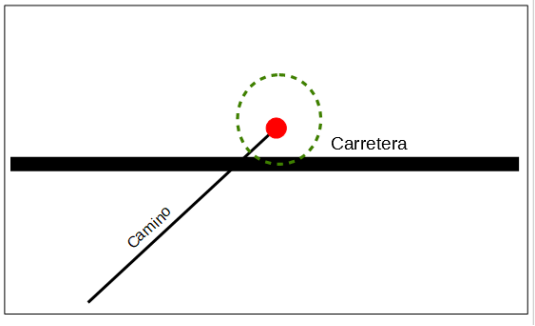
Descripción	Porcentaje de conexiones omitidas por sobretrazos
Función de medición	$$X = \frac{y_{so}}{N} . 100$$ donde y_{so} es el total de conexiones omitidas por sobretrazos; N es el total de conexiones en el universo de discurso.
Ejemplo	 *Representación de una conexión omitida en el conjunto de datos por un sobretrazo.* - *Conjunto de datos*: Hay un total de 5 conexiones. - *Universo de discurso:* Existen 7 conexiones (*N*). - *Conclusión:* Comparando el conjunto de datos con el universo de discurso se identifica que 2 conexiones no figuran en el conjunto de datos debido a sobretrazos (y_{so}). Por lo tanto el Índice de conexiones omitidas por sobretrazos, para este ejemplo, es de 28,6 % (*X*).

Tabla 10: Número de polígonos ficticios no válidos.

Descripción	Número de elementos del conjunto de datos que son polígonos ficticios no válidos.
Función de medición	X es el número de polígonos ficticios no válidos.
Ejemplo	Existe un total de 20 polígonos ficticios no válidos en el conjunto de datos evaluado.

Tabla 11: Índice de errores por autointersecciones no válidas

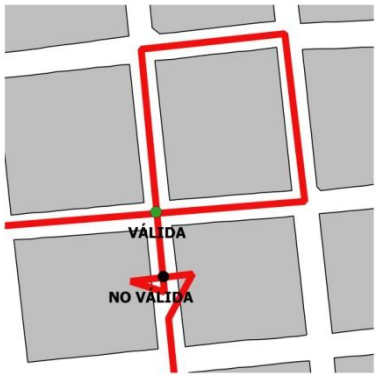
Descripción	Porcentaje de errores por autointersecciones no válidas.
Función de medición	$$X = \frac{y_a}{N}.100$$ donde y_a es el total de autointersecciones no válidas; N es el total de autointersecciones.
Ejemplo	 *Representación de una autointersección no válida y una válida. Una autointersección no válida puede producirse, por ejemplo, cuando se releva una trayectoria con GPS, al detenerse en un punto.*

Tabla 12: Índice de errores por autosuperposiciones no válidas

Descripción	Porcentaje de errores por autosuperposiciones no válidas.
Función de medición	$$X = \frac{y_{as}}{N}.100$$ donde y_{as} es el total de autosuperposiciones no válidas; N es el total de autosuperposiciones.
Ejemplo	*Leyenda:* • *vértices*

3.3.5 Consistencia temporal

Factor de la Subdimensión de la calidad de la información geográfica "Consistencia Lógica" que **define** el grado en que es correcto el orden de los eventos.

La medida de la calidad utilizada para medir este Factor está definida en la ***Tabla 13***

Tabla *13*.

Tabla 13: Índice de eventos ordenados de forma cronológicamente correcta

Descripción	Porcentaje de eventos ordenados de forma cronológicamente correcta.
Función de medición	$$X = \frac{y_{ev}}{N} . 100$$ donde y_{ev} es el total de eventos ordenados de forma cronológicamente correcta; N es el total eventos ordenados cronológicamente.
Ejemplo	En el catálogo de objetos del conjunto de datos el objeto "Plan de Usos de Suelos" tiene los atributos "Fecha de presentación" y "Fecha de modificación". La fecha de modificación no puede ser anterior a la fecha de presentación. De un total de 1.500 PDU se encontraron 30 con fecha de modificación anterior a la fecha de presentación. Por lo tanto el 98% de los eventos están ordenados cronológicamente de forma correcta.

3.4 Exactitud de la IG

Dimensión de la calidad de la información geográfica que **define** el grado en que los atributos de las instancias de los objetos geográficos representan correctamente el valor verdadero o considerado como tal en un contexto específico de uso.

NOTA: Los valores pueden ser obtenidas por relevamiento directo en campo ("valor verdadero") o a partir de un conjunto de datos de mejor calidad que el que se desea evaluar ("valor considerado verdadero").

En la presente Especificación se definen la **"Exactitud Posicional"** (apartado 3.4.1), la "**Exactitud Temática**" (apartado 3.4.2) y la "**Exactitud Temporal**" (apartado 3.4.3) como **subdimensiones** de la Dimensión "Exactitud de la IG".

3.4.1 Exactitud Posicional

Subdimensión de la Dimensión "Exactitud de la Información Geográfica" que **define** el grado de exactitud de la posición de los objetos geográficos en un determinado sistema de referencia espacial (SRS, *Spatial Reference System*).

Los **Factores** de esta subdimensión definidos en los apartados 3.4.1.1 a 3.4.1.3 son "**Exactitud Absoluta o Externa**", "**Exactitud Relativa o Interna**" y "**Exactitud Posicional de Datos en malla**", respectivamente.

3.4.1.1 Exactitud Absoluta o Externa

Factor de la Subdimensión "Exactitud posicional" que **define** el grado de proximidad de los valores de las coordenadas reportados con los valores verdaderos o aceptados como tales.

Las **medidas** de la calidad utilizadas para medir la "Exactitud Posicional absoluta o externa" **se clasifican en tres tipos**:

- "Medidas Generales de Errores Posicionales" (apartado 3.4.1.1.1, Tabla 14 a 16);
- "Medidas de Errores Posicionales Verticales" (apartado 3.4.1.1.2, Tablas 17 a 19);
- "Medidas de Errores Posicionales Horizontales" (apartado 3.4.1.1.3, **Tabla *20*** a 23).

3.4.1.1.1 Medidas generales de errores posicionales

Para todas las medidas generales de errores posicionales **debe** fijarse el criterio por el cual se define la correspondencia entre puntos homólogos, el cual **debe** ser explicitado en el resultado de la evaluación.

Tabla 14: Valor medio de los errores posicionales

Descripción	Valor medio de los errores posicionales para un conjunto de puntos (1D, 2D, 3D)
Función de medición	***1. Error posicional (e_i)*** 1D: $e_i = \left\| x_{mi} - x_{ti} \right\|$ 2D: $e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2}$ 3D: $e_i = \sqrt{(x_{mi} - x_{ti})^2 + (y_{mi} - y_{ti})^2 + (z_{mi} - z_{ti})^2}$ donde $x_{mi}; y_{mi}; z_{mi}$ son las posiciones medidas en 1D, 2D y 3D, respectivamente; x_{ti}, y_{ti}, z_{ti} son las posiciones verdaderas (*true*) en 1D, 2D y 3D, respectivamente. ***2. Valor medio de los errores posicionales absolutos o externos ($\bar{e}$)*** $\bar{e} = \frac{1}{N} \sum_{i=1}^{N} e_i$ donde e_i es el error posicional de cada punto; N es el número de puntos del conjunto de puntos que se está evaluando.

Ejemplo

Punto	x_m	y_m	x_t	y_t	$1D_x:e_x$	$1D_y:e_y$	$2D:e_{2D}$
1	553719,420	6155639,764	553716,833	6155628,358	2,587	11,406	11,696
2	555561,309	6152550,857	555558,890	6152544,819	2,419	6,038	6,505
3	546787,035	6157841,302	546789,733	6157833,684	2,698	7,618	8,082
Suma					7,704	25,062	26,282
N					3	3	3
ē					**2,568**	**8,354**	**8,761**

Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.

Dadas las coordenadas de 3 puntos (x_t;y_t) consideradas como verdaderas[1] y sus correspondientes obtenidas del conjunto de datos espaciales que se desea evaluar (x_m;y_m) el Valor medio del error posicional en X es $\bar{e}_x$**= 2,568 metros**; el Valor medio del error posicional en Y es $\bar{e}_y$ **= 8,354 metros** y el Valor medio del error posicional para la planimetría es $\bar{e}_{2D}$**= 8,761 metros**.

Tabla 15: Sesgo de las posiciones

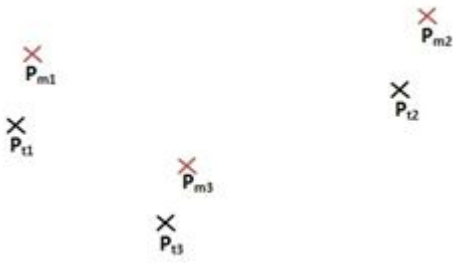
Descripción	Sesgo de las posiciones para un conjunto de puntos.
Función de medición	***1. Desviaciones individuales (e_{xi} ; e_{yi} ; e_{zi})*** $e_{xi} = x_{mi} - x_{ti}$ $e_{yi} = y_{mi} - y_{ti}$ $e_{zi} = z_{mi} - z_{ti}$ donde x_{mi} ; y_{mi} ; z_{mi} son las posiciones medidas en 1D, 2D y 3D, respectivamente; x_{ti} ; y_{ti} ; z_{ti} son las posiciones verdaderas (*true*) en 1D, 2D y 3D, respectivamente. ***2. Sesgo (a_x ; a_y ; a_z)*** $a_x = \frac{\sum e_{xi}}{N_x}$ $a_y = \frac{\sum e_{yi}}{N_y}$ $a_z = \frac{\sum e_{zi}}{N_z}$ donde e_{xi} ; e_{yi} ; e_{zi} son las desviaciones individuales de posiciones en 1D, 2D y 3D, respectivamente; N_x ; N_y ; N_z son el número de puntos en 1D, 2D y 3D, respectivamente. ***3. Sesgo de las posiciones*** $a_p = \sqrt{a_x^2 + a_y^2}$ $a_{3D} = \sqrt{a_x^2 + a_y^2 + a_z^2}$ donde a_x ; a_y ; a_z son los sesgos en 1D, 2D y 3D.

Ejemplo

Punto	x_m	y_m	x_t	y_t	e_x	e_y
1	553719,420	6155639,764	553716,833	6155628,358	2,587	11,406
2	555561,309	6152550,857	555558,890	6152544,819	2,419	6,038
3	546787,035	6157841,302	546789,733	6157833,684	-2,698	7,618
Suma				2,308	25,062	
N				3	3	
a				**0,768**	**8,354**	
a_p					**8,389**	
Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.						

Dadas las coordenadas de 3 puntos (x_t;y_t) consideradas como verdaderas[1] y sus correspondientes obtenidas del conjunto de datos espaciales que se desea evaluar (x_m;y_m) el Sesgo para la componente X es a_x**= 0,768 metros**; el Sesgo para la componente Y es a_y **= 8,354 metros** y el Sesgo para la planimetría es a_p**= 8,389 metros**.

Conclusión: En promedio el conjunto de datos geográficos evaluados se encuentran desplazados de la posición considerada como verdadera 8,389 metros.



Leyenda:

P_m es el punto medido,

P_t es el punto verdadero (*true*).

Tabla 16: Índice de errores posicionales mayores a un umbral

Descripción	Porcentaje de errores posicionales mayores a un umbral dado para un conjunto de posiciones
Función de medición	***1. Error posicional (e_i)*** 1D: $e_i = \lvert x_{mi} - x_{ti} \rvert$ 2D: $e_i = \sqrt{(x_{mi}-x_{ti})^2+(y_{mi}-y_{ti})^2}$ 3D: $e_i = \sqrt{(x_{mi}-x_{ti})^2+(y_{mi}-y_{ti})^2+(z_{mi}-z_{ti})^2}$ donde x_{mi}, y_{mi}, z_{mi} son las posiciones medidas en 1D, 2D y 3D, respectivamente; x_{ti}, y_{ti}, z_{ti} son las posiciones verdaderas (*true*) en 1D, 2D y 3D, respectivamente. ***2. Recuento de errores posicionales mayores a un umbral (A)*** $A = \sum a_i$ **Condiciones:** ▪ Si $e_i > e_{max}$, entonces a_i=1, ▪ Si $e_i \leq e_{max}$, entonces a_i=0. donde e_i es el error posicional; e_{max} es el umbral. **NOTA 1:** El umbral es definido de acuerdo al objetivo de uso del conjunto de datos. **NOTA 2**: Debe explicitarse el umbral definido.

3. Porcentaje de errores posicionales mayores a un umbral (b)

$$b = (A / N) . 100$$

donde

A es el total de errores posicionales mayores a un umbral;

N es el total de errores posicionales.

Ejemplo

Punto	x_m	y_m	x_t	y_t	$1D_x:e_x$	$1D_y:e_y$	$2D:e_{2D}$
1	553719,420	6155639,764	553716,833	6155628,358	2,587	11,406	11,696
2	555561,309	6152550,857	555558,890	6152544,819	2,419	6,038	6,505
3	546787,035	6157841,302	546789,733	6157833,684	-2,698	7,618	8,082

Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.

Se plantea la evaluación de un conjunto de datos con una escala nominal de 1/40000. Para esta medida es necesario definir un error máximo. Si consideramos como límite el límite visual de ¼ de milímetro, el error planimétrico en metros aceptable para un punto sería de 10 metros. Por lo tanto definimos el e_{max} **= 10,000 metros**.

Punto	**Condición**
1	1
2	0
3	0
Resultado	
A	1
N	3
b	**33,33%**

Conclusión: el 33,33% de los puntos tiene un error en distancia mayor al umbral definido.

NOTA: En el Informe de la evaluación debe indicarse cómo se define el umbral

3.4.1.1.2 Medidas de errores posicionales verticales

Entre las medidas de errores posicionales verticales se encuentra la "**Exactitud lineal"** a diferentes niveles de significación. La misma **constituye** la mitad de la longitud del intervalo, definido por un límite superior y otro inferior, en que se sitúa el valor verdadero con una probabilidad determinada (por ejemplo, 50%, 68,3%). **Se calcula** utilizando como medidas básicas las definidas en el apartado A.2. La **Tabla 18** constituye una tabla de correspondencia entre el nombre de la medida específica de Exactitud lineal, su nivel de significación y las medidas básicas a partir de las que se construyen.

Tabla 17: Error cuadrático medio

<table>
<tr><td>Descripción</td><td>Error cuadrático medio</td></tr>
<tr><td>Función de medición</td><td>$$\sigma_z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_{mi} - z_t)^2}$$
donde
z_t es el valor verdadero de una medida Z;
z_{mi} es el valor medido de una medida Z.</td></tr>
<tr><td>Ejemplo</td><td>
<table>
<tr><th>Punto</th><th>z_m</th><th>z_t</th><th>e_z</th><th>e_z^2</th></tr>
<tr><td>1</td><td>13,205</td><td>13,934</td><td>-0,729</td><td>0,531</td></tr>
<tr><td>2</td><td>5,860</td><td>5,303</td><td>0,557</td><td>0,310</td></tr>
<tr><td>3</td><td>7,722</td><td>7,370</td><td>0,352</td><td>0,124</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td colspan="4">σ_z</td><td>0,567</td></tr>
<tr><td colspan="5">Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.</td></tr>
</table>
</td></tr>
</table>

	Dadas las coordenadas de 3 puntos de una medida *Z* consideradas como verdaderas (z_t;) y sus correspondientes obtenidas del conjunto de datos espaciales que se desea evaluar (z_m) el Valor cuadrático medio es 0,567. **Conclusión:** El valor cuadrático medio del conjunto de datos geográficos evaluados es de 0,567.

Tabla 18: Medidas básicas aplicables a las medidas de Exactitud lineal

Nombre de la medida de la calidad	**Nivel de significación**	**Medidas básica**		
		Nombre	***Función de medición***	***Ejemplo*** *(Véase NOTA 1)*
Error lineal probable	50%	LE50 o LE50®	$u_{50\%} \cdot \sigma_z$ donde $u_{50\%} = 0{,}6745$ σ_z es la desviación típica	0,383
Error lineal típico	68,3 %	LE68,3 o LE68,3®	$u_{68,3\%} \cdot \sigma_z$ donde $u_{68,3\%} = 1$ σ_z es la desviación típica	0,567
Exactitud lineal al 90% del nivel de significación	90%	LE90 o LE90®	$u_{90\%} \cdot \sigma_z$ donde $u_{90\%} = 1{,}645$ σ_z es la desviación típica	0,933
Exactitud lineal al 95% del nivel de significación	95%	LE95 o LE95®	$u_{95\%} \cdot \sigma_z$ donde $u_{95\%} = 1{,}960$ σ_z es la desviación típica	1,112
Exactitud lineal al 99% del nivel de significación	99%	LE95 o LE95®	$u_{99\%} \cdot \sigma_z$ donde $u_{99\%} = 2{,}576$ σ_z es la desviación típica	1,461
Error lineal cierto	99,8%	LE99 o LE99®	$u_{99,8\%} \cdot \sigma_z$ donde $u_{99\%} = 2{,}576$ σ_z es la desviación típica	1,702

NOTA 1: Todos los resultados están dados en metros y constituyen ejemplos para una desviación típica (σ_z) de 0.567, calculada a partir de los datos de ejemplo y función de medición brindada en la **¡Error! No se encuentra el origen de la referencia.**.

Tabla 19: Error lineal absoluto al 90% del nivel de significación de datos con sesgo

Descripción	Exactitud vertical absoluta de las coordenadas de los datos, expresadas en términos de error lineal al 90% de probabilidad y con existencia de sesgo.
Funciones de medición	***1. Error absoluto en la dimensión vertical en cada punto*** (δV_i) $$\delta V_i = \delta V_{fi} - \delta V_{ci}$$ donde δV_{fi} es la coordenada verdadera del punto considerado (*i*); δV_{ci} es la coordenada control del punto considerado (*i*). ***2. Valor medio absoluto del sesgo (***$\acute{\delta V}$ v $$\left\lvert\overline{\delta V}\right\rvert = \left\lvert\frac{1}{N}\sum_{i=1}^{N}\delta V_i\right\rvert$$ donde δV_i es el error absoluto en la dimensión vertical en cada punto; N es el número de puntos que se consideraron.

3. Desviación típica lineal de los errores absolutos en la dimensión vertical (σ_m)

$$\sigma_m = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \delta V_i^{\,2}}$$

donde

δV_i es el error absoluto en la dimensión vertical en cada punto;

N es el número de puntos que se consideraron.

4. Desviación típica lineal de los errores absolutos en la dimensión vertical de la fuente de referencia (σ_r)

Se calcula σ_r siguiendo la función de medición presentada en el literal 3.

5. Desviación típica lineal de los errores en el producto evaluado (σ_V)

$$\sigma_V = \sqrt{\sigma_m^2 + \sigma_r^2}$$

donde

σ_m es la desviación típica lineal de los errores absolutos en la dimensión vertical;

σ_r es la desviación típica lineal de los errores absolutos en la dimensión vertical de la fuente de referencia.

6. Ratio del valor absoluto del error medio entre la deviación típica (r)

$$r = \frac{|\overline{\delta V}|}{\sigma_V}$$

donde

$|\overline{\delta V}|$ es el valor medio absoluto del sesgo;

σ_V es la desviación típica lineal de los errores en el producto evaluado.

7. Exactitud vertical absoluta al 90% del nivel de significación de datos verticales sesgados

Alterantiva nº1 de cálculo:

- Condición: Si $r > 1{,}4$, entonces LMAS = σ_V . [1,282 + r]
- Condición: Si $r \leq 1{,}4$,
  entonces LMAS = σ_V . [1,6435 + 0,92 . r^2 – 0,28 . r^2]

Alternativa nº2 de cálculo:

- Condición: Si $r > 1{,}4$, entonces $k = 1{,}2815$
- Condición: Si $r \leq 1{,}4$, entonces se calcula k, el ratio del sesgo vertical entre la desviación típica de las altitudes, usando un ajuste polinómico cúbico mediante los valores tabulados definidos en *Handbook of Tables for Probability and Statistics* (Referencia[20]).

3.4.1.1.3 Medidas de errores posicionales horizontales

Entre las medidas de errores posicionales horizontales se encuentra la "**Exactitud circular"** a diferentes niveles de significación. La misma **constituye** el radio que describe un círculo, en el cual se sitúa la localización verdadera del punto con una probabilidad determinada (por ejemplo, 39,4%, 50%). **Se calcula** utilizando como medidas básicas las definidas en el apartado A.3.3. La **Tabla 20** constituye una tabla de correspondencia entre el nombre de la medida específica de Exactitud circular, su nivel de significación y las medidas básicas a partir de las que se construyen.

Tabla 20: Medidas básicas aplicables a las medidas de Exactitud circular con distintos niveles de significación.

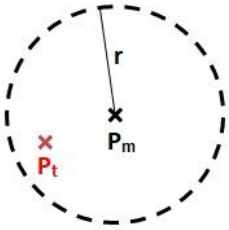
Nombre de la medida de la calidad	**Nivel de significación**	**Medidas básicas**		
		Nombre	***Función de medición***	***Ejemplo (véase NOTA al pie de la Tabla)***
Exactitud circular estándar	39,4%	CE39,4	$\frac{1}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	1,595
Exactitud circular probable	50%	CE50	$\frac{1,1774}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	1,878
Exactitud circular al 90% del nivel de significación	90%	CE90	$\frac{2,146}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	3,422
Exactitud circular al 95% del nivel de significación	95%	CE95	$\frac{2,4477}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	3,904
Exactitud circular al 99,8% del nivel de significación	99,8%	CE99,8	$\frac{3,5}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	5,582

NOTA 1: A continuación se detalla la situación de ejemplo a partir de la cual se calcularon los ejemplos de resultados de errores circulares a distintos niveles de significación (5ª columna de la Tabla 20):

A partir de las coordenadas de 3 puntos consideradas como verdaderas (x_t;y_t) y sus correspondientes obtenidas del conjunto de datos espaciales que se desea evaluar (x_m;y_m) presentadas en la siguiente tabla, se calculan las Exactitudes circulares a distintos niveles de

significación indicadas en la quinta columna de la Tabla anterior.

Punto	x_m	y_m	x_t	y_t	$1D_x:e_x$	$1D_y:e_y$
1	553719,420	6155639,764	553716,833	6155628,358	2,587	11,406
2	555561,309	6152550,857	555558,890	6152544,819	2,419	6,038
3	546787,035	6157841,302	546789,733	6157833,684	2,698	7,618
σ_{ex}					0,114690889	
σ_{ey}					2,252425064	
$\sqrt{{\sigma_{ex}}^2 + {\sigma_{ey}}^2}$					2,2553431 8	
Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.						



La posición verdadera del punto se encuentra dentro de un círculo con centro en el punto del conjunto de datos y radio *r*, para una probabilidad *P*. Donde *r* es el resultado obtenido para la exactitud circular para dicha probabilidad *P*.

Tabla 21: Error cuadrático medio de la planimetría

<table>
<tr><td>Descripción</td><td>Radio de un círculo en torno al punto dado, en el cual se sitúa el valor verdadero.</td></tr>
<tr><td>Función de medición</td><td>$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[(x_{mi} - x_t)^2 + (y_{mi} - y_t)^2\right]}$$
donde
x_{mi} ; y_{mi} son los puntos medidos en 1D y 2D, respectivamente;
x_{ti} ; y_{ti} son los puntos verdaderos (true) en 1D y 2D, respectivamente;
n es la cantidad de puntos considerados para la evaluación.</td></tr>
<tr><td>Ejemplo</td><td>Dadas las coordenadas de 3 puntos consideradas como verdaderas (x_t;y_t) y sus correspondientes obtenidas del conjunto de datos espaciales que se desea evaluar (x_m;y_m), presentadas en la siguiente tabla, se calcula que el Error cuadrático medio de la planimetría es σ= 9,026 metros.
<table>
<tr><th>Punto</th><th>x_m</th><th>y_m</th><th>x_t</th><th>y_t</th></tr>
<tr><td>1</td><td>553719,420</td><td>6155639,764</td><td>553716,833</td><td>6155628,358</td></tr>
<tr><td>2</td><td>555561,309</td><td>6152550,857</td><td>555558,890</td><td>6152544,819</td></tr>
<tr><td>3</td><td>546787,035</td><td>6157841,302</td><td>546789,733</td><td>6157833,684</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td colspan="5">σ= 9,026</td></tr>
<tr><td colspan="5">Las unidades de las coordenadas están dadas en metros y en el mismo sistema de coordenadas.</td></tr>
</table></td></tr>
</table>

Tabla 22: Error circular absoluto al 90% del nivel de significación de datos con sesgo

Descripción	Exactitud horizontal absoluta de las coordenadas de los datos, expresada en términos de error circular al 90% de probabilidad y con existencia de sesgo.
Función de medición	***1. Error absoluto en la dimensión horizontal en cada punto y en cada coordenada X_i e Y_i (δX_i; δY_i)*** $$\delta X_i = \delta X_{fi} - \delta X_{ci}$$ $$\delta Y_i = \delta Y_{fi} - \delta Y_{ci}$$ donde δX_{fi}; δY_{fi} son las coordenadas verdadera del punto considerado (*i*); δX_{ci}; δY_{ci} son las coordenadas control del punto considerado (*i*). ***2. Valor medio del sesgo de cada coordenada X e Y ($\acute{\delta X}$*** $$\overline{\delta X} = \frac{1}{N}\sum_{i=1}^{N} \delta X_i$$ $$\overline{\delta Y} = \frac{1}{N}\sum_{i=1}^{N} \delta Y_i$$ donde δX_i; δY_i son los errores absolutos en la dimensión horizontal en cada coordenada *X* e *Y*; N es el número de puntos que se consideraron.

3. Desviación típica circular de las diferencias que se han medido entre el producto evaluado y el utilizado como referencia (σ_{CM})

$$\sigma_{CM} = \sqrt{\frac{1}{2(N-1)}\left(\sum_{i=1}^{N}\left(\delta Xi - \overline{\delta X}\right)^2 + \sum_{i=1}^{N}\left(\delta Xi - \overline{\delta X}\right)^2\right)}$$

donde

$\delta X_i; \delta Y_i$	son los errores absolutos en la dimensión horizontal en cada coordenada *X* e *Y*;
$\overline{\delta X}; \overline{\delta Y}$	son los valores medio del sesgo de las coordenadas *X* e *Y*;
N	es el número de puntos que se consideraron.

4. Desviación típica circular de los errores en la fuente de referencia (σ_{CR})

Se calcula σ_{CR} siguiendo la función de medición presentada en el literal 3.

5. Desviación típica circular de los errores en el producto evaluado (σ_C)

$$\sigma_{C=}\sqrt{\sigma_{CM}^2 + \sigma_{CR}^2}$$

donde

σ_{CM}	es la desviación típica circular de los errores en el producto evaluado;
σ_{CR}	es la desviación típica circular de los errores en la fuente de referencia.

	6. Error circular absoluto al 90% de nivel de confianza de datos con sesgo (CMAS) $$CMAS = \sigma_C \cdot \left[1{,}2943 + \sqrt{\left(\frac{\overline{\delta X}^2 + \overline{\delta Y}^2}{\sigma_C} \right)} + 0{,}7254 \right]$$ donde $\overline{\delta X}$; $\overline{\delta Y}$ son los valores medio del sesgo de las coordenadas *X* e *Y*; σ_C es la desviación típica circular de los errores en el producto evaluado.

Tabla 23: Elipse de incertidumbre

Descripción	Elipse 2D cuyos dos ejes principales indican la dirección y magnitud de la mayor y menor incertidumbre de un punto 2D.
Función de medición	A partir de una Matriz de covarianzas dada de las coordenadas de un punto 2D, los elementos que describen la elipse de incertidumbre pueden determinarse mediante los valores propios. Para un solo punto *k*, la Matriz de covarianzas viene dada por $$\Sigma^{k}{}_{xx} = \begin{bmatrix} \sigma^2_{x_k} & \sigma_{x_k y_k} \\ \sigma_{y_k x_k} & \sigma^2_{y_k} \end{bmatrix}, \text{ con } \sigma_{x_k y_k} = \sigma_{y_k x_k}$$ La dirección α (orientación) del semieje mayor de la elipse de incertidumbre puede calcularse mediante $$\phi = \frac{1}{2}\arctan\frac{2\sigma_{x_k y_k}}{\sigma^2_{x_k} - \sigma^2_{y_k}}$$ y $$a = \sqrt{\frac{1}{2}\left(\sigma^2_{x_k} + \sigma^2_{y_k} + \sqrt{\left(\sigma^2_{x_k} - \sigma^2_{y_k}\right)^2 + 4\sigma^2_{x_k y_k}}\right)}$$ $$b = \sqrt{\frac{1}{2}\left(\sigma^2_{x_k} + \sigma^2_{y_k} - \sqrt{\left(\sigma^2_{x_k} - \sigma^2_{y_k}\right)^2 + 4\sigma^2_{x_k y_k}}\right)}$$

3.4.1.1.4 Exactitud relativa o interna

Factor de la Subdimensión de la calidad de la información geográfica "Exactitud posicional" que **define** el grado de proximidad de las posiciones relativas de objetos geográficos de un conjunto de datos con sus respectivas posiciones relativas verdaderas o aceptadas como tales.

Este Factor utiliza el mismo conjunto de medidas de la calidad del Factor Exactitud absoluta o externa (véase el apartado 3.4.1.1) teniendo como única diferencia el método de evaluación.

Las medidas de la calidad utilizadas para medir el Error posicional vertical relativo o interno y el Error posicional horizontal relativo o interno se presentan en las **Tablas 25 a 26**, respectivamente.

Tabla 24: Error posicional vertical relativo o interno

Descripción	Errores aleatorios en la posición vertical de un objeto geográfico en relación a otro del mismo conjunto de datos o en el mismo mapa/carta. Depende de los errores aleatorios en las dos elevaciones respecto a un *datum* vertical común.
Función de medición	Se calcula la diferencia entre los datos fuente (medidos) con los datos control (verdaderos) de la siguiente manera: 1. **Cálculo de todas las combinaciones de parejas posibles entre puntos** $$CombinacionesdeParesdePuntos = m = \frac{n(n1)}{2}$$ donde *n* es el tamaño de la muestra. **2. Cálculo del error vertical absoluto en cada punto** $$\Delta Z_i = h_{mi} - h_{ti} \quad \text{para } i = 1...n$$ donde h_{mi} es la altura medida del punto *i*, h_{ti} es la atura verdadera (*true*) del punto *i.*

	3. Cálculo del error vertical relativo para todas las combinaciones de pares de puntos $\Delta Z_{relkj} = \Delta Z_k - \Delta Z_j$ para k= 1...m, j= k + 1,...m **3. Cálculo de la desviación estándar vertical relativa** $\sigma_{Zrel} = \sqrt{\frac{\sum \Delta Z_{rel}^2}{m-1}}$ **4. Cálculo del error vertical relativo pasando en nivel de confianza al 90%** Rel LE90 = 1,645 . $\sigma_{Z\ rel}$

Tabla 25: Error posicional horizontal relativo o interno

Descripción	Errores aleatorios en la posición horizontal de un objeto geográfico en relación a otro del mismo conjunto de datos o en el mismo mapa/carta
Función de medición	Se calcula la diferencia entre los datos fuente (medidos) con los datos control (verdaderos) de la siguiente manera: **1. Cálculo de todas las combinaciones de parejas posibles entre puntos** $$CombinacionesdeParesdePuntos = m = \frac{n(n1)}{2}$$ donde *n* es el tamaño de la muestra. **2. Cálculo del error absoluto en las dimensiones *X* e *Y* de cada punto** $\Delta X_i = X_{mi} - X_{ti}$ para *i* = 1...*n* $\Delta Y_i = Y_{mi} - Y_{ti}$ para *i* = 1...*n* donde $X_{mi}; Y_{mi}$ son los valores medidos de las dimensiones *X* e *Y* del punto *i*, $X_{ti}; Y_{ti}$ son los valores medidos de las dimensiones *X* e *Y* del punto *i*. **3. Cálculo del error relativo de *X* e *Y* para todas las combinaciones de pares de puntos** $\Delta X_{relkj} = \Delta X_k - \Delta X_j$ para *k*= 1...*m*, *j*= *k* + 1,...*m* $\Delta Y_{relkj} = \Delta Y_k - \Delta Y_j$ para *k*= 1...*m*, *j*= *k* + 1,...*m*

	4. Cálculo de la desviación estándar relativa en cada dimensión $\sigma_{Xrel} = \sqrt{\frac{\sum \Delta X_{rel}^2}{m-1}}$ $\sigma_{Xrel} = \sqrt{\frac{\sum \Delta Y_{rel}^2}{m-1}}$ **5. Cálculo de la desviación estándar horizontal relativa** $\sigma_{H\,rel} = \sqrt{\frac{\sigma_{X\,rel}^{\,2} + \sigma_{Y\,rel}^{\,2}}{2}}$ **6. Cálculo del error horizontal relativo pasando el nivel de confianza al 90%** Rel CE90 = 2,146 . $\sigma_{H\,rel}$

3.4.1.2 Exactitud posicional de datos en malla

Factor de la subdimensión de la calidad de la información geográfica "Exactitud posicional" que **define** el grado de proximidad de los valores de las posiciones espaciales de datos en malla con los valores verdaderos o aceptados como tales.

Este Factor utiliza el mismo conjunto de Medidas de incertidumbres posicionales horizontales (véase el apartado 3.4.1.1.3) del Factor Exactitud absoluta o externa. Los valores ráster pueden describirse usando la Exactitud de atributos cuantitativos (véase el apartado 3.4.2.3).

3.4.2 Exactitud temática

Subdimensión de la Dimensión "Exactitud de la Información Geográfica" que **define** el grado de exactitud de los atributos cuantitativos y la corrección tanto de los atributos no cuantitativos como de las clasificaciones de los objetos geográficos y sus relaciones.

Los **Factores** de la subdimensión "Exactitud temática" definidos en los apartados 3.4.2.1 a 3.4.2.3 son "Corrección de la clasificación", "Corrección de los atributos no cuantitativos" y "Exactitud de los atributos cuantitativos", respectivamente.

3.4.2.1 Corrección de la clasificación

Factor de la subdimensión de la calidad de la información geográfica "Exactitud temática" que **define** si son correctas las clasificaciones realizadas comparando las clases asignadas al objeto geográfico o a sus atributos con el universo de discurso.

Las **medidas** de la calidad de este Factor se definen en las **Tablas 27 a 30.**

Tabla 26: Índice de clasificaciones correctas

Descripción	Porcentaje de instancias de un objeto geográfico clasificadas correctamente.
Función de medición	$X = \frac{y_{cc}}{N} . 100$ donde y_{cc} es el total de instancias de un objeto geográfico clasificadas correctamente; N es el total de instancias de un objeto geográfico presentes en el universo de discurso.
Ejemplos	En el conjunto de datos conformado por 100 instancias del objeto geográfico "represa" (generadas por la clasificación de imágenes satelitales) se encontró que 5 estaban mal clasificadas (no eran represas). Por lo tanto el 95% de las clasificaciones en este conjunto de datos son correctas.
	De los 8.500 caminos departamentales de todo el país, 8.100 fueron clasificados como "camino departamental". Por lo tano el 95 % de los caminos fueron clasificados correctamente.

Tabla 27: Matriz de confusión

<table>
<tr><td>Descripción</td><td>Matriz que indica el número de elementos de la clase (i) clasificados como clase (j).</td></tr>
<tr><td>Función de medición</td><td>Matriz cuadrada de n columnas y n filas, donde n indica el número de clases consideradas.

MCM (i,j) = [nº de ítems de la clase (i) clasificados como clase (j)]

Los elementos de la diagonal de la matriz de confusión contienen los ítems clasificados correctamente, y los elementos fuera de la diagonal contienen el número de errores de la clasificación.</td></tr>
<tr><td>Ejemplo</td><td>Clasificaciones para caminos={Camino departamental, Ruta Primaria, Ruta Secundaria}

<table>
<tr><td></td><td colspan="5">Conjunto de datos</td></tr>
<tr><td rowspan="5">Clase verdadera</td><td></td><td>Camino departamental</td><td>Ruta primaria</td><td>Ruta secundaria</td><td>Total</td></tr>
<tr><td>Camino departamental</td><td>8.100</td><td>200</td><td>200</td><td>8.500</td></tr>
<tr><td>Ruta primaria</td><td>20</td><td>105</td><td>25</td><td>150</td></tr>
<tr><td>Ruta secundaria</td><td>55</td><td>120</td><td>200</td><td>375</td></tr>
<tr><td>Total</td><td>8.175</td><td>425</td><td>425</td><td>9.025</td></tr>
</table>
</td></tr>
</table>

Tabla 28: Matriz de confusión relativa

<table>
<tr><td>Descripción</td><td>Matriz que indica el porcentaje de elementos de la clase (i) clasificados como clase (j).</td></tr>
<tr><td>Función de medición</td><td>Matriz cuadrada de n columnas y n filas, donde n indica el número de clases consideradas.

RMCM (i,j) = [nº de ítems de la clase (i) clasificados como clase (j)] / nº de ítems de la clase (i) . 100</td></tr>
<tr><td>Ejemplo</td><td>Teniendo en cuenta la matriz de confusión del ejemplo de la Tabla 28, su matriz de confusión relativa queda como sigue:
<table>
<tr><td></td><td colspan="4">Conjunto de datos</td></tr>
<tr><td rowspan="4">Clase verdadera</td><td></td><td>Camino departamental</td><td>Ruta primaria</td><td>Ruta secundaria</td><td>Total</td></tr>
<tr><td>Camino departamental</td><td>95%</td><td>2,5 %</td><td>2,5%</td><td>100%</td></tr>
<tr><td>Ruta primaria</td><td>13,3%</td><td>70%</td><td>16,7%</td><td>100%</td></tr>
<tr><td>Ruta secundaria</td><td>14,6%</td><td>32%</td><td>53,4%</td><td>100%</td></tr>
</table>
</td></tr>
</table>

Tabla 29: Coeficiente Kappa

Descripción	Coeficiente que cuantifica la proporción de acuerdo de las asignaciones a clases eliminando los errores de la clasificación
Función de medición	Con los elementos de la Matriz de confusión - MCM(*i,j*)- correspondiente a la medida de la Tabla 28 se puede calcular el Coeficiente kappa (*K*) como: $$\kappa = \frac{N \cdot \sum_{i=1}^{r} \mathrm{MCM}(i,i) - \sum_{i=1}^{r}\left(\sum_{j=1}^{r} \mathrm{MCM}(i,j) \cdot \sum_{j=1}^{r} \mathrm{MCM}(j,i)\right)}{N^2 - \sum_{i=1}^{r}\left(\sum_{j=1}^{r} \mathrm{MCM}(i,j) \cdot \sum_{j=1}^{r} \mathrm{MCM}(j,i)\right)}$$ donde *N* es el número de clases consideradas.
Ejemplo	Teniendo en cuenta la matriz de confusión del ejemplo de la Tabla 28, su coeficiente Kappa es: K=0,523 .

3.4.2.2 Corrección de atributos no cuantitativos

Factor de la subdimensión de la calidad de la información geográfica "Exactitud temática" que define si son correctos los valores de los atributos no cuantitativos de un objeto geográfico.

La medida de la calidad de este Factor se define en la **Tablas 31.**

Tabla 30: Índice de valores de atributos no cuantitativos correctos

Descripción	Porcentaje de valores correctos de un atributo no cuantitativo.
Función de medición	$X = \frac{y_{cor}}{N} . 100$ donde y_{cor} es el total de instancias de un atributo no cuantitativo correctas, N es el total de instancias de un atributo no cuantitativo.
Ejemplo	De 150 instancias del objeto geográfico "camino departamental" 100 tienen asignado correctamente su "nombre" (atributo no cuantitativo) en el conjunto de datos, por lo que, el 67% de los caminos departamentales tienen su nombre correcto en el conjunto de datos.

3.4.2.3 Exactitud de atributos cuantitativos

Factor de la subdimensión de la calidad de la información geográfica “Exactitud temática” que define que tan correctos son los valores de los atributos cuantitativos de un objeto geográfico.

Entre las medidas de la Exactitud de atributos cuantitativos se encuentra la **Incertidumbre del valor del atributo** a diferentes niveles de significación. La misma constituye la mitad de la longitud del intervalo, definido por un límite superior y otro inferior, en que se sitúa el valor verdadero del atributo cuantitativo con una probabilidad determinada (por ejemplo, 50%, 68,3%). Se calcula utilizando como medidas base las definidas en el apartado A.3.2. La Tabla 32 constituye una Tabla de correspondencia entre el nombre de la medida específica de la Incertidumbre del valor del atributo, su nivel de significación y las medidas base a partir de la que se construyen.

Tabla 31: Medidas básicas aplicables a las medidas de Incertidumbre del valor del atributo con distintos niveles de significación.

Nombre de la medida de la calidad	Nivel de significación	Medidas básicas	
		Nombre	***Función de medición***
Incertidumbre del valor del atributo al 50% del nivel de significación	50%	LE50 o LE50®	$u_{50\%} \cdot \sigma_z$
Incertidumbre del valor del atributo al 68,3% del nivel de significación	68,3%	LE68,3 o LE68,3®	$u_{68,3\%} \cdot \sigma_z$
Incertidumbre del valor del atributo al 90% del nivel de significación	90%	LE90 o LE90®	$u_{90\%} \cdot \sigma_z$
Incertidumbre del valor del atributo al 95% del nivel de significación	95%	LE95 o LE95®	$u_{95\%} \cdot \sigma_z$
Incertidumbre del valor del atributo al 99% del nivel de significación	99%	LE99 o LE99®	$u_{99\%} \cdot \sigma_z$
Incertidumbre del valor del atributo al 99,8% del nivel de significación	99,8%	LE99,8 o LE99,8®	$u_{99,8\%} \cdot \sigma_z$

3.4.3 Exactitud temporal

Subdimensión de la Dimensión "Exactitud de la Información Geográfica" que **define** el grado de exactitud de las mediciones de tiempo.

La Exactitud temporal se evalúa con diferentes niveles de significación. La misma constituye la mitad de la longitud del intervalo, definido por un límite superior y otro inferior, en que se sitúa el valor verdadero de la instancia de tiempo con una probabilidad determinada (por ejemplo, 50%, 68,3%). Se calcula utilizando como medidas base las definidas en el apartado A.3.2. La Tabla 33 constituye una Tabla de correspondencia entre el nombre de la medida específica de la Exactitud temporal, su nivel de significación y las medidas base a partir de la que se construyen.

Tabla 32: Medidas básicas aplicables a las medidas de Exactitud temporal con distintos niveles de significación.

Nombre de la medida de la calidad	**Nivel de significación**	**Medidas básicas**		
		Nombre	***Función de medición***	***Ejemplo (véase Nota al pie de la tabla)***
Exactitud temporal al 50% del nivel de significación	50%	LE50 o LE50®	$u_{50\%} \cdot \sigma_z$	3,87
Exactitud temporal al 68,3% del nivel de significación	68,3%	LE68,3 o LE68,3®	$u_{68,3\%} \cdot \sigma_z$	5,74
Exactitud temporal al 90% del nivel de significación	90%	LE90 o LE90®	$u_{90\%} \cdot \sigma_z$	9,44
Exactitud temporal al 95% del nivel de significación	95%	LE95 o LE95®	$u_{95\%} \cdot \sigma_z$	11,25
Exactitud temporal al 99% del nivel de significación	99%	LE99 o LE99®	$u_{99\%} \cdot \sigma_z$	14,79
Exactitud temporal al 99,8% del nivel de significación	99,8%	LE99,8 o LE99,8®	$u_{99,8\%} \cdot \sigma_z$	17,22

NOTA 1: A continuación se detalla la situación de ejemplo a partir de la cual se calcularon los ejemplos de resultados de errores circulares a distintos niveles de significación (5ª columna de la **Tabla *32***):

Se presentan 6 eventos de vientos de más de 60km/h en el mes de agosto de 2017.

Para cada evento perteneciente al conjunto de datos se analiza el tiempo real de ocurrencia y el tiempo registrado:

n	t real	t registrado	t registrado - t real	(t registrado - t real)²
1	ago-17	may-17	-3	9
2	ago-17	ago-17	0	0
3	ago-17	ago-17	0	0
4	ago-17	jun-17	-2	4
5	ago-17	abr-17	-4	16
6	ago-17	sep-18	13	169

Siguiendo la función de medición presentada en el apartado A.2.2 se obtiene que
St= 5,74$meses$.

3.5 Usabilidad

La usabilidad se basa en los requerimientos de usuario. Todos los elementos de calidad pueden usarse para evaluar la usabilidad. La evaluación de la usabilidad puede basarse en requerimientos de usuario específicos que no pueden describirse utilizando los elementos de la calidad descritos anteriormente. En este caso, el elemento usabilidad debe usarse para describir información específica de la calidad sobre la idoneidad de un conjunto de datos para una aplicación particular o sobre su conformidad con un conjunto de requisitos.

Anexo A
(normativo)
Medidas básicas de la calidad de datos de errores

A.1 Generalidades

Los valores numéricos obtenidos a partir de mediciones se obtienen con una exactitud determinada. La exactitud de la medición, y por ende el error de la misma, pueden cuantificarse al tratar la magnitud medida como una variable aleatoria. Para la definición de las medidas básicas relacionadas con el error se usan las diferentes formas de describirlo con métodos estadísticos.

Los métodos estadísticos usados para la definición de las medidas básicas de la calidad relacionadas con el error se basan en ciertos supuestos:

— los errores son homogéneos para todos los valores observados;

— no existe correlación entre los valores observados;

— los valores observados siguen una distribución normal.

A.2 Variable aleatoria unidimensional, *Z*

Si bien es imposible brindar la probabilidad de que un único valor sea verdadero para una magnitud medida, sí es posible proporcionar la probabilidad de que el valor verdadero esté contenido en cierto intervalo; dicha probabilidad *P* se denomina “nivel de significación” mientras que dicho intervalo se denomina “intervalo de confianza”.

P (límite inferior ≤ valor verdadero ≤ límite superior) = P

En las Tablas 35 y 37 se presentan las medidas básicas de la calidad de datos para el error de magnitudes unidimensionales. Ambos grupos de medidas pretenden valorar el error proporcionando los límites inferior y superior de un intervalo de confianza. La diferencia reside en cuándo es obtenida la desviación típica:

- Si se conoce *a priori*, véase el apartado A.2.1, en concreto la Tabla 35.
- Si se estima a partir de medidas redundantes, véase el apartado A.2.2, en concreto las Tablas 37 y 36.

A.2.1 Desviación típica conocida

Si la desviación típica es conocida, los límites vienen dados por los cuantiles de la distribución normal (gaussiana).

$$P\,(Z_t - u\,.\,\sigma \leq Z_t \leq Z_t + u\,.\,\sigma) = P$$

donde

Z_t es el valor verdadero (*true*) de la variable aleatoria Z,

σ es la desviación típica,

u son los cuantiles de la distribución normal (gausssiana).

Véase además la **Tabla 35.**

Tabla 33: Relación entre los cuantiles de la distribución normal y el nivel de significación

Nivel de significación (Probabilidad *P)*	Cuantil	Medida básica de la calidad de datos	Nombre
50 %	$u_{50\%} = 0{,}6745$	$u_{50\%}\ .\ \sigma_z$	LE50
68,3 %	$u_{68{,}3\%} = 1$	$u_{68{,}3\%}\ .\ \sigma_z$	LE68.3
90 %	$u_{90\%} = 1{,}645$	$u_{90\%}\ .\ \sigma_z$	LE90
95 %	$u_{95\%} = 1{,}960$	$u_{95\%}\ .\ \sigma_z$	LE95
99 %	$u_{99\%} = 2{,}576$	$u_{99\%}\ .\ \sigma_z$	LE99
99,8 %	$u_{99{,}8\%} = 2{,}576$	$u_{99{,}8\%}\ .\ \sigma_z$	LE99.8

A.2.2 Desviación típica desconocida

Si la desviación típica es desconocida pero la variable aleatoria unidimensional se mide de forma redundante mediante *N* observaciones independientes, es posible determinar la desviación típica a partir de dichas observaciones.

- Si el **valor verdadero de la variable aleatoria *Z*** es **conocido** y la redundancia se corresponde con el número de las observaciones (*N*), entonces la desviación típica puede calcularse de la siguiente manera:

$$s_Z = \sqrt{\frac{1}{r}\sum_{i=1}^{N}(z_{mi} - z_t)^2}$$

donde

z_{mi} es el valor medido de la variable aleatoria *Z*,

z_{ti} es el valor verdadero (*true*) de la variable aleatoria *Z*,

r es la redundancia.

- Si el **valor verdadero de la variable aleatoria *Z*** es **desconocido**, entonces éste puede calcularse como la media aritmética de las observaciones:

$$z_t = \sum_{i=1}^{N} z_{mi}$$

donde

z_{mi} es el valor medido de la variable aleatoria *Z*.

Y la desviación típica se calcula con la fórmula anterior con un valor de redundancia

$$r = N - 1$$

Si la desviación típica se estima con medidas redundantes, el intervalo de confianza puede derivarse a partir de la distribución *t* de *Student* con parámetro *r*:

$P\,(-t.s_z \le Z - z_t \le t.s_z) = P$ con $\dfrac{(Z - z_t)}{s_z} \sim t(r)$

Tabla 34: Relación entre los cuantiles de la distribución *t* de *Student*

Nivel de significación (Probabilidad *P)*	Cuantil para *r* = 10	Cuantil para *r* = 5	Cuantil para *r* = 4	Cuantil para *r* = 3	Cuantil para *r* = 2	Cuantil para *r* = 1
50 %	*t* = 1,221	*t* = 1,301	*t* = 1,344	*t* = 1,423	*t* = 1,604	*t* = 2,414
68,3 %	*t* = 1,524	*t* = 1,657	*t* = 1,731	*t* = 1,868	*t* = 2,203	*t* = 3,933
90 %	*t* = 2,228	*t* = 2,571	*t* = 2,776	*t* = 3,182	*t* = 4,303	*t* = 12,706
95 %	*t* = 2,634	*t* = 3,163	*t* = 3,495	*t* = 4,177	*t* = 6,205	*t* = 25,452
99 %	*t* = 3,581	*t* = 4,773	*t* = 5,598	*t* = 7,453	*t* = 14,089	*t* = 127,321
99,8 %	*t* = 4,587	*t* = 6,869	*t* = 8,610	*t* = 12,924	*t* = 31,599	*t* = 636,619

Tabla 35: Medidas básicas de la calidad de datos para diferentes probabilidades *P* para una magnitud unidimensional, donde la desviación típica se estima a partir de medidas redundantes

Nivel de significación (Probabilidad *P)*	Medida básica de la calidad de datos	Nombre
50,0 %	$t_{50\%}(r)\cdot s_Z$	LE50(r)
68,3 %	$t_{68,3\%}(r)\cdot s_Z$	LE68.3(r)
90,0 %	$t_{90\%}(r)\cdot s_Z$	LE90(r)
95,0 %	$t_{95\%}(r)\cdot s_Z$	LE95(r)
99,0 %	$t_{99\%}(r)\cdot s_Z$	LE99(r)
99,8 %	$t_{99,8\%}(r)\cdot s_Z$	LE99.8(r)
NOTA: Los valores de *t* para un número de redundancias *r* pueden obtenerse de la **Tabla *34***.		

A.3 Variable aleatoria bidimensional, *X e Y*

El caso de una variable aleatoria unidimensional *Z* puede expandirse a dos dimensiones cuando siempre se obtienen dos valores para la magnitud medida. El resultado viene dado por la tupla *X, Y*. Este caso tiene los mismos supuestos que el de una variable aleatoria unidimensional (véase el apartado A.2).

Las observaciones son x_{mi} e y_{mi}. La equivalencia con el intervalo de confianza en una dimensión es el área de confianza, que usualmente se describe como un círculo en torno a la mejor estimación del valor verdadero. La probabilidad de que el valor verdadero esté situado dentro de dicha área se calcula mediante la integración de superficies sobre la función de densidad bidimensional de la distribución normal. Un área circular se caracteriza por su radio. Este radio, *R*, se usa como medida de la exactitud de variables aleatorias bidimensionales (véase además la **Tabla *36***):

$$P(R,\sigma_X,\sigma_Y)=\frac{1}{2\pi\sigma_X\sigma_Y}\iint_{(x-x_t)^2+(y-y_t)^2=R^2} e^{-\frac{1}{2}\left(\frac{(x-x_t)^2}{\sigma_X^2}+\frac{(y-y_t)^2}{\sigma_Y^2}\right)}dxdy$$

Puede calcularse el radio para algunas probabilidades concretas dependiendo de las desviaciones típicas σ_x y σ_y.

Tabla 36: Relación entre la probabilidad P y el radio correspondiente del área circular

Nivel de significación (Probabilidad *P)*	Medida de la calidad de datos	Nombre
39,4 %	$\frac{1}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE39.4
50 %	$\frac{1,1774}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE50
90 %	$\frac{2,146}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE90
95 %	$\frac{2,4477}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE95
99,8 %	$\frac{3,5}{\sqrt{2}}\sqrt{\sigma_x^2+\sigma_y^2}$	CE99.8

A.4 Variable aleatoria tridimensional, *X, Y, Z*

El caso de una variable aleatoria unidimensional *Z* puede expandirse a tres dimensiones cuando siempre se obtienen tres valores para la magnitud medida. El resultado viene dado por la tupla *X, Y, Z*. Este caso tiene los mismos supuestos que el de una variable aleatoria unidimensional (véase el apartado A.2).

Las observaciones son x_{mi}, y_{mi} y z_{mi}. La equivalencia con el intervalo de confianza en una dimensión es el volumen de confianza, que usualmente se describe como una esfera en torno a la mejor estimación del valor verdadero. La probabilidad de que el valor verdadero esté situado dentro de este volumen se calcula mediante la integración de volúmenes sobre la función de densidad tridimensional de la distribución normal. Un volumen esférico se caracteriza por su radio. Este radio se usa como medida de la exactitud de variables aleatorias tridimensionales.

Nivel de significación (Probabilidad *P*)	**Medida básica de la calidad de datos**	**Nombre**
50 %	$0{,}51 \cdot \left(\sigma_x + \sigma_y + \sigma_z\right)$	error esférico probable (SEP, por sus siglas en ingles, *Spherical Error Probable*)
61 %	$\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$	error esférico radial medio (MRSE, por sus siglas en ingles, *mean radial spherical error*)
90 %	$0{,}833 \cdot \left(\sigma_x + \sigma_y + \sigma_z\right)$	exactitud estándar esférica al 90 %
P = 99 %	$1{,}122 \cdot \left(\sigma_x + \sigma_y + \sigma_z\right)$	exactitud estándar esférica al 90 %

Bibliografía

[1] ISO 19157:2003, *Geographic information—Data quality*

[2] ISO 19101-1:2014, *Geographic information—Reference model—Part 1: Fundamentals*