
NORMA CUBANA

NC

ISO 19114: 2010
(Publicada por la ISO en 2003)

**INFORMACIÓN GEOGRÁFICA — PROCEDIMIENTOS DE
EVALUACIÓN DE LA CALIDAD
(ISO 19114:2003, IDT)**

Geographic information — Quality evaluation procedures

ICS: 35.240.70

1. Edición Diciembre 2010
REPRODUCCIÓN PROHIBIDA

Oficina Nacional de Normalización (NC) Calle E No. 261 Vedado, Ciudad de La Habana, Cuba. Teléfono: 830-0835 Fax: (537) 836-8048 Correo electrónico: nc@ncnorma.cu Sitio Web: www.nc.cubaindustria.cu



Cuban National Bureau of Standards

Prefacio

La Oficina Nacional de Normalización (NC), es el Organismo Nacional de Normalización de la República de Cuba y representa al país ante las organizaciones internacionales y regionales de normalización.

La elaboración de las Normas Cubanas y otros documentos normativos relacionados se realiza generalmente a través de los Comités Técnicos de Normalización. Su aprobación es competencia de la Oficina Nacional de Normalización y se basa en las evidencias del consenso.

Esta Norma Cubana:

- Ha sido elaborada por el Comité Técnico de Normalización NC/CTN 113 de Geomática, en el que están representadas las siguientes entidades:
 - Comité de la Infraestructura de Datos Espaciales de la República de Cuba
 - Instituto de Geología y Paleontología, MINBAS
 - Oficina Nacional de Recursos Minerales, MINBAS
 - Oficina Nacional de Hidrografía y Geodesia, MINFAR
 - Grupo Empresarial GEOCUBA, MINFAR
 - Instituto de Planificación Física, MEP
 - Universidad Agraria de la Habana, MES
 - Instituto de Geografía Tropical, CITMA
 - Instituto Nacional de Recursos Hidráulicos
 - Ministerio de la Agricultura
 - Ministerio de la Industria Azucarera
 - Oficina Nacional de Normalización
- Es una adopción idéntica de la Norma Internacional ISO 19113:2003. Geographic information — Quality evaluation procedures
- En esta versión se incluyeron las modificaciones a la Norma Internacional ISO 19114:2003 contempladas en el erratum ISO 19114:2003/Cor. 1:2005.
- Para los propósitos de esta norma, se ha incluido un anexo nacional informativo como guía para los usuarios.
- En el Anexo Nacional se da una lista de las Normas Cubanas idénticas a las Normas Internacionales a las que se hace referencia en la ISO 19113.

© NC, 2010

Todos los derechos reservados. A menos que se especifique, ninguna parte de esta publicación podrá ser reproducida o utilizada en alguna forma o por medios electrónicos o mecánicos, incluyendo las fotocopias, fotografías y microfilmes, sin el permiso escrito previo de:

Oficina Nacional de Normalización (NC)

Calle E No. 261, Vedado, Ciudad de La Habana, Habana 4, Cuba.

Impreso en Cuba

Índice	Página
Prólogo	4
Introducción.....	5
1 Objeto y campo de aplicación.....	6
2 Conformidad.....	6
3 Normas para consulta	6
4 Términos y definiciones	6
5 Abreviaturas	7
6 Proceso para la evaluación de la calidad de datos	8
6.1 Generalidades	8
6.2 Componentes del proceso	8
7 Métodos de evaluación de la calidad de datos	10
7.1 Clasificación de los métodos de evaluación de la calidad de datos	10
7.2 Métodos de evaluación directa	11
7.3 Métodos indirectos de evaluación.....	13
7.4 Ejemplos de evaluación de la calidad de datos	13
8 Informe sobre la evaluación de la calidad de datos.....	13
8.1 Reporte como metadatos	13
8.2 Reporte en un informe sobre la evaluación de la calidad.....	13
8.3 Informe del resultado agregado de la calidad de datos.....	14
Anexo A (normativo) conjuntos de pruebas genéricas.....	15
A.1 Introducción	15
A.2 Procedimientos de evaluación de la calidad	15
A.3 Evaluación de la calidad de datos	15
A.4 Informe sobre la evaluación de la calidad	15
Anexo B (informativo) usos de los procedimientos de evaluación de la calidad	16
B.1 Introducción	16
B.2 Desarrollo de las especificaciones de producto o de los requisitos de usuario	16
B.3 Control de calidad durante la creación del conjunto de datos.....	16
B.4 Inspección para determinar la conformidad a una especificación de producto	16
B.5 Evaluación de la conformidad de un conjunto de datos respecto a los requisitos de usuario	16
B.6 Control de calidad durante la actualización de los datos	17
Anexo C (informativo) aplicación de los procedimientos de evaluación de la calidad a conjuntos de datos de carácter dinámico	18
C.1 Introducción	18
C.2 Determinación e informe de la calidad de de conjunto de datos dinámico	18
C.3 Establecimiento de procedimientos continuos para la evaluación de la calidad	18
C.4 Restablecimiento periódico de la calidad de referencia del conjunto de datos.....	19
Anexo D (informativo) ejemplos de medidas de la calidad de datos.....	20
D.1 Introducción	20
D.2 Relaciones entre las componentes de la calidad de datos.....	20
D.3 Ejemplos de medidas de la calidad relativas a la completitud.....	21

D.4 Ejemplos de medidas de la calidad relativas a la consistencia lógica	22
D.5 Ejemplos de medidas de la calidad relativas a la exactitud posicional	24
D.6 Ejemplos de medidas de la calidad relativas a la exactitud temporal	28
D.7 Ejemplos de medidas de la calidad relativas a la calidad temática	34
Anexo E (informativo) guía para la aplicación de métodos de muestreo a conjuntos de datos geográficos.....	38
E.1 Introducción.....	38
E.2 Lote e ítem.....	38
E.3 Tamaño de la muestra	38
E.4 Estrategias de muestreo	39
E.5 Muestreos de base probabilística	42
Anexo F (informativo) ejemplo de prueba de exactitud temática y completitud.....	44
F.1 Introducción	44
F.2 Proceso de evaluación de la calidad.....	44
F.3 Método para la evaluación de la calidad de datos.....	44
F.4 Inspección para la calidad.....	46
F.5 Determinación de los resultados de la calidad de datos y de la conformidad	46
F.6 Presentación de los resultados de calidad	47
Anexo G (informativo) ejemplo de medida e informe de la completitud y de la exactitud temática.....	50
G.1 Introducción	50
G.2 Descripción del conjunto de datos.....	50
G.3 Evaluación de la calidad de datos	56
G.4 Informe de los resultados de calidad	58
Anexo H (informativo) ejemplo de un resultado agregado de calidad de datos.....	61
H.1 Introducción	61
H.2 Descripción del conjunto de datos	61
H.3 Universo de discurso.....	62
H.4 Conjunto de datos.....	63
H.5 Agregación de los resultados de la evaluación e informe	63
Anexo I (normativo) _presentación de la información sobre la calidad mediante un informe de evaluación de la calidad.....	65
I.1 Introducción	65
I.2 Componentes del informe de evaluación de la calidad	65
ANEXO J (Informativo) Agregación de resultados de la calidad	69
J.1 Introducción	69
J.2 Conforme / no conforme al 100%	69
J.3 Conforme / no conforme ponderado	69
J.4 Subconjunto de resultados suficientes para los propósitos del producto	70
J.5 Valor máximo/mínimo	70
Bibliografía.....	71

Introducción

Para evaluar la calidad de un conjunto de datos se tienen que utilizar procedimientos claramente definidos y de una manera consistente. Esto permite, por un lado, a los productores de datos expresar de cuán bien sus productos se ajustan a los criterios establecidos en las especificaciones de producto, y por otro, que los usuarios establezcan en qué medida un conjunto de datos satisface sus requisitos. La calidad de un conjunto de datos se describe usando dos componentes, una cuantitativa y otra no cuantitativa. El objetivo de este documento es proveer una guía para los procedimientos de evaluación de la componente cuantitativa de la calidad de los datos geográficos de acuerdo con los principios de calidad descritos en la Norma NC ISO 19113. Este documento también ofrece orientación sobre cómo reportar la información relativa a la calidad.

Esta norma internacional considera que productores y usuarios pueden ver la calidad de los datos desde diferentes perspectivas. Los niveles de conformidad de la calidad de datos pueden establecerse usando las especificaciones de producto realizadas por el productor, o bien los requisitos de calidad del usuario de los datos. Si el usuario requiere más información sobre la calidad que aquella suministrada por el productor, aquél puede seguir el flujo del proceso de evaluación de la calidad de datos realizado por éste para obtener información adicional. En este caso, los requisitos de usuario se consideran como especificaciones de producto con el propósito de usar el flujo de procesos del productor.

Los procedimientos de evaluación de la calidad descritos en esta norma internacional, cuando se aplican en conformidad con la Norma NC ISO 19113, proporcionan una manera consistente y normalizada para determinar e informar sobre la calidad de un conjunto de datos.

INFORMACIÓN GEOGRÁFICA — PROCEDIMIENTOS DE EVALUACIÓN DE LA CALIDAD

1 Objeto y campo de aplicación

Esta norma internacional proporciona un marco de procedimientos para determinar y evaluar la calidad aplicable a conjuntos de datos geográficos digitales de manera consistente con los principios definidos en la Norma NC ISO 19113. También se establece un marco para evaluar e informar sobre los resultados de la calidad de datos como parte de los metadatos solamente, o también como un informe de evaluación de la calidad.

Esta norma internacional es aplicable a productores que suministren información sobre calidad acerca cuán bien sus productos se ajustan a los criterios establecidos en las especificaciones de producto, así como a los usuarios que pretendan determinar si un conjunto de datos tiene calidad suficiente para que resulte adecuado frente a su uso en aplicaciones particulares.

Aunque esta norma internacional es aplicable a todo tipo de datos geográficos digitales, sus principios pueden ser extendidos a muchas otras formas de datos geográficos como mapas, cartas y documentos de texto.

2 Conformidad

Esta norma internacional define tres clases de conformidad, una para los procedimientos de evaluación de la calidad, otra para la evaluación de la calidad de datos y la tercera para reportar la información sobre la calidad. En el anexo A se presentan los conjuntos/paquetes de pruebas para las tres clases de conformidad.

3 Normas para consulta

Las normas que a continuación se indican son indispensables para la aplicación de esta norma. Para las referencias con fecha, sólo se aplica la edición citada. Para las referencias sin fecha se aplica la última edición de la norma (incluyendo cualquier modificación de ésta). Los miembros de la ISO y la IEC llevan un registro de las normas internacionales vigentes.

NC ISO 19113:2010 — *Información geográfica. Principios de calidad.*

ISO 19115:2003 — *Información geográfica. Metadatos.*

4 Términos y definiciones

Para los fines de este documento se aplican los términos y definiciones dados en las Normas NC ISO 19113 e ISO 19115 (por comodidad, algunos de ellos se repiten) y los siguientes:

4.1 nivel de conformidad de la calidad

Valor umbral o conjunto de valores umbrales para los resultados de la calidad de los datos que son empleados para determinar la adecuación del conjunto de datos a los criterios establecidos en la especificación del producto o en los requisitos del usuario.

4.2 conjunto de datos

Colección identificable de datos. [ISO 19115]

NOTA - Un conjunto de datos puede ser una agrupación reducida de datos que, aunque esté limitada por alguna restricción, como la extensión espacial o el tipo de objeto geográfico, físicamente está localizada dentro de un conjunto mayor. Teóricamente, un conjunto de datos puede ser tan pequeño como un único objeto geográfico o atributo de objeto geográfico contenido dentro de un conjunto de datos mayor.

4.3 serie de conjuntos de datos

Colección de conjuntos de datos que comparten la misma especificación de producto. [ISO 19115]

4.4 método directo de evaluación

Método de evaluación de la calidad de un conjunto de datos basado en la inspección de sus ítems.

4.5 inspección completa

Inspección de todos y cada uno de los ítems de un conjunto de datos.

NOTA - La inspección completa también se conoce como inspección al 100%.

4.6 método indirecto de evaluación

Método de evaluación de la calidad de un conjunto de datos basado en conocimiento externo.

NOTA - Ejemplo de conocimiento externo es el linaje del conjunto de datos, como pueden ser el método de producción o la fuente de datos.

4.7 ítem

Aquel que puede ser individualmente descrito o considerado. [ISO 2859-1]

NOTA - Un ítem puede ser cualquier parte de un conjunto de datos, tal como un objeto geográfico, una relación entre objetos geográficos, un atributo de objeto geográfico, o cualquier combinación de estos.

4.8 población

Totalidad de ítems bajo consideración. [ISO 3534-2]

Ejemplo 1: Todos los puntos de un conjunto de datos.

Ejemplo 2: Los nombres de todas las carreteras en un área geográfica determinada.

4.9 datos de referencia

Datos aceptados como representativos del universo de discurso, para ser usados como referencia para los métodos directos de evaluación externa de la calidad.

5 Abreviaturas

RACD resultado agregado de la calidad de datos

NCA nivel de calidad aceptable [ISO 3534-2]

EMC error medio cuadrático

6 Proceso para la evaluación de la calidad de datos

6.1 Generalidades

Un proceso de evaluación de la calidad puede ser usado en diferentes fases del ciclo de vida de un producto, teniendo diferentes objetivos en cada una de ellas. Las fases del ciclo de vida aquí consideradas son: especificación, producción, entrega, uso y actualización. El anexo B describe algunas operaciones específicas, relacionadas con los conjuntos de datos, a las cuales les son aplicables procedimientos de evaluación de la calidad.

El proceso para la evaluación de la calidad de datos es una secuencia de pasos orientados a producir e informar sobre un resultado de la calidad; consiste en la aplicación de procedimientos de evaluación de la calidad a unas operaciones específicas, relacionadas con un conjunto de datos, que son realizadas por el productor y por el usuario.

Los procesos para la evaluación de la calidad de datos son aplicables tanto a conjuntos de datos estáticos como dinámicos. Los conjuntos dinámicos son aquellos cuyas actualizaciones son tan frecuentes que, a efectos prácticos, están cambiando continuamente. El anexo C describe la aplicación del proceso para evaluar la calidad de conjuntos de datos dinámicos.

6.2 Componentes del proceso

6.2.1 Flujo del proceso

El proceso de evaluación de la calidad es una secuencia de pasos realizados para producir un resultado de la evaluación. La figura 1 ilustra el diagrama de flujo del proceso para evaluar e informar sobre los resultados de la calidad de datos.

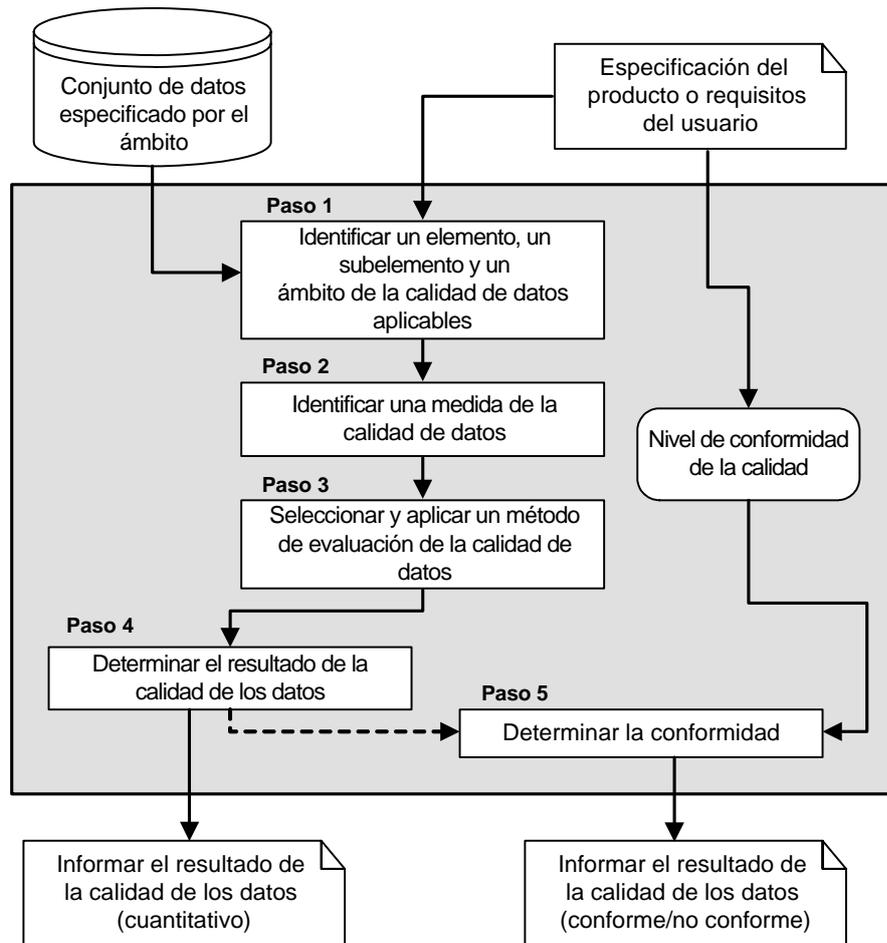


Figura 1 — Evaluación e informe sobre los resultados de la calidad de datos

6.2.2 Pasos del proceso

La tabla 1 especifica los pasos del proceso.

Tabla 1 — Pasos del proceso

Paso del proceso	Acción	Descripción
1	Identificar un elemento, un subelemento y el ámbito de calidad de datos aplicable	De acuerdo con los requisitos de la Norma NC ISO 19113 , se deben identificar: el elemento, subelemento y ámbito de la calidad a comprobar. Esto se repite para tantas pruebas diferentes como lo requieran las especificaciones de producto o los requisitos de usuario.
2	Identificar una medida de calidad de datos	Para cada prueba a realizar se debe identificar: una medida de la calidad, el tipo de valor y, si es de aplicación, la unidad de medida. El anexo D proporciona ejemplos de medidas para los elementos y subelementos dados en la Norma NC ISO 19113 . Por medio de estos ejemplos, el anexo D proporciona ayuda al usuario para la selección de una medida.

Tabla 1 (continuación)

Paso del proceso	Acción	Descripción
3	Seleccionar y aplicar un método de evaluación de la calidad de datos	Se selecciona un método para la evaluación de la calidad de datos para cada medida que se haya identificado. NOTA - Podría ser útil una descripción espacial de los resultados (por medio de una interpolación de los resultados, representación, etc.) correspondiente, no con un resultado, sino con un conjunto de diferentes, pero relacionados.
4	Determinar el resultado de la calidad de datos	Al aplicar el método se obtiene: un resultado cuantitativo de la calidad de datos, un valor o conjunto de valores, una unidad de medida y una fecha.
5	Determinar la conformidad	Siempre que se haya especificado un nivel de conformidad de la calidad en las especificaciones de producto o en los requisitos de usuario, el resultado de la calidad se compara con aquel para determinar la conformidad. El resultado de la conformidad (conforme / no conforme) es la comparación del resultado cuantitativo de la calidad con un nivel de conformidad para la calidad.

7 Métodos de evaluación de la calidad de datos

7.1 Clasificación de los métodos de evaluación de la calidad de datos

Los procedimientos de evaluación de la calidad se realizan mediante la aplicación de uno o más métodos de evaluación. Los métodos se dividen en dos clases principales: directos e indirectos. Los métodos directos determinan la calidad de datos mediante la comparación de los datos con información de referencia interna y/o externa. Los métodos indirectos infieren o estiman la calidad de datos usando información sobre los propios datos, como el linaje. Los métodos directos de evaluación pueden ser a su vez subclasificados atendiendo a la fuente de información necesaria para realizar la evaluación. La figura 2 representa esta estructura de clasificación.

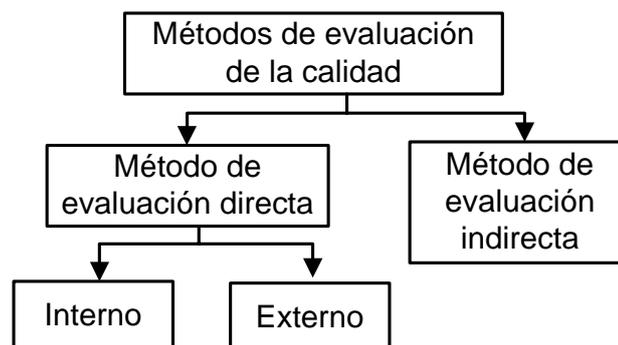


Figura 2 — Clasificación de los métodos de evaluación de la calidad de datos (informativa)

7.2 Métodos de evaluación directa

7.2.1 Tipos de métodos de evaluación directa

Los métodos de evaluación directa se subdividen en internos y externos. Todos los datos necesarios para llevar a cabo un método de evaluación directa de la calidad del tipo interno son datos internos al propio conjunto que se está evaluando.

Ejemplo 1: Todos los datos necesarios para realizar una prueba de consistencia lógica sobre la consistencia topológica del cierre de límites residen en el conjunto de datos topológicamente estructurado.

La evaluación directa externa de la calidad requiere datos de referencia externos al propio conjunto de datos que se está evaluando.

Ejemplo 2: Para realizar una prueba de completitud para los nombres de carreteras de un conjunto de datos se requiere información proveniente de otra fuente de información de nombres de carreteras.

Ejemplo 3: Una comprobación de la exactitud posicional requiere un conjunto de datos de referencia o una nueva medición.

7.2.2 Forma de efectuar la evaluación directa

Para ambos métodos de evaluación, externo e interno, existen dos opciones a tener en cuenta: automatizada o no automatizada, e inspección completa o muestreo.

Entre los elementos y subelementos de la calidad que son fácilmente comprobables de manera automática se incluyen los siguientes:

a) consistencia lógica: consistencia del formato;

Ejemplo: Comprobar que en los campos de datos haya entradas positivas.

b) consistencia lógica: consistencia topológica;

Ejemplo: Cierre de polígonos.

c) consistencia lógica: consistencia del dominio;

Ejemplo: Violaciones de límites, violaciones de los valores especificados para el dominio.

d) completitud: omisión;

Ejemplo: Comparación de control de los nombres de las calles empleando otro archivo.

e) completitud: comisión;

Ejemplo: Comparación de control de los nombres de las calles empleando otro archivo.

f) exactitud temporal: consistencia temporal;

Ejemplo: Comprobar que todos los registros presenten un rango de fechas apropiado.

7.2.3 Inspección completa

La inspección completa requiere la comprobación de cada ítem de la población especificada por el ámbito de la calidad de datos. La tabla 2 describe el procedimiento que se debe utilizar para la inspección completa.

Tabla 2 — Procedimiento para la inspección completa

Paso del proceso	Descripción
Definir los ítems	Un ítem es la mínima unidad a inspeccionar. Un ítem puede ser un objeto geográfico, un atributo de objeto geográfico o una relación de objetos geográficos.
Inspeccionar los ítems en el ámbito de la calidad de datos	Inspeccionar todos los ítems existentes en el ámbito de la calidad definido.

NOTA - La inspección completa es más apropiada para poblaciones pequeñas o para pruebas que pueden ser realizadas de manera automatizada.

7.2.4 Muestreo

El muestreo requiere la comprobación de suficientes ítems de la población para alcanzar un resultado representativo de calidad de datos. La tabla 3 describe el proceso de muestreo que debe usarse.

Tabla 3 — Procedimiento de muestreo

Paso del proceso	Descripción
Definir un método de muestreo	En el anexo E se presentan algunos ejemplos de métodos de muestreo. Estos métodos incluyen el muestreo aleatorio simple, el muestreo estratificado (por ejemplo guiados por un tipo de objeto geográfico, una relación de objetos geográficos o un área), el muestreo polietápico y el muestreo no aleatorio.
Definir los ítems	Un ítem es la mínima unidad a inspeccionar. Un ítem puede ser un objeto geográfico, un atributo de objeto geográfico o una relación entre objetos geográficos.
Dividir el ámbito de la calidad (población) en lotes	Un lote es un conjunto de ítems en el ámbito de calidad del cual se extrae e inspecciona una muestra. Cada lote debería estar formado, en la medida de lo posible, por ítems producidos bajo las mismas condiciones y a un mismo tiempo.
Dividir los lotes en unidades de muestreo	Las unidades de muestreo son las áreas del lote a las que se dirige la inspección.
Definir la razón de muestreo o el tamaño de la muestra	Una razón de muestreo da información acerca del promedio de ítems que son extraídos de cada lote para la inspección.
Seleccionar las unidades de muestreo	Se selecciona el número requerido de unidades de muestreo de forma que se satisfaga la razón de muestreo o el tamaño de la muestra.
Inspeccionar los ítems en las unidades de muestreo	Se inspecciona cada ítem de las unidades de muestreo.

Se debe informar sobre el procedimiento de muestreo de acuerdo con el capítulo 8 de esta norma internacional.

La serie de Normas ISO 2859 y la Norma ISO 3951-1 pueden ser aplicadas a los muestreos para evaluar la conformidad frente a una especificación de producto. Originalmente estas normas fueron desarrolladas para un uso no espacial. El anexo E de esta norma internacional proporciona ejemplos que describen como aplicar la serie de Normas ISO 2859 y la Norma ISO 3951-1, y presenta orientaciones sobre como definir muestras e idear métodos de muestreo que tengan en cuenta la naturaleza geográfica de los datos.

Cuando se utiliza el muestreo debería analizarse la fiabilidad del resultado, especialmente cuando las muestras son de pequeño tamaño y los métodos aplicados son distintos del muestreo aleatorio simple.

7.3 Métodos indirectos de evaluación

El método indirecto de evaluación es un método de evaluación de la calidad de un conjunto de datos basado en conocimiento externo. Dicho conocimiento puede incluir, pero sin limitarse a ello, los elementos generales de la calidad de datos y otros informes sobre la calidad del conjunto de datos o sobre los datos usados para producirlo.

NOTA 1 - Este método sólo se recomienda si no se pueden usar métodos directos de evaluación.

NOTA 2 - La información sobre el uso registra los usos de un conjunto de datos. Ésta es de ayuda para la búsqueda de conjuntos de datos que han sido producidos o utilizados con fines determinados.

NOTA 3 - La información sobre el linaje registra información acerca de la producción y la historia del conjunto de datos. Por ejemplo, incluye información sobre las fuentes utilizadas para producir el conjunto de datos o sobre los procesos de producción aplicados. Esto es útil para determinar la adecuación de un conjunto de datos a un uso dado. Un ejemplo son los metadatos de linaje relativos a un fichero correspondiente a un modelo digital del terreno que ha sido creado por medio de estéreo-correlación a partir de imágenes capturadas bajo ciertas condiciones. La experiencia le dice al evaluador que el EMC posicional horizontal es de 10 m para este tipo de imágenes. Otro ejemplo puede ser el siguiente: los metadatos que recogen el linaje de un mapa topográfico a escala 1:25 000 digitalizado indican la conformidad a los requisitos para un mapa base de un urbanista.

NOTA 4 - La información sobre el propósito describe el fin para el cual se produjo el conjunto de datos. Un propósito puede ser la base de un requisito específico, pero el conjunto de datos puede ser de propósito general, para varios usos. Esto es útil para identificar el posible valor de un conjunto de datos.

7.4 Ejemplos de evaluación de la calidad de datos

En los anexos F, G y H se describen ejemplos de métodos típicos utilizados y de cómo se pueden aplicar.

8 Informe sobre la evaluación de la calidad de datos

8.1 Reporte como metadatos

Los resultados cuantitativos de la calidad deben aparecer como metadatos en conformidad con la Norma ISO 19115, la cual contiene el modelo y el diccionario de datos.

8.2 Reporte en un informe sobre la evaluación de la calidad

Existen dos condiciones bajo las cuales se debe generar un informe sobre la evaluación de la calidad:

- a) cuando los resultados de la calidad de datos, presentados como metadatos, sólo son reportados como Conforme / no conforme;
- b) cuando se generan resultados agregados de calidad de datos.

Es bajo la última circunstancia cuando se requiere un informe para explicar cómo se ha realizado la agregación y cómo interpretar el significado del resultado agregado. No obstante, se puede generar un informe sobre la evaluación de la calidad en cualquier otro caso (por

ejemplo, para proporcionar mayor detalle que el incluido en los metadatos) pero en este caso no podrá ser utilizado en lugar del informe de metadatos.

El informe sobre la evaluación de la calidad debe ser generado de conformidad según lo estipulado en el anexo I, el cual contiene el modelo pertinente y el diccionario de datos.

8.3 Informe del resultado agregado de la calidad de datos

Cuando para informar sobre la calidad de un conjunto de datos se agregan varios resultados de la calidad para generar un resultado único, el resultado agregado debe aparecer como metadato y debe estar incluido en el informe sobre la calidad de datos. El resultado debe aparecer tipificado como “agregado”. En el anexo J se describe cómo se generan resultados agregados y en el anexo H se incluye un ejemplo.

ANEXO A (Normativo)

CONJUNTOS DE PRUEBAS ABSTRACTAS

A.1 Introducción

Este anexo define tres clases de conformidad:

- Procedimiento de evaluación de la calidad (capítulo A.2);
- Evaluación de la calidad de datos (capítulo A.3);
- Informe sobre la calidad de datos (capítulo A.4).

Cualquier procedimiento de evaluación de la calidad que pretenda alcanzar la conformidad con esta norma internacional debe cumplir todos los requisitos que se describen en el capítulo A.2. Cualquier evaluación de la calidad de datos que pretenda alcanzar la conformidad con esta norma internacional debe cumplir todos los requisitos que se describen en el capítulo A.3. Cualquier informe sobre la calidad de datos que pretenda alcanzar la conformidad con esta norma internacional debe cumplir todos los requisitos que se describen en el capítulo A.4.

NOTA - Todos los casos son del tipo “básico”.

A.2 Procedimientos de evaluación de la calidad

El conjunto de pruebas para la clase 1 debe ser el siguiente:

- a) Propósito de la prueba: Asegurar que el procedimiento de evaluación de la calidad fue generado de acuerdo con esta norma internacional.
- b) Método de comprobación: Se cumplen todos los requisitos descritos en los capítulos A.3 y A.4.
- c) Referencias: Capítulos A.3 y A.4.

A.3 Evaluación de la calidad de datos

El conjunto de pruebas para la clase 2 debe ser el siguiente:

- a) Propósito de la prueba: Asegurar que el procedimiento de evaluación de la calidad fue generado de acuerdo con el proceso de evaluación de la calidad del capítulo 6.
- b) Método de comprobación: Se compara la adecuación del procedimiento de evaluación de la calidad con la evaluación de la calidad.
- c) Referencias: Norma NC ISO 19114:2010, capítulo 6.

A.4 Informe sobre la evaluación de la calidad

El conjunto de pruebas para la clase 3 debe ser el siguiente:

- a) Propósito de la prueba: Asegurar que se ha informado sobre la calidad de datos de acuerdo con el capítulo 8.
- b) Método de comprobación: Se revisa el informe sobre la evaluación de la calidad para asegurar que los resultados de la calidad han sido reportados apropiadamente, de acuerdo con el capítulo 8 y los anexos que le sean aplicables.
- c) Referencias: Norma NC ISO 19114:2010, capítulo 8.

ANEXO B (Informativo)

USOS DE LOS PROCEDIMIENTOS DE EVALUACIÓN DE LA CALIDAD

B.1 Introducción

Los procedimientos de evaluación de la calidad pueden usarse en diferentes fases del ciclo de vida de un producto. En este anexo se muestran ejemplos de etapas del ciclo de vida de un producto durante las cuales se pueden aplicar procedimientos de evaluación de la calidad.

B.2 Desarrollo de las especificaciones de producto o de los requisitos de usuario

Cuando se desarrollan las especificaciones de un producto, o los requisitos de usuario, los procedimientos de evaluación de la calidad pueden utilizarse para ayudar al establecimiento de los niveles de conformidad de la calidad de datos que deberían ser alcanzados por el producto final. Una especificación de producto, o un requisito de usuario, debería incluir niveles de conformidad para la calidad del conjunto de datos, y también los procedimientos de evaluación de la calidad a aplicar durante la producción y actualización.

B.3 Control de calidad durante la creación del conjunto de datos

Durante la etapa de producción, como parte del proceso de control de calidad, el productor puede aplicar procedimientos de evaluación de la calidad establecidos (explícitamente o no) en las especificaciones del producto. Cuando son utilizados para el control de calidad en la producción, se debería informar en los metadatos con una descripción de los procedimientos de evaluación de la calidad aplicados, en forma de linaje; incluyendo, pero no necesariamente limitado a los mismos, los procedimientos de evaluación de la calidad aplicados, los niveles de conformidad de la calidad de datos establecidos y los resultados.

B.4 Inspección para determinar la conformidad a una especificación de producto

Una vez finalizada la producción se desarrolla un proceso de evaluación de la calidad, se generan resultados de la calidad de datos y se informa sobre los mismos. Estos resultados pueden utilizarse para determinar si el conjunto de datos es conforme a sus especificaciones de producto. Si el conjunto de datos supera la inspección (compuesta por un conjunto de procedimientos de evaluación de la calidad), se puede considerar que el conjunto de datos está listo para su uso. Se debería informar sobre los resultados de la inspección de acuerdo con el capítulo 8 de esta norma.

A consecuencia de una inspección se producirá la aceptación o rechazo del conjunto de datos. Si se rechaza, tras la corrección pertinente, se requerirá una nueva inspección antes de que el producto pueda ser considerado conforme a la especificación del producto.

B.5 Evaluación de la conformidad de un conjunto de datos respecto a los requisitos de usuario

Se usan procedimientos de evaluación de la calidad para establecer los niveles de calidad de un conjunto de datos respecto a un requisito del usuario. En este análisis de la conformidad se pueden utilizar tanto métodos directos como indirectos. Los resultados pueden incluirse como metadatos de uso para el conjunto de datos.

B.6 Control de calidad durante la actualización de los datos

Los procedimientos de evaluación de la calidad también se aplican a las operaciones de actualización de los datos, tanto a los ítems usados para la actualización, como para determinar la calidad del conjunto de datos después de su actualización. El anexo C proporciona una guía para la aplicación de la Norma NC ISO 19113 y de esta norma internacional a conjuntos de datos de carácter dinámico.

ANEXO C

(Informativo)

APLICACIÓN DE LOS PROCEDIMIENTOS DE EVALUACIÓN DE LA CALIDAD A CONJUNTOS DE DATOS DE CARÁCTER DINÁMICO

C.1 Introducción

Este anexo describe cómo se pueden aplicar procedimientos de evaluación de la calidad a conjuntos de datos de carácter dinámico. Se consideran aquí conjuntos de datos dinámicos como aquellos cuyas actualizaciones son tan frecuentes que a los efectos prácticos están cambiando continuamente. Por ejemplo, un conjunto de datos catastrales en línea puede recibir actualizaciones cada pocos minutos. Básicamente existen dos formas de determinar e informar sobre la calidad de un conjunto de datos dinámicos: periódica y continua.

C.2 Determinación e informe de la calidad de conjuntos de datos dinámicos

C.2.1 Procedimiento periódico

El procedimiento periódico se basa en el establecimiento de una frecuencia adecuada para la emisión de informes (por ejemplo, semanal o trimestral). En la fecha del informe se realiza una copia del conjunto de datos, y se evalúa su calidad como si fuera de carácter estático. Este procedimiento informa de la calidad del conjunto de datos en la fecha y hora de la copia.

C.2.2 Procedimiento continuo

El procedimiento continuo se basa en la comprobación de las actualizaciones y en la evaluación del impacto de las mismas. Esto es equivalente a incluir los procedimientos de evaluación de la calidad presentados en esta Norma Internacional en un procedimiento orientado al proceso (por ejemplo, el dado por la Norma ISO 9001). Dado que este procedimiento sólo puede informar del estado actual de la calidad de los ítems actualizados, se hace necesario combinar los procedimientos periódico y continuo, como se describe en el capítulo C.3, en orden a establecer la calidad de la base de datos actualizada.

C.3 Establecimiento de procedimientos continuos para la evaluación de la calidad

C.3.1 Identificación de las partes

De acuerdo con los pasos marcados en el apartado 6.2, se identifican los elementos aplicables de la calidad de datos y sus subelementos adecuados, los ámbitos de la calidad, las medidas de calidad y los niveles de conformidad a usar en la evaluación e informe de los resultados.

C.3.2 Selección del método a aplicar

Se selecciona el método a aplicar para la evaluación de la calidad de datos. La evaluación se hará sobre el objeto geográfico actualizado y la relación de este objeto geográfico con los demás dentro del ámbito marcado para la calidad. En los procedimientos continuos de evaluación de la calidad sólo pueden aplicarse métodos indirectos o directos internos.

Ejemplos:

- 1) ¿Procede la actualización de una fuente de confianza?
- 2) ¿Preserva la actualización la consistencia topológica?

3) ¿Mantiene el objeto geográfico actualizado la consistencia lógica?

C.3.3 Establecimiento de una referencia para la calidad del conjunto de datos

Utilice el procedimiento periódico para establecer valores de referencia de la calidad de un conjunto de datos para los objetos geográficos y atributos de objetos geográficos dentro del ámbito a verificar durante la comprobación de carácter continuo.

C.3.4 Integración de las pruebas de carácter continuo en el proceso de actualización

Integrar las pruebas de carácter continuo en el flujo del proceso de actualización de forma tal que cada actualización propuesta sea comprobada y aceptada previamente a su introducción en el conjunto de datos.

C.3.5 Actualización dinámica de los resultados de la calidad de datos

Por medio de la integración de las pruebas de carácter continuo en el flujo del proceso de actualización, cada actualización aceptada ocasiona un reajuste de los resultados de la calidad. Esto permite la generación inmediata de informes sobre la calidad del conjunto de datos.

C.4 Restablecimiento periódico de la calidad de referencia del conjunto de datos

Todos los aspectos de la calidad de un conjunto de datos no tienen por qué comprobarse mediante operaciones basadas en procesos continuos. Por ejemplo, la omisión de objetos geográficos puede no ser determinada cuando sólo se comprueban los ítems actualizados. El conjunto de datos debería someterse a evaluaciones periódicas de la calidad.

ANEXO D

(Informativo)

EJEMPLOS DE MEDIDAS DE LA CALIDAD DE DATOS

D.1 Introducción

Este anexo proporciona ejemplos elementales de medidas de la calidad para cada uno de los elementos de la calidad y subelementos asociados definidos en la Norma NC ISO 19113, para demostrar cómo se relacionan las componentes de la calidad de datos durante un proceso de evaluación. En otros anexos de esta norma internacional se pueden encontrar ejemplos más detallados.

Para cada combinación de elemento y subelemento de la calidad de datos se da un ejemplo del ámbito y de los parámetros del conjunto de datos. De esta forma, se presentan tres medidas de la calidad, cada una diseñada para demostrar una manera distinta de evaluar la calidad. Así, los ejemplos serán tan completos como sea posible; proporcionando la fecha y el nivel de conformidad de la calidad alcanzado. Finalmente, a modo de ejemplo del significado de los resultados de la calidad, se incluye una interpretación de éstos.

Si bien los ejemplos que se presentan en este anexo son elementales, puede ser deseable el referirse a ellos en perfiles y otros documentos. Por tanto, este anexo dispone de un código de identificación de la medida de la calidad de datos que relaciona el ejemplo con el elemento y subelemento.

D.2 Relaciones entre las componentes de la calidad de datos

La tabla D.1 presenta las relaciones entre las componentes de la calidad de datos.

Para ahorrar espacio, a cada componente de la calidad de datos se le ha asignado una abreviatura, que será utilizada a lo largo de este anexo.

Tabla D.1 — Relaciones entre las componentes de la calidad de datos

Componentes de calidad de los datos	Nombre corto *	Dominio de la componente	Ejemplo
Ámbito de calidad de datos	DQ_Scope	Texto libre	Todos los elementos clasificados como casas
Elemento de calidad de datos	DQ_Element	Dominio enumerado 1. Completitud 2. Consistencia lógica 3. Exactitud posicional 4. Exactitud temporal 5. Exactitud temática	1. Completitud Elemento de calidad de datos que describe la presencia o ausencia de objetos geográficos, de sus atributos y de sus relaciones.
Subelemento de la calidad de datos	DQ_Subelement	Dominio enumerado (depende del elemento de la calidad de datos)	1. Comisión Datos en exceso en el conjunto de datos
Medida de calidad de datos	DQ_Measure		
Descripción de la medida de calidad de datos	DQ_MeasureDesc	Texto libre	Existencia de elementos en exceso
Código de identificación de la medida de la calidad de datos	DQ_MeasureID	Dominio enumerado	10101
Método de evaluación de la calidad de datos	DQ_EvalMethod		
Tipo de método de evaluación de la calidad de datos	DQ_EvalMethodType	Dominio enumerado 1. Interno (directo) 2. Externo (directo) 3. Indirecto	2. Externo
Descripción de método de evaluación de la calidad de datos	DQ_EvalMethodDesc	Texto libre o cita (dependiendo del tipo de método de evaluación de la calidad de datos)	Compare el total de ítems en el conjunto de datos frente al total de ítems en el universo de discurso
Resultado de la calidad de datos	DQ_QualityResult		
Tipo del valor de la calidad de datos	DQ_ValueType	Dominio enumerado 1. Variable booleana 2. Número 3. Razón 4. Porcentaje 5. Muestra 6. Tabla 7. Imagen binaria 8. Matriz 9. Cita (ISO 19115) 10. Texto libre 11. Otro	1. Variable booleana
Valor de la calidad de datos	DQ_Value	Registro (Norma ISO 11404) (Depende del tipo de valor de la calidad de datos)	Verdadero
Unidad de valor de la calidad de datos	DQ_ValueUnit	(Depende del valor de la calidad de datos)	No aplicable
Fecha de calidad de datos	DQ_Date	Norma ISO 8601:1988	2000-03-05
Nivel de conformidad de la calidad de datos	DQ_ConformanceLevel	Valor o conjunto de valores	Diferencia nula entre el número de casos del conjunto de datos y del universo de discurso

* NOTA – El nombre corto es para uso dentro de este anexo.

D.3 Ejemplos de medidas de la calidad relativas a la completitud

La completitud es la presencia o ausencia de objetos geográficos, sus atributos y sus relaciones. Presenta los siguientes subelementos:

- comisión: datos excedentes en el conjunto de datos;
- omisión: datos ausentes en el conjunto de datos.

La tabla D.2 muestra algunos ejemplos relativos a los subelementos.

Tabla D.2 — Ejemplos de medidas de la calidad relativas a la completitud

Componente de calidad de datos	Ejemplo 1	Ejemplo 2	Ejemplo 3
DQ_Scope	Todos los ítems clasificados como casas en el conjunto de datos.	Todos los ítems clasificados como casas delimitados por las longitudes -83.1 -83.3 y las latitudes +38.3 +38.4	Todos los ítems clasificados como casas en la ciudad de Augusta, Georgia, EE.UU.
DQ_Element	1. Completitud	1. Completitud	1. Completitud
DQ_Subelement	1. Comisión	1. Comisión	1. Comisión
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de comisiones	Porcentaje de comisiones
DQ_MeasureID	10101	10102	10103
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Se compara el total de ítems en el conjunto de datos frente al total de ítems en el universo de discurso.	Se compara el total de ítems en el conjunto de datos frente al total de ítems en el universo de discurso.	Se divide el total de ítems en el conjunto de datos entre el total de ítems en el universo de discurso, y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	10	10
DQ_ValueUnit	No aplicable	Casas	Tanto por ciento
DQ_Date	2000-03-05	2000-03-06	2000-03-04
DQ_ConformanceLevel	Ninguna comisión en el conjunto de datos.	Menos de 9 comisiones en el conjunto de datos.	Menos del 9% de comisiones en el conjunto de datos.
Ejemplo de parámetros del conjunto de datos	110 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.	110 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.	110 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Existe un exceso de ítems. En el conjunto de datos aparecen clasificados como casas más ítems de los que lo están en el universo de discurso.	El conjunto de datos no es conforme. El número de ítems en exceso en el conjunto de datos excede el nivel de conformidad de la calidad de datos.	El conjunto de datos no es conforme. El porcentaje ítems en exceso en el conjunto de datos excede el nivel de conformidad de la calidad de datos.

Tabla D.2 (continuación)

Componente de calidad de datos	Ejemplo 4	Ejemplo 5	Ejemplo 6
DQ_Scope	Todos los ítems clasificados como casas en el conjunto de datos.	Todos los ítems clasificados como casas delimitados por las longitudes -83.1 -83.3 y las latitudes +38.3 +38.4	Todos los ítems clasificados como casas existentes en la ciudad de Estocolmo, Suecia.
DQ_Element	1. Completitud	1. Completitud	1. Completitud
DQ_Subelement	2. Omisión	2. Omisión	2. Omisión
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de omisiones	Porcentaje de omisiones
DQ_MeasureID	10201	10202	10203
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Se compara el total de ítems en el conjunto de datos frente al total de ítems en el universo de discurso.	Se compara el total de ítems en el conjunto de datos frente al total de ítems en el universo de discurso.	Se divide el total de ítems en el conjunto de datos entre el total de ítems en el universo de discurso, y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	10	10
DQ_ValueUnit	No aplicable	Casas	Tanto por ciento
DQ_Date	2000-03-06	2000-03-03	2000-03-04
DQ_ConformanceLevel	Ninguna omisión en el conjunto de datos.	Menos de 9 omisiones en el conjunto de datos.	Menos del 9% de omisiones en el conjunto de datos.
Ejemplo de parámetros del conjunto de datos	90 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.	90 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.	90 ítems del conjunto de datos están dentro del ámbito; 100 ítems del universo de discurso están dentro del ámbito.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Existen omisiones. En el conjunto de datos aparecen clasificados como casas menos ítems de los que lo están en el universo de discurso.	El conjunto de datos no es conforme. El número de ítems omitidos en el conjunto de datos excede el nivel de conformidad de la calidad de datos.	El conjunto de datos no es conforme. El porcentaje ítems omitidos en el conjunto de datos excede el nivel de conformidad de la calidad de datos.

D.4 Ejemplos de medidas de la calidad relativas a la consistencia lógica

La consistencia lógica es el grado de adherencia de la estructura de los datos, de los atributos y de las relaciones, a un conjunto de reglas lógicas (la estructura de los datos puede ser conceptual, lógica o física). La consistencia lógica posee los siguientes subelementos:

- consistencia conceptual: adherencia a las reglas del esquema conceptual;
- consistencia de dominio: adherencia de los valores al dominio marcado para los mismos;
- consistencia del formato: grado en que los datos son almacenados de acuerdo con la estructura física del conjunto de datos;
- consistencia topológica: corrección de las características topológicas codificadas explícitamente para un conjunto de datos.

La tabla D.3 presenta algunos ejemplos relativos a los subelementos.

Tabla D.3 —Ejemplos de medidas de Consistencia lógica de calidad de datos

Componente de calidad de datos	Ejemplo 1	Ejemplo 2	Ejemplo 3
DQ_Scope	Todos los ítems clasificados como casas en el conjunto de datos.	Todos los ítems clasificados como casas delimitados por las longitudes -83,1 -83,3 y las latitudes +38,3 +38,4	Todos los ítems clasificados como casas existentes en la ciudad de Helsinki, Finlandia.
DQ_Element	2. Consistencia lógica	2. Consistencia lógica	2. Consistencia lógica
DQ_Subelement	1. Consistencia conceptual	1. Consistencia conceptual	1. Consistencia conceptual
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de violaciones	Porcentaje de violaciones
DQ_MeasureID	20101	20102	20103
DQ_EvalMethod			
DQ_EvalMethodType	1. Interno	1. Interno	1. Interno
DQ_EvalMethodDesc	Se contabiliza el número de objetos geográficos y relaciones de objetos geográficos que violan el esquema conceptual del conjunto de datos.	Se contabiliza el número de objetos geográficos y relaciones de objetos geográficos que violan el esquema conceptual del conjunto de datos.	Se divide el número de objetos geográficos y relaciones de objetos geográficos que violan el esquema conceptual entre el total existente en el conjunto de datos, y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	1	1,0%
DQ_ValueUnit	No aplicable	Casos	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ninguna violación en el conjunto de datos.	Ninguna violación en el conjunto de datos.	Cero por ciento de violaciones en el conjunto de datos.
Ejemplo de parámetros del conjunto de datos	80 objetos geográficos y 20 relaciones de objetos geográficos están dentro de ámbito. Una relación de objetos geográficos no está definida en el esquema conceptual.	80 objetos geográficos y 20 relaciones de objetos geográficos están dentro de ámbito. Una relación de objetos geográficos no está definida en el esquema conceptual.	80 objetos geográficos y 20 relaciones de objetos geográficos están dentro de ámbito. Una relación de objetos geográficos no está definida en el esquema conceptual.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Existe violación del esquema conceptual.	El conjunto de datos no es conforme. El número de violaciones excede el nivel de conformidad de la calidad de datos.	El conjunto de datos no es conforme. El porcentaje de violaciones excede el nivel de conformidad de la calidad de datos.

Tabla D.3 (continuación)

Componente de calidad de datos	Ejemplo 4	Ejemplo 5	Ejemplo 6
DQ_Scope	Todos los ítems clasificados como viviendas en el conjunto de datos.	Todos los ítems clasificados como viviendas delimitados por las longitudes -91.3 -91.4 y las latitudes +40.0 +40.2	Todos los ítems clasificados como viviendas existentes en la ciudad de Londres, Inglaterra, Reino Unido.
DQ_Element	1. Consistencia lógica	1. Consistencia lógica	1. Consistencia lógica
DQ_Subelement	2. Consistencia del dominio	2. Consistencia del dominio	2. Consistencia del dominio
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de inconsistencias del dominio	Porcentaje de inconsistencias del dominio
DQ_MeasureID	20201	20202	20203
DQ_EvalMethod			
DQ_EvalMethodType	1. Interno	1. Interno	1. Interno
DQ_EvalMethodDesc	Se comparan los atributos de los ítems dentro de ámbito con el dominio aceptado para los atributos y se determina si alguno está fuera de éste.	Se comparan los atributos de los ítems dentro de ámbito con el dominio aceptado para los atributos y se determina si alguno está fuera de éste.	Se divide el total de ítems con violaciones de atributos entre el total de ítems en el ámbito y se multiplica el resultado por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	8	8,0%
DQ_ValueUnit	No aplicable	Violaciones de atributos	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún ítem con violación de atributos	Diez o menos ítems con violaciones de atributos	Menos del 5,0% de los ítems con violaciones de atributos
Ejemplo de parámetros del conjunto de datos	100 ítems dentro del ámbito en el conjunto de datos. Ocho de ellos poseen atributos que violan su dominio.	100 ítems dentro del ámbito en el conjunto de datos. Ocho de ellos poseen atributos que violan su dominio.	100 ítems dentro del ámbito en el conjunto de datos. Ocho de ellos poseen atributos que violan su dominio.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Los atributos de al menos un elemento violan el dominio de los atributos.	El conjunto de datos es conforme. Existen menos de 10 ítems que tienen atributos que violan su dominio.	El conjunto de datos no es conforme. Más del 5% de los ítems tienen atributos que violan su dominio.

Tabla D.3 (continuación)

Componente de calidad de datos	Ejemplo 7	Ejemplo 8	Ejemplo 9
DQ_Scope	Todos los registros en el conjunto de datos para ítems clasificados como viviendas.	Todos los registros en el conjunto de datos para ítems clasificados como viviendas delimitados por las longitudes: +139 +140 y las latitudes: +36,0 +37,0	Todos los registros en el conjunto de datos para ítems clasificados como viviendas existentes en la ciudad de Tokio, Japón.
DQ_Element	2. Consistencia lógica	2. Consistencia lógica	2. Consistencia lógica
DQ_Subelement	3. Consistencia del formato	3. Consistencia del formato	3. Consistencia del formato
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de inconsistencias en el formato	Porcentaje de inconsistencias en el formato
DQ_MeasureID	20301	20302	20303
DQ_EvalMethod			
DQ_EvalMethodType	1. Interno	1. Interno	1. Interno
DQ_EvalMethodDesc	Se compara la estructura de los registros de todos los ítems dentro de ámbito con las definiciones de los campos y estructura especificados, y se contabilizan aquellos que son inconsistentes. Se asegura específicamente que el campo para el código de tipo de vivienda sea un campo alfabético de 5 caracteres de longitud.	Se compara la estructura de los registros de todos los ítems dentro de ámbito con las definiciones de los campos y estructura especificados, y se contabilizan aquellos que son inconsistentes. Se asegura específicamente que el campo para el código de tipo de vivienda sea un campo alfabético de 5 caracteres de longitud.	Se compara la estructura de los registros de todos los ítems dentro de ámbito con las definiciones de los campos y estructura especificados, y se contabilizan aquellos que son inconsistentes. Se asegura específicamente que el campo para el código de tipo de vivienda sea un campo alfabético de 5 caracteres de longitud. Se divide el total de casos encontrados por el número de registros comprobados, y se multiplica el resultado por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	8	8,0%
DQ_ValueUnit	No aplicable	Violaciones del formato	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún elemento puede tener violaciones del formato	Ningún elemento puede tener violaciones del formato	El 0% de los ítems pueden tener violaciones del formato
Ejemplo de parámetros del conjunto de datos	100 ítems dentro del ámbito, 8 violan el formato especificado.	100 ítems dentro del ámbito, 8 violan el formato especificado.	100 ítems dentro del ámbito, 8 violan el formato especificado.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Se han encontrado violaciones del formato.	El conjunto de datos no es conforme. Se han encontrado violaciones del formato.	El conjunto de datos no es conforme. Se han encontrado violaciones del formato.

Tabla D.3 (continuación)

Componente de calidad de datos	Ejemplo 10	Ejemplo 11	Ejemplo 12
DQ_Scope	Todos los límites provinciales en el conjunto de datos.	Todos los límites estatales en los Estados Unidos.	Todos los límites estatales en los Estados Unidos.
DQ_Element	2. Consistencia lógica	2. Consistencia lógica	2. Consistencia lógica
DQ_Subelement	4. Consistencia topológica	4- Consistencia topológica	4. Consistencia topológica
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de ítems con inconsistencias topológicas	Porcentaje de ítems con inconsistencia topológicas
DQ_MeasureID	20401	20402	20403
DQ_EvalMethod			
DQ_EvalMethodType	1. Interno	1. Interno	1. Interno
DQ_EvalMethodDesc	Para cada provincia se comprueban sus límites para asegurar el cierre. Se contabiliza el número de provincias cuyos límites no cierran.	Para cada estado se comprueban sus límites para asegurar el cierre. Se contabiliza el número de estados cuyos límites no cierran.	Para cada estado se comprueban sus límites para asegurar el cierre. Se contabiliza el número de estados cuyos límites no cierran. Se divide el total de casos encontrados por el número de registros comprobados y se multiplica el resultado por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	2	2,0%
DQ_ValueUnit	No aplicable	Inconsistencias topológicas.	Tanto por ciento de inconsistencias topológicas.
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún ítem puede tener violaciones topológicas.	Ningún ítem puede tener violaciones topológicas.	El 0% de los ítems pueden tener violaciones topológicas.
Ejemplo de parámetros del conjunto de datos	100 ítems del conjunto de datos están dentro de ámbito. Dos de ellos presentan inconsistencias topológicas.	100 ítems del conjunto de datos están dentro de ámbito. Dos de ellos presentan inconsistencias topológicas.	100 ítems del conjunto de datos están dentro de ámbito. Dos de ellos presentan inconsistencias topológicas.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Se han encontrado inconsistencias topológicas.	El conjunto de datos no es conforme. El número de inconsistencias topológicas excede el nivel de conformidad de la calidad de datos.	El conjunto de datos no es conforme. El porcentaje de inconsistencias topológicas excede el nivel de conformidad de la calidad de datos.

D.5 Ejemplos de medidas de la calidad relativas a la exactitud posicional

La exactitud posicional es la exactitud de la posición de un objeto geográfico. Esta componente tiene los siguientes subelementos:

- exactitud absoluta o externa: proximidad de los valores de las coordenadas a los valores verdaderos, o considerados como tales;
- exactitud interna o relativa: proximidad de las posiciones relativas entre objetos geográficos en el conjunto de datos frente a sus posiciones relativas verdaderas, o aceptadas como tales;
- exactitud posicional de datos malla: proximidad de los valores de posición en estructura de malla regular a los valores reales o considerados como tales.

La tabla D.4 presenta algunos ejemplos relativos a los subelementos.

Tabla D.4 — Ejemplos de medidas de la calidad relativas a exactitud posicional

Componente de calidad de datos	Ejemplo 1	Ejemplo 2
DQ_Scope	Todos los nodos que conforman los límites de carreteras del conjunto de datos.	Todos los nodos que conforman los límites de carreteras del conjunto de datos delimitados por las longitudes +139 +140 y las latitudes +36,0 37,0
DQ_Element	3. Exactitud posicional	3. Exactitud posicional
DQ_Subelement	1. Exactitud absoluta o externa	1. Exactitud absoluta o externa
DQ_Measure		
DQ_MeasureDesc	EMC	Porcentaje de ítems cuyo error en coordenadas supera el límite especificado.
DQ_MeasureID	30101	30102
DQ_EvalMethod		
DQ_EvalMethodType	2. Externo	2. Externo
DQ_EvalMethodDesc	Para cada nodo se mide el error en la distancia entre los valores de las coordenadas absolutas del nodo en el conjunto de datos y en el universo de discurso. Se calcula el EMC a partir de los errores en distancia.	Para cada nodo se mide el error en la distancia entre los valores de las coordenadas absolutas del nodo en el conjunto de datos y en el universo de discurso. Se contabiliza el número de nodos cuyo error en distancia excede el límite de la especificación (por ejemplo, 1 m). Se divide el número de nodos no conformes entre el número total en el ámbito. El resultado se multiplica por 100.
DQ_QualityResult		
DQ_ValueType	2. Número	4. Porcentaje
DQ_Value	1,70 m	25%
DQ_ValueUnit	Metro	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06
DQ_ConformanceLevel	No especificado	No especificado
Ejemplo de parámetros del conjunto de datos	Omitido	Omitido
Ejemplo de interpretación del resultado de la calidad	El EMC en distancia de los nodos es de 1,70 m. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del EMC.	El 25% de los nodos en el ámbito posee un error en distancia mayor a 1m. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del porcentaje de casos.

Tabla D.4 (continuación)

Componente de calidad de datos	Ejemplo 3	Ejemplo 4
DQ_Scope		
DQ_Element	3. Exactitud posicional	3. Exactitud posicional
DQ_Subelement	2. Exactitud relativa o interna	2. Exactitud relativa o interna
DQ_Measure		
DQ_MeasureDesc	EMC	Porcentaje de ítems cuyo error en coordenadas supera un límite especificado.
DQ_MeasureID	30201	30202
DQ_EvalMethod		
DQ_EvalMethodType	2. Externo	2. Externo
DQ_EvalMethodDesc	Para cada nodo se mide el error en distancia entre los valores de las coordenadas relativas del nodo en el conjunto de datos y en el universo de discurso. Se calcula el EMC a partir de los errores en distancia.	Para cada nodo se mide el error en distancia entre los valores de coordenadas relativas del nodo en el conjunto de datos y en el universo de discurso. Se contabiliza el número de los nodos cuyo error en distancia excede el límite de la especificación (por ejemplo, 1m). Se divide el número de nodos no conformes entre el número total de nodos en el ámbito. El resultado se multiplica por 100.
DQ_QualityResult		
DQ_ValueType	2. Número	4. Porcentaje
DQ_Value	1,50m	20%
DQ_ValueUnit	Metro	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06
DQ_ConformanceLevel	No especificado	No especificado
Ejemplo de parámetros del conjunto de datos	Omitido	Omitido
Ejemplo de interpretación del resultado de la calidad	El EMC en distancia de los nodos es de 1,50m. Dado que no se especifica un nivel de conformidad de calidad, sólo se informa del EMC.	El 20% de los nodos en el ámbito de posee un error en distancia mayor a 1m. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del porcentaje.

Tabla D.4 (continuación)

Componente de calidad de datos	Ejemplo 5	Ejemplo 6	Ejemplo 7
DQ_Scope	Todos puntos de elevación de la malla del MDE en el conjunto de datos.	Todos puntos de elevación de la malla del MDE en la zona delimitada por las longitudes +139 +140 y las latitudes +36,0 +37,0	Todos puntos de elevación de la malla del MDE correspondientes a la ciudad de Bangkok, Tailandia.
DQ_Element	3. Exactitud posicional	3. Exactitud posicional	3. Exactitud posicional
DQ_Subelement	3. Exactitud posicional de datos malla	3. Exactitud posicional de datos malla	3. Exactitud posicional de datos malla
DQ_Measure			
DQ_MeasureDesc	EMC	Porcentaje de ítems con error en coordenadas superior a un límite especificado.	Conforme / no conforme
DQ_MeasureID	30301	30302	30303
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Para cada punto de la malla se mide la diferencia entre el valor de la altura absoluta del punto en el conjunto de datos y en el universo de discurso. Se calcula el EMC a partir de las diferencias en altura.	Para cada punto de la malla se mide la diferencia entre el valor de la altura absoluta del punto en el conjunto de datos y en el universo de discurso. Se contabiliza el número de puntos cuya diferencia de altura exceda el límite de la especificación (por ejemplo, 1m). Se divide el número de puntos no conformes entre el número total de puntos en el ámbito. El resultado se multiplica por 100.	Para cada punto de la malla se mide la diferencia entre el valor de la altura absoluta del punto en el conjunto de datos y en el universo de discurso. Se contabiliza el número de puntos cuya diferencia de altura exceda el límite de la especificación (por ejemplo, 1m). Se divide el número de puntos no conformes entre el número total de puntos en el ámbito. El resultado se multiplica por 100. Se compara el porcentaje de puntos no conformes con el nivel de conformidad de la calidad establecido.
DQ_QualityResult			
DQ_ValueType	2. Número	4. Porcentaje	1. Variable booleana
DQ_Value	0,8 m	8%	Falso
DQ_ValueUnit	Metro	Porcentaje de puntos cuyo error en altura supera el límite de la especificación.	No aplicable
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	No especificado	No especificado	Menos del 5% de los ítems pueden presentar errores mayores que el límite de la especificación.
Ejemplo de parámetros del conjunto de datos	Omitido	Omitido	Omitido
Ejemplo de interpretación del resultado de la calidad	El EMC de las alturas es 0.8 m. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del EMC.	El 8% de los puntos de la malla en el ámbito tienen errores en altura mayores a 1 m. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del porcentaje.	El conjunto de datos no es conforme. El porcentaje de puntos no conformes excede el nivel de conformidad de la calidad.

D.6 Ejemplos de medidas de la calidad relativas a la exactitud temporal

La exactitud temporal es la exactitud de los atributos y de las relaciones temporales. Sus subelementos son los siguientes:

- exactitud de una medida de tiempo: corrección de las referencias temporales de un instante (informando del error en una medida de tiempo);
- consistencia temporal: corrección de eventos ordenados o secuencias;
- validez temporal: validez de los datos respecto al tiempo.

La tabla D.5 presenta algunos ejemplos de exactitud temática para los subelementos de exactitud de una medida de tiempo, consistencia temporal y validez temporal.

Tabla D.5 — Ejemplos de medidas de la calidad relativas a exactitud temporal

Componente de calidad de datos	Ejemplo 1	Ejemplo 2	Ejemplo 3
DQ_Scope	Todos los datos de accidentes del tráfico del conjunto de datos.	Todos los datos de accidentes del tráfico en la zona delimitada por las longitudes +139 +140 y las latitudes +36,0 +37,0	Todos los datos de accidentes del tráfico en la ciudad de Londres, Reino Unido.
DQ_Element	4. Exactitud temporal	4. Exactitud temporal	4. Exactitud temporal
DQ_Subelement	1. Exactitud de una medida de tiempo	1. Exactitud de una medida de tiempo	1. Exactitud de una medida de tiempo
DQ_Measure			
DQ_MeasureDesc	EMC	Porcentaje de ítems con error en el atributo temporal mayor que el límite de la especificación	Conforme / no conforme
DQ_MeasureID	40101	40102	40103
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Para cada dato de un accidente de tráfico, se mide la diferencia entre el tiempo de ocurrencia en el conjunto de datos y en el universo de discurso. Se calcula el EMC de las diferencias entre los tiempos de ocurrencia.	Para cada dato de un accidente de tráfico, se mide la diferencia entre el tiempo de ocurrencia en el conjunto de datos y en el universo de discurso. Se contabiliza el número de accidentes cuyas diferencias en los tiempos de ocurrencia excede el límite de la especificación (por ejemplo, 2h). Se divide el número de datos de accidentes no conformes entre el número total de accidentes en el ámbito. El resultado se multiplica por 100.	Para cada dato de un accidente de tráfico, se mide la diferencia entre el tiempo de ocurrencia en el conjunto de datos y en el universo de discurso. Se contabiliza el número de accidentes cuyas diferencias en los tiempos de ocurrencia excede el límite de la especificación (por ejemplo, 2h). Se divide el número de datos de accidentes no conformes entre el número total de accidentes en el ámbito. El resultado se multiplica por 100. Se compara el porcentaje de accidentes no conformes con el nivel de conformidad de la calidad.
DQ_QualityResult			
DQ_ValueType	2. Número	4. Porcentaje	1. Variable booleana
DQ_Value	1,5 h	18	Falso
DQ_ValueUnit	Horas	Tanto por ciento	No aplicable
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	No especificado	No especificado	10%
Ejemplo de parámetros del conjunto de datos	Omitido	Omitido	Omitido
Ejemplo de interpretación del resultado de la calidad	El EMC de las diferencias de tiempo de ocurrencia es de 1,5 h. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del EMC.	El 18 % de los datos de accidentes dentro del ámbito tienen errores en el tiempo de ocurrencia mayores de 2 h. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa del EMC.	El conjunto de datos no es conforme. El porcentaje de datos de accidentes no conformes excede el nivel de conformidad de la calidad.

Tabla D.5 (continuación)

Componente de calidad de datos	Ejemplo 4	Ejemplo 5	Ejemplo 6
DQ_Scope	Todos los datos de eventos históricos en el conjunto de datos.	Todos los datos de eventos históricos en la zona delimitada por las longitudes +139 +140 y las latitudes +36.0 +37.0	Todos los datos de eventos históricos en China.
DQ_Element	4. Exactitud temporal	4. Exactitud temporal	4. Exactitud temporal
DQ_Subelement	2. Consistencia temporal	2. Consistencia temporal	2. Consistencia temporal
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de ítems con relaciones temporales inconsistentes.	Porcentaje de ítems con relaciones temporales inconsistentes.
DQ_MeasureID	40201	40202	40203
DQ_EvalMethod			
DQ_EvalMethodType	1. Externo	1. Externo	1. Externo
DQ_EvalMethodDesc	Se comprueba que cada evento histórico está ordenado correctamente respecto al resto de los datos de eventos.	Se comprueba que cada evento histórico está ordenado correctamente respecto al resto de los datos de eventos. Se contabilizan aquellos que no estén correctamente ordenados.	Se comprueba que cada evento histórico está ordenado correctamente respecto al resto de los datos de eventos. Se contabilizan aquellos que no estén correctamente ordenados. Se divide el resultado por el total de ítems en el ámbito y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	3	60%
DQ_ValueUnit	No aplicable	Inconsistencias temporales	Tanto por ciento de inconsistencias temporales
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún ítem puede tener inconsistencia temporal.	Ningún ítem puede tener inconsistencia temporal.	El cero por ciento de los ítems pueden tener inconsistencia temporal.
Ejemplo de parámetros del conjunto de datos	Hay 5 eventos históricos en el ámbito. La secuencia correcta es {A,B,C,D,E}. En el conjunto de datos, los cinco eventos se registran en el orden {A,B,D,E,C}. Un evento individual (A,B,C,D,E) se considera como un ítem. Los ítems con orden inconsistentes son (C,D,E).	Hay 5 eventos históricos en el ámbito. La secuencia correcta es {A,B,C,D,E}. En el conjunto de datos, los cinco eventos se registran en el orden {A,B,D,E,C}. Un evento individual (A,B,C,D,E) se considera como un ítem. Los ítems con orden inconsistentes son (C,D,E).	Hay 5 eventos históricos en el ámbito. La secuencia correcta es {A,B,C,D,E}. En el conjunto de datos, los cinco eventos se registran en el orden {A,B,D,E,C}. Un evento individual (A,B,C,D,E) se considera como un ítem. Los ítems con orden inconsistentes son (C,D,E).
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Se ha encontrado inconsistencia temporal.	El conjunto de datos no es conforme. El número de inconsistencias temporales excede el nivel de conformidad de la calidad.	El conjunto de datos no es conforme. El número de inconsistencias temporales excede el nivel de conformidad de la calidad.

Tabla D.5 (continuación)

Componente de calidad de datos	Ejemplo 7	Ejemplo 8	Ejemplo 9
DQ_Scope	Todos los datos de precios de los terrenos en el conjunto de datos.	Todos los datos de precios de los terrenos en la zona delimitada por las longitudes +139 +140 y las latitudes +36,0 +37,0	Todos los datos de precios de los terrenos en la ciudad de Tokio, Japón.
DQ_Element	4. Exactitud temporal	4. Exactitud temporal	4. Exactitud temporal
DQ_Subelement	3. Validez temporal	3. Validez temporal	3. Validez temporal
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de ítems con invalidez temporal	Porcentaje de ítems con invalidez temporal
DQ_MeasureID	40301	40302	40303
DQ_EvalMethod			
DQ_EvalMethodType	1. Interno	1. Interno	1. Interno
DQ_EvalMethodDesc	Se comprueban los datos de precio de los terrenos para asegurar que fueron recogidos en 1995.	Se comprueban los datos de precio de los terrenos para asegurar que fueron recogidos en 1995. Se cuentan aquellos que no fueron capturados en 1995.	Se comprueban los datos de precio de los terrenos para asegurar que fueron recogidos en 1995. Se cuentan aquellos que no fueron capturados en 1995. Se divide el resultado por el número total de casos en el ámbito y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	5	5%
DQ_ValueUnit	No aplicable	Invalidez temporal	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún ítem puede presentar invalidez temporal.	A lo sumo 10 ítems pueden tener invalidez temporal.	Menos del 10% de los ítems pueden tener invalidez temporal.
Ejemplo de parámetros del conjunto de datos	Hay 100 ítems en el conjunto de datos con fecha de captura de 1995; 95 de ellos fueron recogidos realmente en 1995; 5 se recolectaron en 1985.	Hay 100 ítems en el conjunto de datos con fecha de captura de 1995; 95 de ellos fueron recogidos realmente en 1995; 5 se recolectaron en 1985.	Hay 100 ítems en el conjunto de datos con fecha de captura de 1995; 95 de ellos fueron recogidos realmente en 1995; 5 se recolectaron en 1985.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Al menos un ítem presenta invalidez temporal.	El conjunto de datos es conforme. Hay menos de diez ítems con invalidez temporal.	El conjunto de datos es conforme. Hay menos del 10 por ciento de ítems con invalidez temporal.

D.7 Ejemplos de medidas de la calidad relativas a la exactitud temática

La exactitud temática es la exactitud de atributos cuantitativos y el grado de corrección de atributos no cuantitativos, así como de la clasificación de los objetos geográficos y de sus relaciones. La exactitud temática tiene los siguientes subelementos:

- corrección de la clasificación: comparación de las clases asignadas a los objetos geográficos, o a sus atributos, con el universo de discurso (verdad terreno o un conjunto de datos de referencia);
- corrección de atributos no cuantitativos: grado de corrección de atributos no cuantitativos;
- corrección de atributos cuantitativos: exactitud de atributos cuantitativos.

La tabla D.6 presenta algunos ejemplos de exactitud temática para los subelementos.

Tabla D.6 — Ejemplos de medidas de la calidad relativas a exactitud temática

Componente de calidad de datos	Ejemplo 1	Ejemplo 2	Ejemplo 3																																																																																										
DQ_Scope	Todos los ítems clasificados como A, B y C en el conjunto de datos.	Todos los ítems clasificados como A, B y C en la zona delimitada por las longitudes +139 +140 y las latitudes +36,0 +37,0.	Todos los ítems clasificados como A, B y C en Arabia Saudita.																																																																																										
DQ_Element	5. Exactitud temática	5. Exactitud temática	5. Exactitud temática																																																																																										
DQ_Subelement	1. Corrección de la clasificación	1. Corrección de la clasificación	1. Corrección de la clasificación																																																																																										
DQ_Measure																																																																																													
DQ_MeasureDesc	Conforme / no conforme	Porcentaje correctamente clasificado (PCC).	Matriz de porcentaje de confusión.																																																																																										
DQ_MeasureID	50101	50102	50103																																																																																										
DQ_EvalMethod																																																																																													
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo																																																																																										
DQ_EvalMethodDesc	Para cada ítem del conjunto de datos se compara la clase asignada con la clase real en el universo de discurso.	Para cada ítem del conjunto de datos se compara la clase asignada con la clase real en el universo de discurso. Se cuentan los ítems correctamente clasificados. El resultado se divide por el total de ítems en el ámbito y se multiplica por 100.	Para cada ítem del conjunto de datos se compara la clase asignada con la clase real en el universo de discurso. Se genera una matriz de $N(i,j)$, donde $N(i,j)$ es el número de ítems de la clase (i) que son clasificados en clase (j) en el conjunto de datos. Se divide $N(i,j)$ entre el número total de ítems de clase (i) y se multiplica por 100.																																																																																										
DQ_QualityResult																																																																																													
DQ_ValueType	1. Variable booleana	4. Porcentaje	8. Matriz																																																																																										
DQ_Value	Falso	60%	<table border="1"> <thead> <tr> <th colspan="5">Clases del conjunto de datos</th> </tr> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>%</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>70</td> <td>20</td> <td>10</td> <td>100</td> </tr> <tr> <th>B</th> <td>20</td> <td>40</td> <td>40</td> <td>100</td> </tr> <tr> <th>C</th> <td>20</td> <td>20</td> <td>60</td> <td>100</td> </tr> <tr> <th>%</th> <td>100</td> <td>100</td> <td>100</td> <td></td> </tr> </tbody> </table>	Clases del conjunto de datos						A	B	C	%	A	70	20	10	100	B	20	40	40	100	C	20	20	60	100	%	100	100	100																																																													
Clases del conjunto de datos																																																																																													
	A	B	C	%																																																																																									
A	70	20	10	100																																																																																									
B	20	40	40	100																																																																																									
C	20	20	60	100																																																																																									
%	100	100	100																																																																																										
DQ_ValueUnit	No aplicable	Tanto por ciento	Tanto por ciento																																																																																										
DQ_Date	2000-03-06	2000-03-06	2000-03-06																																																																																										
DQ_ConformanceLevel	Ningún ítem puede tener errores de clasificación.	Más del 80 % de los ítems deben estar correctamente clasificados	No especificado.																																																																																										
Ejemplo de parámetros del conjunto de datos	<table border="1"> <thead> <tr> <th colspan="5">Clases del conjunto de datos</th> </tr> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>Casos</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>7</td> <td>2</td> <td>1</td> <td>10</td> </tr> <tr> <th>B</th> <td>1</td> <td>2</td> <td>2</td> <td>5</td> </tr> <tr> <th>C</th> <td>1</td> <td>1</td> <td>3</td> <td>5</td> </tr> <tr> <th>Casos</th> <td>9</td> <td>5</td> <td>6</td> <td>20</td> </tr> </tbody> </table>	Clases del conjunto de datos						A	B	C	Casos	A	7	2	1	10	B	1	2	2	5	C	1	1	3	5	Casos	9	5	6	20	<table border="1"> <thead> <tr> <th colspan="5">Clases del conjunto de datos</th> </tr> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>Casos</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>7</td> <td>2</td> <td>1</td> <td>10</td> </tr> <tr> <th>B</th> <td>1</td> <td>2</td> <td>2</td> <td>5</td> </tr> <tr> <th>C</th> <td>1</td> <td>1</td> <td>3</td> <td>5</td> </tr> <tr> <th>Casos</th> <td>9</td> <td>5</td> <td>6</td> <td>20</td> </tr> </tbody> </table>	Clases del conjunto de datos						A	B	C	Casos	A	7	2	1	10	B	1	2	2	5	C	1	1	3	5	Casos	9	5	6	20	<table border="1"> <thead> <tr> <th colspan="5">Clases del conjunto de datos</th> </tr> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>Casos</th> </tr> </thead> <tbody> <tr> <th>A</th> <td>7</td> <td>2</td> <td>1</td> <td>10</td> </tr> <tr> <th>B</th> <td>1</td> <td>2</td> <td>2</td> <td>5</td> </tr> <tr> <th>C</th> <td>1</td> <td>1</td> <td>3</td> <td>5</td> </tr> <tr> <th>Casos</th> <td>9</td> <td>5</td> <td>6</td> <td>20</td> </tr> </tbody> </table>	Clases del conjunto de datos						A	B	C	Casos	A	7	2	1	10	B	1	2	2	5	C	1	1	3	5	Casos	9	5	6	20
Clases del conjunto de datos																																																																																													
	A	B	C	Casos																																																																																									
A	7	2	1	10																																																																																									
B	1	2	2	5																																																																																									
C	1	1	3	5																																																																																									
Casos	9	5	6	20																																																																																									
Clases del conjunto de datos																																																																																													
	A	B	C	Casos																																																																																									
A	7	2	1	10																																																																																									
B	1	2	2	5																																																																																									
C	1	1	3	5																																																																																									
Casos	9	5	6	20																																																																																									
Clases del conjunto de datos																																																																																													
	A	B	C	Casos																																																																																									
A	7	2	1	10																																																																																									
B	1	2	2	5																																																																																									
C	1	1	3	5																																																																																									
Casos	9	5	6	20																																																																																									
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. 8 ítems están mal clasificados.	El conjunto de datos no es conforme. El 40% de los ítems están mal clasificados.	Dado que no se especifica un nivel de conformidad de la calidad, se informa la matriz de confusión.																																																																																										

Tabla D.6 (continuación)

Componente de calidad de datos	Ejemplo 4	Ejemplo 5	Ejemplo 6
DQ_Scope	Todos los ítems del conjunto de datos con nombres geográficos.	Todos los ítems con nombres geográficos en la zona delimitada por las longitudes +139 +140 y las latitudes +36.0 +37.0	Todos los ítems con nombres geográficos en la ciudad de Lisboa, Portugal.
DQ_Element	5. Exactitud temática	5. Exactitud temática	5. Exactitud temática
DQ_Subelement	3. Corrección de atributos no cuantitativos.	3. Corrección de atributos no cuantitativos.	3. Corrección de atributos no cuantitativos.
DQ_Measure			
DQ_MeasureDesc	Conforme / no conforme	Número de ítems con nombres geográficos incorrectos.	Porcentaje de ítems con nombres geográficos incorrectos.
DQ_MeasureID	50201	50202	50203
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Se comparan los nombres geográficos del conjunto de datos con los del universo de discurso.	Se comparan los nombres geográficos del conjunto de datos con los del universo de discurso. Se contabiliza el número de ítems con nombres geográficos incorrectos.	Se comparan los nombres geográficos del conjunto de datos con los del universo de discurso. Se contabiliza el número de ítems con nombres geográficos incorrectos. Se divide el resultado entre el número total de ítems en el ámbito y se multiplica por 100.
DQ_QualityResult			
DQ_ValueType	1. Variable booleana	2. Número	4. Porcentaje
DQ_Value	Falso	5	5%
DQ_ValueUnit	No aplicable	Número de ítems con nombres geográficos incorrectos.	Tanto por ciento
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	Ningún ítem puede tener el nombre geográfico incorrecto.	Menos de 3 ítems pueden tener nombres geográficos incorrectos.	Menos del 3% de los ítems pueden tener nombres geográficos incorrectos.
Ejemplo de parámetros del conjunto de datos	Hay 100 ítems con nombres geográficos en el ámbito; 5 nombres están mal escritos.	Hay 100 ítems con nombres geográficos en el ámbito; 5 nombres están mal escritos.	Hay 100 ítems con nombres geográficos en el ámbito; 5 nombres están mal escritos.
Ejemplo de interpretación del resultado de la calidad	El conjunto de datos no es conforme. Al menos un ítem tiene incorrecto el nombre geográfico.	El conjunto de datos no es conforme. Más de 3 ítems presentan nombres geográficos incorrectos.	El conjunto de datos no es conforme. Más del 3% de los ítems presentan nombres geográficos incorrectos.

Tabla D.6 (continuación)

Componente de calidad de datos	Ejemplo 7	Ejemplo 8	Ejemplo 9
DQ_Scope	Todos los ítems del conjunto de datos que poseen atributos de temperatura.	Todos los ítems que poseen atributos de temperatura en la zona delimitada por las longitudes +139 +140 y las latitudes +36,0 +37,0	Todos los ítems que poseen atributos de temperatura en la ciudad de Munich, Alemania.
DQ_Element	5. Exactitud temática	5. Exactitud temática	5. Exactitud temática
DQ_Subelement	3. Exactitud del atributo cuantitativo	3. Exactitud del atributo cuantitativo	3. Exactitud del atributo cuantitativo
DQ_Measure			
DQ_MeasureDesc	EMC	Porcentaje de ítems con error en temperatura mayor que el límite de la especificación.	Conforme / no conforme
DQ_MeasureID	50301	50302	50303
DQ_EvalMethod			
DQ_EvalMethodType	2. Externo	2. Externo	2. Externo
DQ_EvalMethodDesc	Para cada ítem se mide la diferencia entre el valor de temperatura del conjunto de datos y el del universo de discurso. A partir de las diferencias se calcula el EMC.	Para cada ítem se mide la diferencia entre el valor de temperatura del conjunto de datos y el del universo de discurso. A partir de las diferencias se calcula el EMC. Se cuenta el número de ítems cuya diferencia de temperatura excede el límite de la especificación (por ejemplo, 1°). Se divide el número de ítems no conformes entre el número de ítems total en el ámbito. El resultado se multiplica por 100.	Para cada ítem se mide la diferencia entre el valor de temperatura del conjunto de datos y el del universo de discurso. A partir de las diferencias se calcula el EMC. Se cuenta el número de ítems cuya diferencia de temperatura excede el límite de la especificación (por ejemplo, 1°). Se divide el número de ítems no conformes entre el número de ítems total en el ámbito. El resultado se multiplica por 100. Se compara el porcentaje de ítems no conformes con el nivel de conformidad de la calidad.
DQ_QualityResult			
DQ_ValueType	2. Número	4. Porcentaje	1. Variable booleana
DQ_Value	0.5	5	Falso
DQ_ValueUnit	Grado	Tanto por ciento de ítems con error en temperatura mayor que el límite de la especificación.	No aplicable
DQ_Date	2000-03-06	2000-03-06	2000-03-06
DQ_ConformanceLevel	No especificado	No especificado	Menos de un 1% de los ítems pueden tener error de temperatura mayor que el límite de la especificación.
Ejemplo de parámetros del conjunto de datos	Omitido	Omitido	Omitido
Ejemplo de interpretación del resultado de la calidad	El EMC de la temperatura es 0,5°. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa el EMC.	El 5% de los ítems en el ámbito tienen errores en temperatura mayor de 1°. Dado que no se especifica un nivel de conformidad de la calidad, sólo se informa el porcentaje.	El conjunto de datos no es conforme. El porcentaje de ítems no conformes excede el nivel de conformidad de la calidad.

ANEXO E (Informativo)

GUÍA PARA LA APLICACIÓN DE MÉTODOS DE MUESTREO A CONJUNTOS DE DATOS GEOGRÁFICOS

E.1 Introducción

Este anexo proporciona directrices para definir muestras e idear métodos de muestreo. En los muestreos para evaluar la conformidad de un producto frente a una especificación se pueden aplicar la serie de Normas NC ISO 2859 y la Norma ISO 3951-1. Originariamente, estas normas fueron desarrolladas para un uso no espacial. Este anexo describe como aplicar la serie de Normas NC ISO 2859 y la Norma ISO 3951-1, y otras técnicas de muestreo espaciales a la información geográfica.

E.2 Lote e ítem

El lote y el ítem son conceptos importantes en los métodos de inspección por muestreo especificados en la serie de Normas NC ISO 2859 y en la Norma ISO 3951-1. Un lote es la mínima unidad para la que la calidad puede ser evaluada. Un ítem es la mínima unidad a inspeccionar y debería ser definido por el productor de los datos de acuerdo con la especificación del producto.

E.3 Tamaño de la muestra

El tamaño de una población, y consecuentemente el tamaño de las muestras, puede definirse de acuerdo a diferentes bases sobre los ítems. La definición de un tamaño de muestra requiere una indicación explícita de los ítems. La tabla E.1 presenta ejemplos de diferentes bases.

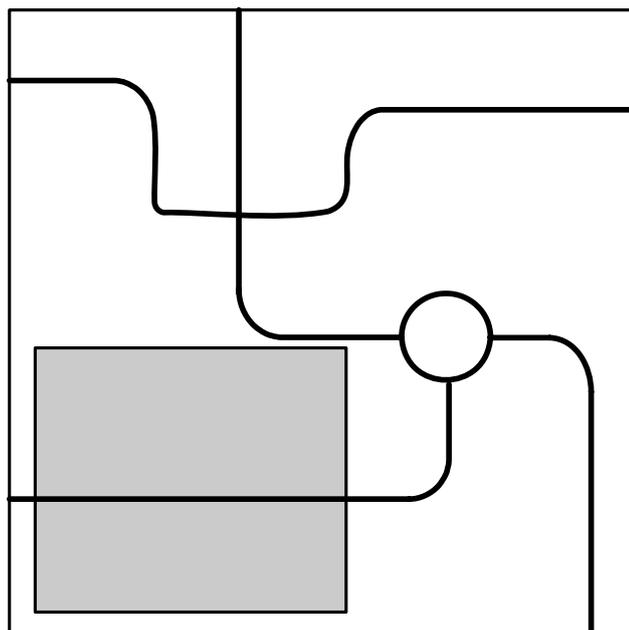
La diferencia entre las perspectivas se ilustra en la figura E.1. La figura en su conjunto representa los datos dentro ámbito de la calidad. En la figura se marca una posible área de muestreo de aproximadamente el 15% del área total del ámbito de calidad, pero en este caso sólo se encuentran dentro de esta área alrededor del 10% de la longitud de las curvas y un 0% de los vértices.

Para salvar las dificultades del muestreo, como las comentadas sobre la figura E.1, el tamaño y localización de la muestra podría ser definido usando una combinación de diferentes criterios, reforzando de esta manera la representatividad de la muestra.

Ejemplo: La muestra debería incluir un 10% del área cubierta por el conjunto de datos y contener no menos de un 5% de la longitud total de las curvas que describen los objetos del conjunto de datos.

Tabla E.1— Diferentes bases para definir una población

Base	Tamaño del conjunto de datos	Tamaño de la muestra
Objetos geográficos	Número de objetos geográficos de un tipo determinado	Número de objetos geográficos de un tipo determinado, expresado como porcentaje del número total de objetos.
Área cubierta	Área cubierta por el conjunto de datos	Área cubierta por la muestra expresada como porcentaje del área total.
Curvas	Longitud total de las curvas en el conjunto de datos	Longitud de las curvas de la muestra expresadas como porcentaje de la longitud total.
Vértices	Número total de vértices que describen las curvas o áreas en el conjunto de datos	Número de vértices en la muestra expresado como un porcentaje del número total de vértices.



NOTA - El ámbito de la calidad de datos es el área del marco exterior. El área de muestreo es la del rectángulo sombreado.

Figura E.1 — Efecto de la localización del área de muestreo sobre la representatividad de los ítems en la muestra

E.4 Estrategias de muestreo

E.4.1 Introducción

Teniendo en consideración aspectos particulares de los datos geográficos, este capítulo presenta una guía para la definición de muestras de los métodos de muestreo. Las estrategias de muestreo que se describen en este anexo se muestran gráficamente en la figura E.2. Hay dos aspectos a considerar para una estrategia de muestreo, los ítems a muestrear (área u objeto geográfico) y la manera de seleccionar los ítems (por probabilidad o por criterios).

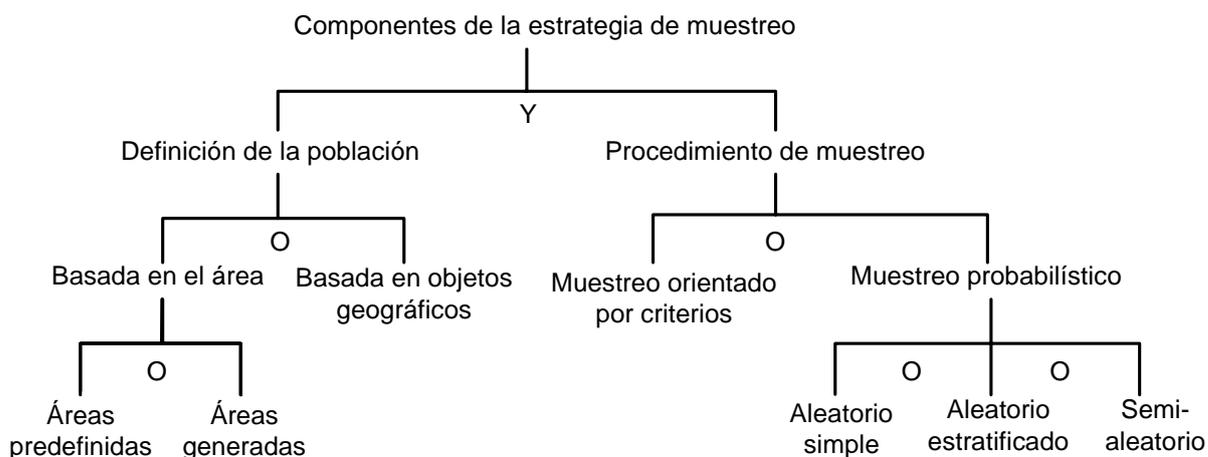


Figura E.2 — Relaciones entre estrategias de muestreo

E.4.2 Muestreo probabilístico frente a muestreo orientado por criterios

E.4.2.1 Diferencias

El muestreo probabilístico aplica la teoría de muestras y conlleva una selección aleatoria de los ítems de la muestra. La característica esencial de un muestreo probabilístico es que cada uno de los miembros de la población de la que se selecciona la muestra tiene una probabilidad conocida de ser seleccionado. Cuando se usa el muestreo probabilístico, se pueden realizar inferencias estadísticas sobre la población muestreada. El diseño de muestras orientado por criterios conlleva la selección de las muestras basada en el conocimiento experto o en el juicio profesional.

E.4.2.2 Muestreo aleatorio simple

El muestreo aleatorio simple se basa en la probabilidad y supone la selección aleatoria de las muestras. Las muestras (por ejemplo, objetos geográficos, localizaciones, tiempos) se determinan usando números aleatorios para identificar los ítems, y todas las selecciones son igualmente probables. El muestreo aleatorio simple es útil cuando la población de interés es relativamente homogénea respecto a las características a muestrear, es decir, sin patrones notables ni agrupamientos. Este método puede resultar en una cobertura no representativa de un área, esto es, es posible que la muestra seleccionada proceda sólo de una parte del área.

E.4.2.3 Muestreo aleatorio estratificado

El muestreo estratificado requiere la separación de la población en estratos, o subpoblaciones, que no se solapan, de forma que en cada estrato exista mayor homogeneidad entre sus ítems que entre los de estratos diferentes. Cuando se aplica a una misma población, esta estrategia tiene el potencial de permitir alcanzar mayores precisiones en la estimación de medias y varianzas que las estrategias no estratificadas.

E.4.2.4 Muestreo semialeatorio

El muestreo semialeatorio o sistemático aplica la selección aleatoria a los ítems iniciales del muestreo (por ejemplo: situación, tiempo, objeto geográfico) y reglas para la selección de los restantes. Un ejemplo de muestreo semialeatorio o sistemático es una malla de muestreo donde la posición inicial de la malla se determina aleatoriamente y las muestras se toman a intervalos de espacio regulares (celdas de la malla). El muestreo sistemático en malla se utiliza para la búsqueda de agrupaciones y para inferir medias, porcentajes y otros parámetros. Es útil para estimar tendencias espaciales o patrones. Este método proporciona una forma práctica y fácil de asegurar la cobertura de un área.

E.4.3 Muestreo orientado por objetos geográficos frente al muestreo orientado por áreas

E.4.3.1 Muestreo orientado por objetos geográficos (muestreo no espacial)

En una estrategia de muestreo orientada por objetos geográficos, se seleccionan los ítems de la muestra según los atributos no espaciales de los objetos geográficos y no por su situación espacial. Dentro del ámbito, la muestra puede ser elegida aleatoriamente asumiendo características de producción homogéneas para todo el ámbito. En algunos casos, el muestreo aleatorio simple puede generar una muestra no satisfactoria en el caso de que sólo exista homogeneidad en subconjuntos, por ejemplo, patrones globales o agrupamientos que tengan lugar en las características a muestrear, cuando el método requiere una distribución homogénea de las muestras. En estos casos, un muestreo estratificado o semialeatorio puede proporcionar mejores resultados.

NOTA - Si el método de muestreo se define mediante una selección aleatoria de objetos geográficos, existe el riesgo de que se seleccione una muestra que se concentre en un área reducida (lo cual no es aceptable).

El muestreo semialeatorio puede utilizarse para asegurar la verificación de diferentes criterios sobre el tamaño de muestra y/o localización, para satisfacer limitaciones complementarias sobre las muestras o para reducir los costes del proceso de inspección.

Ejemplo: Una compañía energética necesita evaluar la corrección de los atributos capturados para diferentes tipos de objetos geográficos. Se consideraron dos métodos: una selección aleatoria y una selección semialeatoria (seleccionando aleatoriamente los objetos geográficos de un tipo y capturando los objetos de los diferentes tipos restantes en la vecindad del primero hasta que se consigan las muestras necesarias para cada tipo), llevando a una reducción de los costos de inspección.

E.4.3.2 Muestreo orientado por áreas (muestreo espacial)

En una estrategia de muestreo basado en el área, la selección de las unidades de muestreo se basa en consideraciones espaciales. Las unidades de muestreo pueden ser áreas geográficas existentes (por ejemplo, áreas políticas o estadísticas) o cualquier otra partición del universo de discurso a la cual se dirige el muestreo. Este tipo de muestreo puede ser usado como la primera etapa del muestreo, que será seguido por un muestreo orientado por objetos geográficos dentro de cada una de las subáreas.

Ejemplo 1: Una selección aleatoria de áreas de cuadrículas UTM de 1 km x 1 km para evaluar los atributos de los objetos contenidos en el área de interés.

Ejemplo 2: La figura E.3 ilustra el resultado de la definición de las áreas a someter a la inspección, que se han obtenido mediante la generación aleatoria de las coordenadas del centro de unos rectángulos de superficies iguales (con la limitación de no solaparse).

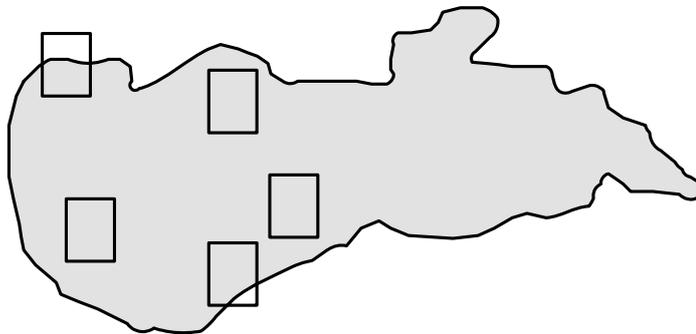


Figura E.3 — Ejemplo de muestreo aleatorio orientado por áreas

Cuando es importante cubrir toda el área, la situación de las muestras debería ser determinada de acuerdo a un patrón regular o semi-regular. La figura E.4 ilustra un ejemplo de un muestreo semi-aleatorio (sistemático) con los objetos geográficos muestreados distribuidos según un patrón regular usado para evaluar la exactitud posicional del conjunto de datos.

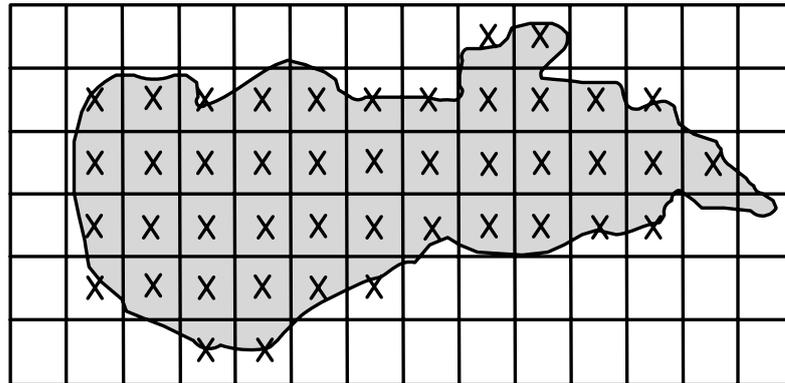


Figura E.4 — Ejemplo de muestreo regular y no aleatorio basado en las áreas

NOTA - El símbolo X denota las celdas de la malla elegidas mediante una regla para su inclusión en la muestra.

Si la distribución de los objetos geográficos no es homogénea, en muestreos semi-aleatorios se puede necesitar una partición espacial con tamaños distintos en diferentes áreas del conjunto de datos. Cuando se utiliza una malla de tamaño constante se necesita una regla para incluir o excluir celdas que no estén completamente incluidas dentro del área de interés.

E.5 Muestreos de base probabilística

E.5.1 Consideraciones generales

Cuando se realiza un muestreo se deben tener en cuenta las siguientes consideraciones:

- Las áreas cubiertas por un conjunto de datos geográficos pueden formar un espacio continuo. Cuando se divide el conjunto de datos en lotes se debería prestar especial atención a la omisión o comisión de ítems que cruzan los límites de los lotes.
- La calidad de los datos geográficos puede verse afectada por una gran variedad de factores, incluyendo la calidad de las fuentes de datos y las habilidades de los operarios. El productor de los datos debería ser cuidadoso a la hora de definir los lotes para alcanzar una cierta homogeneidad en términos de calidad.

E.5.2 Normas existentes para la inspección por muestreo

Basándose en las características de la producción y de acuerdo con las especificaciones de producto, del conjunto de normas internacionales se deben elegir aquellas adecuadas para la inspección por muestreo. La Norma NC ISO 2859-1 se aplica principalmente a la inspección de series continuas de lotes. La Norma ISO 2859-2 puede aplicarse a lotes individuales o aislados, mientras que la Norma NC ISO 2859-3 se aplica a procedimientos de muestreo por lotes salteados. La Norma ISO 3951-1 es para la inspección por variables mediante porcentajes de ítems no conformes.

El nivel de conformidad de la calidad de un conjunto de datos se especifica como LCA (Límite de Calidad Aceptable) en la Norma ISO 3534-2. Fue previamente denominado como NCA (Nivel de Calidad Aceptable) en las Normas NC ISO 2859-1, ISO 2859-3 e ISO 3951, y como CL (Calidad Límite) en el caso de la Norma NC ISO 2859-2 basada en la especificación del producto.

Cuando se aplica la serie de Normas ISO 2859, basada en la especificación del producto, se deberían establecer límites de la especificación para determinar la conformidad de cada ítem. En la aplicación de la Norma ISO 3951-1, los estadísticos de la calidad deberían ser establecidos basándose en la especificación del producto.

E.5.3 El proceso de muestreo

E.5.3.1 Definición de los ítems

Los ítems deberían ser definidos de acuerdo a la especificación del producto o a sus requisitos. Si los ítems no conformes poseen una alta correlación estadística, son tratados como un ítem individual.

E.5.3.2 Definición del ámbito de la calidad del conjunto de datos a inspeccionar

Si el ámbito de la calidad de datos no es homogéneo, debería ser dividido en subconjuntos homogéneos. Estos subconjuntos deberían ser considerados como ámbitos separados.

NOTA - Se puede hablar de homogeneidad cuando ocurren las siguientes condiciones:

- La fuente de datos de la producción tiene casi la misma calidad;
- Los sistemas de producción (hardware, software, habilidad del operario) son esencialmente los mismos;
- Otros factores que pueden afectar a la probabilidad de ocurrencia de no conformidades, como la complejidad y la densidad de objetos geográficos, son esencialmente los mismos.

E.5.3.3 División en lotes del ámbito de la calidad de datos

Los lotes se generan dividiendo el ámbito de la calidad. Cuando exista una fuerte autocorrelación espacial positiva en la ocurrencia de no conformidades es preferible reducir el tamaño de los lotes.

E.5.3.4 División del lote en unidades de muestreo

Una unidad de muestreo puede ser un área geográfica existente o cualquier otra partición del universo de discurso al que se dirige la inspección. Cuando la unidad de muestreo es un área geográfica, se deberían establecer reglas para la inclusión parcial de ítems en una unidad de muestreo.

E.5.3.5 Selección de unidades de muestreo para la inspección por muestreo aleatorio simple

El número total de ítems que pertenecen a las unidades de muestreo seleccionadas debería ser el especificado por las normas internacionales pertinentes.

NOTA - Si los lotes son estadísticamente heterogéneos, el muestreo aleatorio simple no puede ser aplicado con el mismo nivel de muestreo. La serie de Normas ISO 2859 permite adicionalmente los muestreos estratificados.

E.5.3.6 Inspección de las unidades de muestreo seleccionadas

Se inspeccionan todos los ítems que pertenecen a las unidades de muestreo seleccionadas. Los ítems del conjunto de datos son comparados con el universo de discurso de acuerdo con la medida de calidad elegida.

ANEXO F
(Informativo)

EJEMPLO DE PRUEBA DE EXACTITUD TEMÁTICA Y COMPLETITUD

F.1 Introducción

Este ejemplo se basa en las técnicas usadas para un levantamiento topográfico nacional realizado en Europa. A partir del levantamiento topográfico nacional se está produciendo una Base de Datos Topográficos (BDT). Esta BDT se emplea para hacer mapas topográficos impresos a escala 1:10 000. La BDT también se emplea para confeccionar algunos conjuntos de datos generalizados. Los niveles de conformidad de la calidad han sido definidos en la especificación de producto.

El objetivo de este ejemplo es ilustrar un procedimiento de evaluación de la calidad que se utiliza para la medida de la exactitud temática y la completitud en un conjunto de datos de una BDT. La exactitud posicional no se estudia porque no es el objetivo de este tipo de informe, sin embargo, en general, también se evalúa mediante levantamientos en campo empleando muestreos no aleatorios.

En F.2 se explica el proceso de evaluación de calidad por parte del productor de los datos y en F.6 la elaboración del reporte de información sobre la calidad.

F.2 Proceso de evaluación de la calidad

La tabla F.1 muestra las operaciones del proceso de evaluación de la calidad realizado en este ejemplo.

Tabla F.1— Proceso de evaluación de la calidad

Pasos del proceso	Ejemplo 1	Ejemplo 2
Identificar un elemento de la calidad de datos aplicable	Completitud	Exactitud temática
Identificar un subelemento aplicable	Comisión y omisión	Corrección de la clasificación
Identificar un ámbito	Base de datos topográficos / conjuntos de datos seleccionados (hojas del mapa 1:10 000)	
Identificar una medida	Conformidad/número de errores	Conformidad/número de errores
Seleccionar y aplicar un método de evaluación	Evaluación de la calidad directa externa	Evaluación de la calidad directa externa
Describir el método de muestreo	Muestreo polietápico	Muestreo polietápico
Especificar el nivel de conformidad de la calidad	NCA = 4 (nivel de calidad aceptable)	NCA = 4
Determinar el resultado cuantitativo	véase el capítulo F.4 y la figura F.1	
Evaluar la conformidad según la especificación de producto	véase el capítulo F.5 y la figura F.2	
Realizar un informe de los resultados de la evaluación de la calidad	véase el capítulo F.6	

F.3 Método para la evaluación de la calidad de datos

F.3.1 Procedimiento de muestreo

Las pruebas de completitud y exactitud temática se han llevado a cabo aplicando los principios de la Norma NC ISO 2859-1, *Procedimientos de muestreo para la inspección por atributos - Parte 1: Planes de muestreo indexados por el nivel de calidad aceptable (NCA) para una inspección por lote*. En la tabla F.2 se explica el procedimiento de muestreo conforme a esta norma internacional.

Tabla F.2 — Procedimiento para el muestreo

Pasos del proceso	Ejemplo
Definir un método de muestreo	Muestreo polietápico (por etapas). Se seleccionan suficientes unidades de muestreo para satisfacer la razón de muestreo. El muestreo está basado en la ponderación de los objetos geográficos.
Definir los ítems	Todos los objetos geográficos.
Dividir el ámbito de la calidad de datos (la población) en lotes	Número de conjuntos de datos.
Dividir los lotes en unidades de muestreo	N números de cuadrados de 1 km x 1 km.
Definir la razón de muestreo o el tamaño de la muestra	El tamaño de la muestra depende del valor NCA para el lote.
Seleccionar las unidades de muestreo	Se selecciona el número requerido de unidades de muestreo tal que satisfaga la razón de muestreo o el tamaño de la muestra para los ítems.
Inspeccionar los ítems en las unidades de muestreo	Se inspecciona cada ítem en las unidades de muestreo.

F.3.2 Métodos de muestreo

Si los requisitos de calidad para el objeto geográfico especifican una no conformidad por cada 100 unidades (NCA = 1), entonces se examinan todos los objetos geográficos a partir de la fuente de datos. La inspección por muestreo se realiza cuando el NCA = 4 ó 15. El nivel de inspección es el nivel 1 de inspección general, o sea, el programa de muestreo simple para una inspección normal. En la Norma ISO 2859-1 hay tres niveles generales de inspección (I, II, III) y cuatro niveles especiales adicionales (S-1, S-2, S-3 y S-4). En general, el nivel de inspección define el tamaño de la muestra a partir del tamaño del lote.

Un lote utilizado para el control debería estar formado por conjuntos de datos que, en la medida de lo posible, hayan sido producidos al mismo tiempo y mediante los mismos métodos. A partir del lote, se seleccionan las unidades de muestreo de N números de cuadrados de 1 km x 1 km, de forma tal que el número de objetos geográficos en la muestra sea suficiente para un NCA igual a 4.

Para realizar el muestreo se emplean los pesos preestablecidos por defecto para los objetos geográficos. En la ponderación, se asigna un valor por defecto de 1 a los objetos geográficos que sean abundantes en el lote, o a aquellos para los que no se haya establecido un valor de NCA para la completitud. A aquellos objetos geográficos con NCA = 4 ó 15 para la completitud se les asigna un peso de 2 o de 3. A aquellos objetos geográficos que son escasos en el lote se les da un peso de 3. Para el resto se usa un peso de 2. Existe un procedimiento para las

situaciones en las que no se consigue obtener el tamaño de muestra requerido. En general el programa trata de usar los niveles de inspección especiales S-1 a S-4.

Todos los objetos geográficos presentes en los cuadrados de muestreo se examinan en campo. Un objeto geográfico es no conforme si falta (omisión) o si el objeto geográfico que aparece en el conjunto de datos no existe en el campo (comisión).

F.3.3 Inspección completa

Para aquellos objetos geográficos que tienen un requisito de calidad de NCA = 1 se lleva a cabo una inspección completa.

F.4 Inspección para la calidad

Para cada área de muestreo se elabora en campo un informe como el que se muestra más adelante, y los resultados se resumen para generar el informe que se muestra en la figura F.1. En el área de muestreo había, por ejemplo, 28 edificios de tipo no residencial; uno faltaba en el conjunto de datos (omisión) y existían 11 objetos geográficos que no deberían haber sido capturados según la especificación de producto.

Nombre del conjunto de datos: L213101C Unidad de muestreo: coordenadas (Norte (m), Este (m)): 6741000 2509000, largo (m), ancho (m): 1000 1000				
Objeto geográfico	Número de ítems	Compleitud		Exactitud temática Corrección de la clasificación Número de errores
		Omisión Número de errores	Comisión Número de errores	
Carretera de II orden	4			
Carretera de III orden	6			
Edificio, residencial, (una o dos plantas)	10			
Edificio, otro edificio, (una o dos plantas)	28	1	11	

Figura F.1 — Ejemplo de anotaciones de una comprobación en campo para una porción de un área

F.5 Determinación de los resultados de la calidad de datos y de la conformidad

Para cada comprobación de calidad realizada se produce un informe generado por computadora. El informe completo de las pruebas de calidad incluye más de 65 objetos geográficos, algunos con uno o más atributos. En la figura F.2 se muestra un ejemplo de un informe de evaluación de la completitud y la exactitud temática de los conjuntos de datos topográficos.

En la figura F.2 hay 16 bases de datos (hojas del mapa 1:10 000) que han sido elegidas para la muestra. Se ha empleado un algoritmo por computadora para la selección de los cuadrados de 1 km x 1 km de aquellas bases de datos. En la figura F.2 se muestra un ejemplo de una unidad de muestreo. En campo se emplea una copia impresa de esta unidad de muestreo junto con la figura F.1, y se examina la completitud y exactitud temática de cada uno de los ítems. Los resultados son resumidos en la figura F.2.

Por ejemplo, el objeto geográfico “Carretera” puede tener cuatro errores por cada cien unidades para la completitud y cuatro errores por cada cien unidades en la clasificación. En las bases de datos había 4712 carreteras diferentes (una carretera es una línea entre nodos). En la muestra había 184 ítems. La Norma NC ISO 2859-1 requiere 80 ítems para este tamaño de lote y nivel de inspección, luego se cumplen los requisitos mínimos. El valor de aceptación para este tamaño de muestra es 10, luego puede haber 10 errores en completitud o en clasificación. En la muestra solamente había dos errores, considerando los de completitud y los de clasificación, por lo que se supera la prueba (se acepta el lote). Como referencia, se ha incluido también en la figura F.2 el valor de aceptación para un NCA = 1 error por cada cien unidades. En esta ocasión también este criterio habría sido satisfecho.

NOMBRE DE LA PRUEBA: 213101_04 Fecha: 09.09.1996 15:15:56 Área: L213101A L213101B L213101C L213101D L213102A L213102B L213102C L213102D L213103A L213103B L213103C L213103D L213104A L213104B L213104C L213104D (El área está definida por las hojas del mapa)										
Tipo de objeto geográfico	Tipo de dato del atributo	Tamaño del lote	Tamaño de la muestra	Tamaño en el nivel 1 de inspección	Nivel de inspección	NCA	Valor de aceptación	Valor de aceptación para NCA 1	Compleción (omisión o comisión) Número de errores	Exactitud temática corrección de la clasificación Número de errores
Carretera		4712	184	80	I	4	10	7	2	—
	Orden (I, II, III)					4	10	7	2	
	Número de la carretera					1	0	0		
	Número de sección de carretera					1	0	0		
	Categoría vertical					1	0	0		
	Tráfico en un solo sentido					0	184	80		
	Tipo de pavimento					0	184	80		
	Estado					0	184	80		
	Altura libre					0	184	0		
Edificio		6447	222	80	I	4	14	7	4	4
	Uso					0	222	80	2	
	Número de plantas					0	222	80		

NOTA - Para mayor claridad, algunos objetos geográficos no se han mostrado en este ejemplo.

Figura. F.2 — Completitud y exactitud temática de la base de datos topográficos

F.6 Presentación de los resultados de calidad

Las figuras F.3 y F.4 proporcionan ejemplos de cómo informar sobre los resultados de calidad. En la figura F.3 se han presentado los resultados de calidad según los metadatos descritos en la Norma ISO 19115. Para presentar información más detallada de la calidad se emplea un Informe de Evaluación de la Calidad (figura F.4). Entre paréntesis hay una explicación de los

códigos usados que se pueden encontrar en la Norma ISO 19115, pero que no forman parte del informe.

DataQuality		
dqScope		
scplvl	012 (tipo de objeto geográfico)	003 (clase de atributo de objeto geográfico)
scpExt	Extensión	
exDesc	Área del lote	
geoEle		
exTypeCode	1 (inclusión)	
BoundPoly		
polygon	6740000,2500000,6770000,2500000,6770000,2510000,6760000,2510000,6750000,2510000,6750000,2520000,6740000,2520000,6740000,2500000	
dqReport		
eleTypCode	001 (Compleitud)	005 (exactitud temática)
subEleCode	001 (adicional) comisión y omisión	002 (corrección de clasificación)
addSubEle		
addName	comisión y omisión	
addDesc	Comisión y omisión del conjunto de datos	
dqResult		
measName	núm. de exceso o ítems que falta	número de errores
dateTime	1996-09-09	1996-09-09
measResult		
Result		
ConResult		
conSpec		
resTitle	Modelo de Calidad del TDB	
resRefDate	1996	
conExpl	conformidad con la especificación del producto	
Pass	1 (cumple)	1 (cumple)
QuanResult		
quanValDom	número	número
quanRes	2	2

Figura F.3 — Informe de metadatos de acuerdo con la Norma ISO 19115

reportidentification	Informe de Evaluación de calidad para la Base de Datos Topográficos	
Reportscope	Ámbito definido en metadatos (ver: dqScope)	
compQuantDesc		
dataQualityMeasure		
mathDesc	Número de ítems faltantes o en exceso	Número de errores
compMeasValue	2	2
valType	número	número
realibilityValue	99	99
realibilityValueUnit	por ciento	por ciento
conformRealibility		
conformRelValues	NCA=4	NCA=4
conformRelDom	número	número
ReferenceDoc	Manual de calidad de datos topográficos, Modelo de calidad de datos topográficos, Instrucciones para la compilación de datos topográficos, Catálogo de objetos geográficos y definiciones.	
dqeMethodTypeInfo		
dqeMethodType	1 (directo externo)	
dqeSamplingApplied	1 (muestreo)	
dqeMethodInfo		
dqeTheory	Consulte la ISO 2859 y el Modelo de calidad de datos topográficos	
dqeProcAlgorithm	El programa siguiente se utiliza para pruebas: procedimientos de comando MLAATU.EXE: LAADUNTARKISTUS.COM y parámetros se definen en los archivos: P99.p99,P97.p97, p98.p98.	
dqeParamInfo		
dqeParamDefinition	Nivel de calidad aceptable (consulte ISO 2859)	Nivel de calidad aceptable (consulte ISO 2859)
dqeParamValues	4	0 (no definido)
dqeParamDomain	NCA numérico	NCA numérico
dqeParamInfo		
dqeParamDefinition	Tamaño del lote	Tamaño del lote
dqeParamValues	4712	6447
dqeParamDomain	número	número
dqeParamInfo		
dqeParamDefinition	Tamaño de la muestra	Tamaño de la muestra
dqeParamValues	184	222
dqeParamDomain	número	número
dqeParamInfo		
dqeParamDefinition	Tamaño de la muestra requerido en el nivel de inspección 1	
dqeParamValues	80	
dqeParamDomain	número	
dqeParamInfo		
dqeParamDefinition	Nivel de inspección	
dqeParamValues	I	
dqeParamDomain	clase	
dqeParamInfo		
dqeParamDefinition	Valor de aceptación	Valor de aceptación
dqeParamValues	10	222
dqeParamDomain	número	número
dqeSampleMethod		
dqeSamplingScheme	Del lote, se muestrea un área de tantos cuadrados de 1 x 1 km de manera que el número de caminos en la muestra es al menos el mismo requerido para NCA=4.	
dqeltemDescription	El ítem es una línea de camino entre nodos	El ítem es un edificio.
dqeLotDescription	Un lote es un grupo de bases de datos (hoja de la carta 1:10 000) que se inspecciona. El tamaño del lote es el número de objetos geográficos en el lote.	
dqeSamplingRatio	Por regla general, se recomienda como un tamaño práctico para lote un área que comprende 4 hojas de la carta (16 bases de datos) con 6 a 10 cuadrados de 1 x 1 km.	

Figura F.4. — Informe de Evaluación de la Calidad de acuerdo a la Norma NC ISO 19114:2010, Anexo I

ANEXO G

(Informativo)

EJEMPLO DE MEDIDA E INFORME DE LA COMPLETITUD Y LA EXACTITUD TEMÁTICA

G.1 Introducción

Este anexo proporciona un ejemplo de la medida y presentación de informe sobre la exactitud temática y la completitud. Los objetivos de este ejemplo son los siguientes:

- demostrar cómo se pueden aplicar los procedimientos de evaluación de la calidad para medir y realizar informes sobre los resultados cuantitativos de la calidad de datos;
- dar un ejemplo de la medida y la realización de un informe sobre exactitud temática y completitud;
- demostrar la utilidad de las matrices de confusión como herramientas para la evaluación de la calidad de datos.

El ejemplo incluye detalles acerca de cómo presentar los resultados de la calidad de datos en forma de metadatos y también como informe de evaluación de la calidad

G.2 Descripción del conjunto de datos

El “mundo real” está representado en la figura G.2. La especificación de producto, dada en la Figura G.1, describe el universo de discurso. La especificación define aquellos objetos geográficos, atributos y relaciones que son considerados importantes y que deberían estar en el conjunto de datos.

Con el propósito de demostrar cómo puede haber sido producido el conjunto de datos, el universo de discurso (esto es, el conjunto de datos ideal que se ajusta a la especificación de producto) se representa gráficamente en la figura G.3. En todas las figuras,

- el dígito o la letra que representa el dominio, y que aparecen bajo el símbolo del árbol, indican la altura del árbol en metros,
- el dígito que aparece dentro del símbolo de la casa indica su número de ocupantes, y
- el nombre de la familia ocupante aparece al lado del símbolo de la casa.

La relación entre las tres figuras es la siguiente:

- La figura G.2 representa el “mundo real”, el cual generalmente contiene más objetos geográficos que los que habrá en el conjunto de datos.
- La figura G.3 representa el “universo de discurso” según la especificación de producto; es aquella parte del ‘mundo real’ que va a ser incluida en el conjunto de datos, si éste está completo y se ha producido con exactitud.
- La figura G.4 representa el conjunto de datos tal y como se ha producido.

Cada ítem subrayado es un tipo de objeto geográfico. Debajo de cada tipo de objeto geográfico se enumeran, por su nombre, sus atributos, que pueden ser varios o ninguno. Cada nombre de atributo es seguido por dos puntos y el tipo de valor, cadena de caracteres o entero. La indicación del dominio, que aparece tras cada tipo de valor y entre corchetes, es opcional.

Tipos de objetos geográficos

Edificio industrial

Casa

Nombre de la familia: cadena de caracteres

Número de ocupantes: entero

Árbol

Clase de altura: cadena de caracteres {A: de 1 a 3 m, B: de 3 a 5 m, C: de 5 a 10 m, D: más de 10m}

Camino

Carretera

Estado: cadena de caracteres {asfaltado, sin asfaltar}

Reglas de la especificación de producto

- los árboles con una altura menor de 1 m no deben ser representados
- el atributo “estado” de una carretera puede no presentar valor alguno (“valor indeterminado”)
- los atributos “nombre” y “número de ocupantes” de una casa pueden no presentar valor alguno (“valor indeterminado”)

Figura G.1 — Especificación de producto

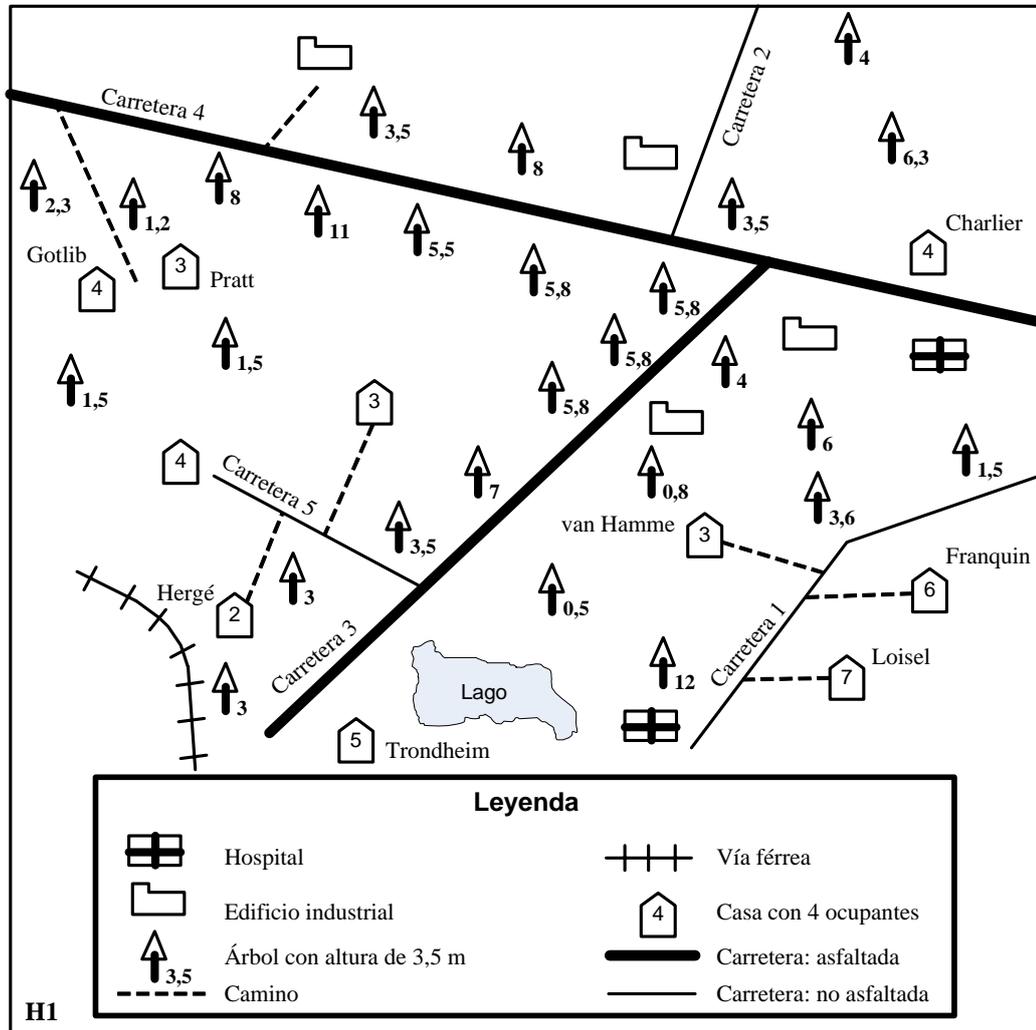


Figura G.2 — Representación gráfica del 'mundo real'

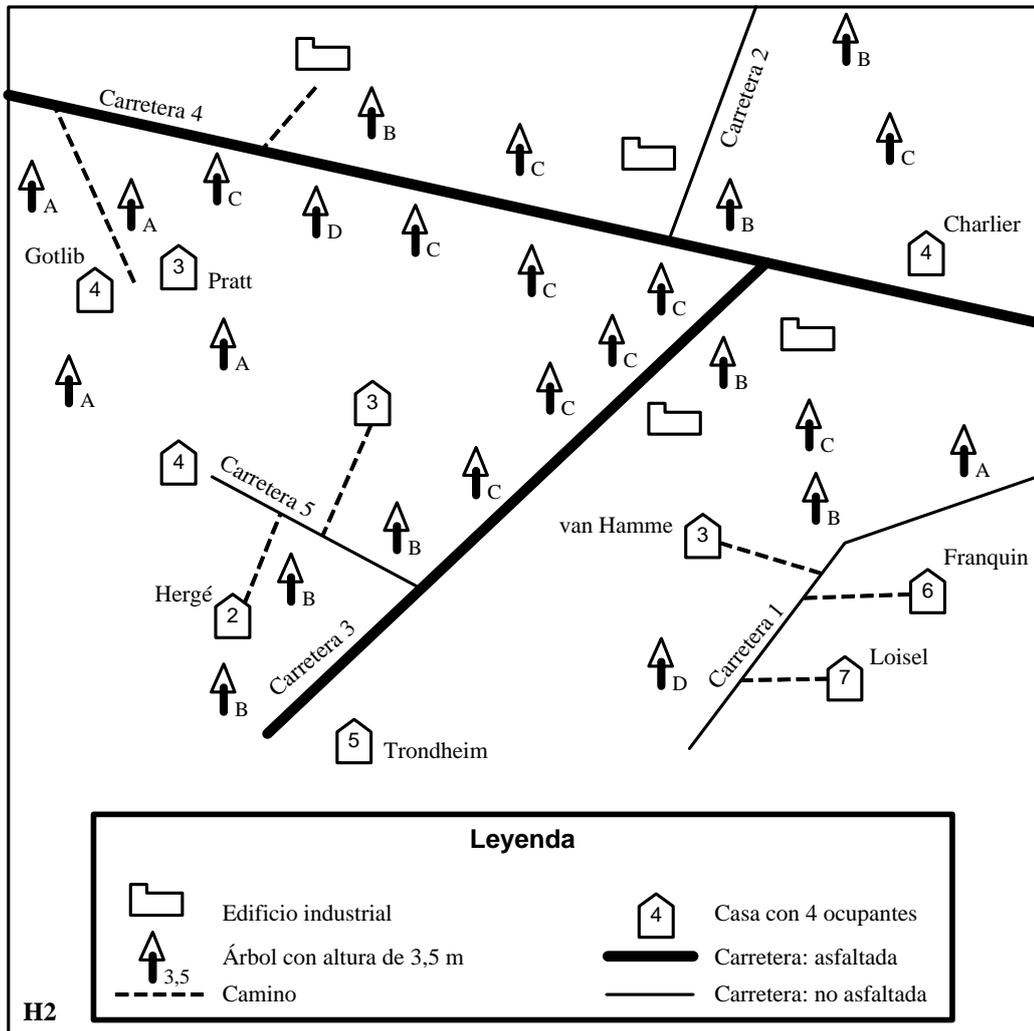


Figura G.3 — Representación gráfica del universo de discurso

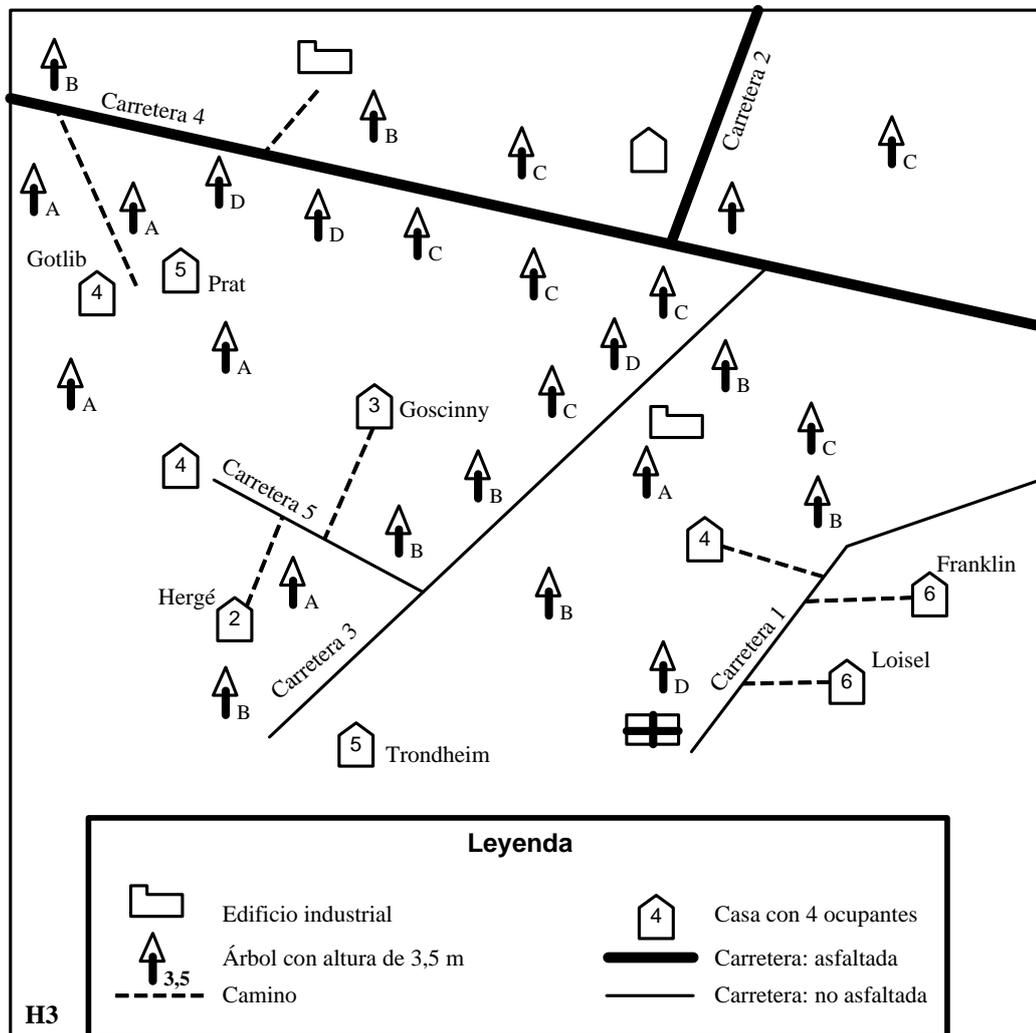


Figura. G.4 — Representación gráfica del conjunto de datos

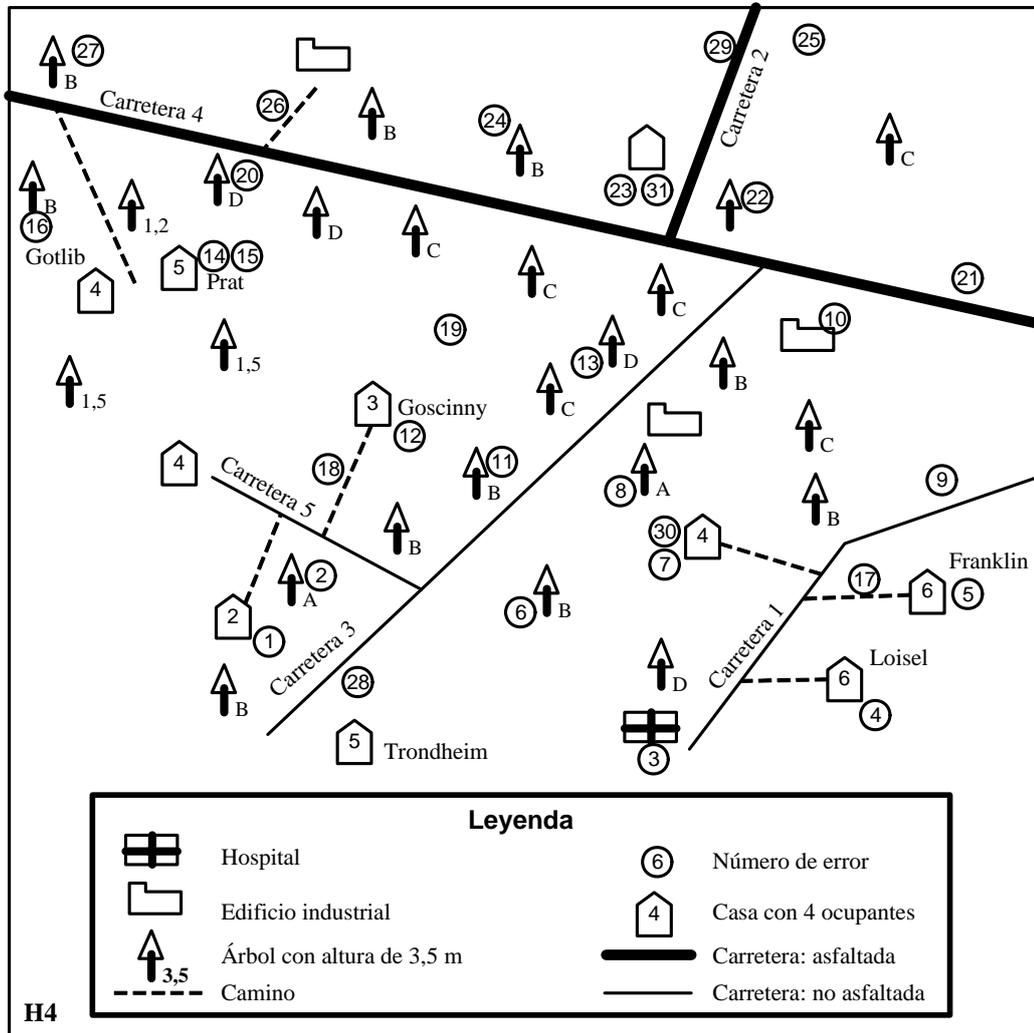


Figura. G.5 — Representación gráfica de la localización de los errores del conjunto de datos

G.3 Evaluación de la calidad de datos

G.3.1 Identificación de los errores

Se puede realizar un listado de los errores del conjunto de datos de ejemplo mediante la comparación del conjunto de datos, representado por la figura G.4, con el universo de discurso, representado por la figura G.3. A continuación se presenta un listado de los errores detectados cada uno con un número de error asignado como referencia.

- a) Errores de omisión y comisión al registrar los árboles. Hay tres árboles (nº 6, nº 8, nº 27) en exceso y dos que faltan (nº 9, nº 25).
- b) Errores de omisión y comisión al registrar los caminos. Hay un camino (nº 18) que falta y uno que está en exceso (nº 19).
- c) Hay una casa en el lugar de un edificio industrial (nº 23).
- d) Dos caminos están mal codificados como carreteras (nº 17, nº 26).
- e) Falta una casa (nº 21).
- f) Errores de atributo en las carreteras. En dos carreteras el atributo “estado” es erróneo (nº 29, nº 28).
- g) Se ha representado un hospital en el conjunto de datos (nº 3).
- h) Se han representado dos árboles de altura menor a 1 m en el conjunto de datos (nº 6, nº 8).
- i) Faltan códigos de atributo de la clase de altura de los árboles. En un árbol falta el código mientras que en el universo de discurso aparece clasificado como B (nº 22).
- j) Error de clasificación en el atributo altura de los árboles. Hay errores en la clase de altura asignada a seis árboles (nº 2, nº 11, nº 13, nº 16, nº 20, nº 24).
- k) Errores en el atributo “nombre de la familia” de las casas. Las casas denominadas en el universo de discurso como “van Hamme” (nº 7) y “Hergé” (nº 1) no tienen nombre en el conjunto de datos. La casa denominada como “Gosciny” en el conjunto de datos (nº 12) no tiene nombre en el universo de discurso.
- l) Errores en el atributo “nombre de la familia” de las casas. Las casas denominadas como “Franquin” (nº 5) y “Pratt” (nº 15) en el universo de discurso, se han denominado como “Franklin” y “Prat” respectivamente en el conjunto de datos.
- m) Errores en el atributo del número de ocupantes de la casa. Falta el atributo en una casa (nº 31) y en tres casas está equivocado (nº 4, nº 14, nº 30).
- n) Error de omisión en edificios industriales. Falta un edificio industrial (nº 10).

NOTA - La clasificación de los errores como omisión/comisión, completitud o exactitud temática es subjetiva. Por ejemplo, un error de clasificación de una casa como un edificio industrial podría ser considerado tanto como un error de omisión de uno como un error de comisión del otro.

G.3.2 Completitud

La Norma NC ISO 19113 define la completitud como la presencia o ausencia de objetos geográficos, de sus atributos y de sus relaciones. En este ejemplo, la completitud se clasifica por la clase de objeto geográfico. Los tipos de medidas evaluados son la comisión y la omisión. La tabla G.1 representa una forma de clasificar la completitud.

Tabla G.1 — Completitud según la clase de objeto geográfico

Clase de objeto geográfico	Número de casos en el universo de discurso	Cuenta de comisiones	Porcentaje de comisiones ^a	Cuenta de omisiones	Porcentaje de omisiones ^b
Camino	7	0	0	2	29
Carretera	5	2	40	0	0
Árbol	25	3	12	2	7
Edificio industrial	4	0	0	2	50
Casa	10	1	10	1	10
Hospital	0	1	100	0	0

^a Porcentaje de comisiones = número de ítems incluidos / número de ítems en el universo de discurso * 100
^b Porcentaje de omisiones = número de ítems omitidos / número de ítems en el universo de discurso * 100

G.3.3 Exactitud temática

La Norma NC ISO 19113 define la exactitud temática como la exactitud de atributos cuantitativos y la corrección de atributos no cuantitativos y de las clasificaciones de los objetos geográficos y sus relaciones. Una forma de representar los errores asociados con la exactitud temática es mediante el empleo de la “matriz de confusión”.

NOTA 1 - Una matriz de confusión es una matriz cuadrada donde el elemento i, j corresponde a la cantidad clasificada como perteneciente a la clase j cuando realmente pertenece a la clase i .

La tabla G.2 es una matriz de confusión en la que se muestran los errores por clase de objeto geográfico. Ésta explica en qué medida se han clasificado correctamente los diferentes casos del conjunto de datos. Los porcentajes deberían referirse siempre a la población del conjunto de datos.

NOTA 2 - En las matrices, el número tras el nombre del tipo de objeto geográfico indica el número de ocurrencias y el valor en la celda es el porcentaje de error de clasificación.

Tabla G.2 — Matriz de confusión de los objetos geográficos

Universo de Discurso	Conjunto de datos					
	Ruta 5	Camino 7	Árbol 26	Edificio industrial 2	Casa 10	Ninguno (objetos geográficos omitidos)
Ruta 7	4/5 = 80%	2/7 = 29%	0%	0%	0%	1/5 =20%
Camino 5	0%	5 / 7 = 71%	0%	0%	0%	0%
Árbol 25	0%	0%	23/26 = 88%	0%	0%	2/5 =40%
Edificio industrial 4	0%	0%	0%	2/2 =100%	1/10 =10%	1/5 =20%
Casa 10	0%	0%	0%	0%	9/10 =90%	1/5 =20%
Ninguno 3 (objetos geográficos excedentes)	0%	0%	3/26 = 11%	0%	0%	
Suma de población conjunto de datos	5/5 =100%	7/7 =100%	26/26 =100%	2/2 =100%	10/10 =100%	

En las tablas G.3 y G.4 sólo se han tenido en cuenta los objetos geográficos que tienen homólogo el mismo tipo de objeto geográfico (“clase”).

El atributo altura de los árboles se muestra en la tabla G.3.

Tabla G.3 — Matriz de confusión de los atributos de objeto geográfico — altura de árbol

Universo de discurso	Conjunto de datos				
	clase A 1 a 3 m 5	clase B 3 a 5 m 5	clase C 5 a 10 m 10	clase D > 10 m 4	indeterminado (omisión de valores) 4
clase A 5	3 / 5 = 60%	1 / 10 = 10%	0%	0%	1 / 4 = 25%
clase B 8	1 / 5 = 20%	5 / 10 = 50%	0%	0%	2 / 4 = 50%
clase C 10	0%	2 / 10 = 20%	5 / 5 = 100%	2 / 4 = 50%	1 / 4 = 25%
clase D 2	0%	0%	0%	2 / 4 = 50%	0%
indeterminado 3 (comisión)	1 / 5 = 20%	2 / 10 = 20%	0%	0%	0%
Suma de población conjunto de datos	5 / 5 = 100%	10/10 = 100%	5 / 5 = 100%	4 / 4 = 100%	4 / 4 = 100%

El atributo de estado de las carreteras se muestra en la tabla G.4.

Tabla G.4 — Matriz de confusión de los atributos de objeto geográfico — Estado de las carreteras

Universo de discurso	Conjunto de datos	
	Asfaltadas 2	No asfaltadas 3
Asfaltadas 2	1 / 2 = 50%	1 / 3 = 33%
No asfaltadas 3	1 / 2 = 50%	2 / 3 = 67%

El atributo “número de ocupantes” de las casas es un ejemplo de exactitud de un atributo cuantitativo de un objeto geográfico definido por un valor. A continuación se demuestra una forma de medir los elementos de calidad de exactitud temática y completitud, y de expresar los resultados de las mediciones en forma de texto, razones comisión/omisión y estadísticos de error.

- 1/9 de las casas no presentan valor del número de ocupantes;
- sesgo $2/8 = -0,25$ ocupantes;
- EMC 0,87 ocupantes;
- tamaño de la muestra: 8.

G.4 Informe de los resultados de calidad

G.4.1 Ejemplo de un error de comisión

En los apartados G.4.2 y G.4.3 se presenta un ejemplo sobre cómo realizar un informe de los resultados de calidad para un tipo de error, los errores de comisión para el tipo de objeto geográfico “camino”. En primer lugar los resultados de calidad han sido presentados como

metadatos. Después, se ha realizado un informe de evaluación de la calidad para presentar información más detallada sobre la calidad.

G.4.2 Informe como metadatos

La figura G.6 es un ejemplo sobre cómo realizar un informe de los resultados de la calidad en forma de metadatos, tal como se describe en la Norma ISO 19115. Entre paréntesis se ofrece la explicación de los códigos empleados en la Norma ISO 19115, pero estos no forman parte del informe.

DataQuality		
	DQ_Scope	
	sepLv1	012 (<i>tipo de objeto geográfico</i>)
	Extent	
	exDesc	Extensión del conjunto de datos
	geoEle	
	exTypeCode	1 (<i>inclusión</i>)
	GeoBndBox	
	westBL	+005.0134
	eastBL	+005.0228
	southBL	+22.956
	nortBL	+23.003
	DQ_Completeness	
	DQ_Omission	
	DQ_Measure	
	nameOfMeasure	conteo
	domainOfMeasure	número
	description	Número de árboles que faltan
	EvaluationProcedure	comparación de la cuenta de árboles en la fuente y en el conjunto de datos
	dateTime	2000-09-14
	DQ_Result	
	DQ_QuantitativeResult	
	valueDomain	{0 ... n}
	result	2

Figura G.6 — Informe como metadatos según la Norma ISO 19115

G.4.3 Informe sobre la evaluación de la calidad La figura G.7 es un ejemplo sobre cómo presentar los resultados de calidad a modo de informe.

	addQualityReport	
	reportIdentification	Informe de Calidad de Ejemplo en este anexo
	reportScope	Conjunto de datos
	compQuantDesc	
	dataQualMeasure	
	mathDesc	Número de ítems del conjunto de datos dividido por número de ítems del universo de discurso multiplicado por 100.
	compMeasValue	proporción
	valType	real
	realibilityValue	100
	realibilityValueUnits	
	conformConfidence	
	conformConfValue	
	conformConfValDesc	
	referenceDoc	
	dqeMethodTypeInfo	
	dqeMethodType	2 (interno directo)
	dqeSamplingApplies	3 (no aplicable)
	dqeMethodInfo	
	dqeAssumptions	
	dqeProcAlgorithm	Compárese la cuenta visual de árboles en la fuente con el conjunto de datos.
	dqeParamInfo	
	dqeParamDefinition	
	dqeParamValues	
	dqeParamDomain	
	dqeFullInspectMetho	
	dqeFullInspectType	Cuenta de árboles
	dqeItemDesc	Árboles por especificación del producto
	referenceDoc	
	dqeSampleMethod	
	dqeSamplingScheme	
	dqeItemDescription	
	dqeLotDescription	
	dqeSamplingRation	
	dqeDeductiveSource	
	dqeDeductRefDocs	
	referenceDoc	
	aggSourceValues	
	aggResult	
	aggValueDomain	
	aggMeasureValue	
	aggErrorStat	
	aggQEPreport	
	qepOtherDesc	

Figura G.7 — Informe de Evaluación de la Calidad de acuerdo a la Norm aNC ISO 19114:2010, Anexo I

ANEXO H (Informativo)

EJEMPLO DE UN RESULTADO AGREGADO DE CALIDAD DE DATOS

H.1 Introducción

La información de este ejemplo está basada en técnicas empleadas en la industria privada en Europa, Norteamérica y Asia. El objetivo del ejemplo descrito es ilustrar las técnicas de medición y agregación de la exactitud temática, completitud y exactitud posicional en un conjunto de datos de carreteras.

Este ejemplo sólo trata de presentar un resultado agregado de la calidad de datos. No se realiza ninguna comparación con un nivel de conformidad de la calidad.

H.2 Descripción del conjunto de datos

H.2.1 Representación del mundo real

El mundo real está representado por la figura H.1, donde también se representa un lote extraído del conjunto completo de datos sobre carreteras. El área rectangular sombreada, casilla B-2, representa la unidad de muestreo seleccionada aleatoriamente para ser evaluada.

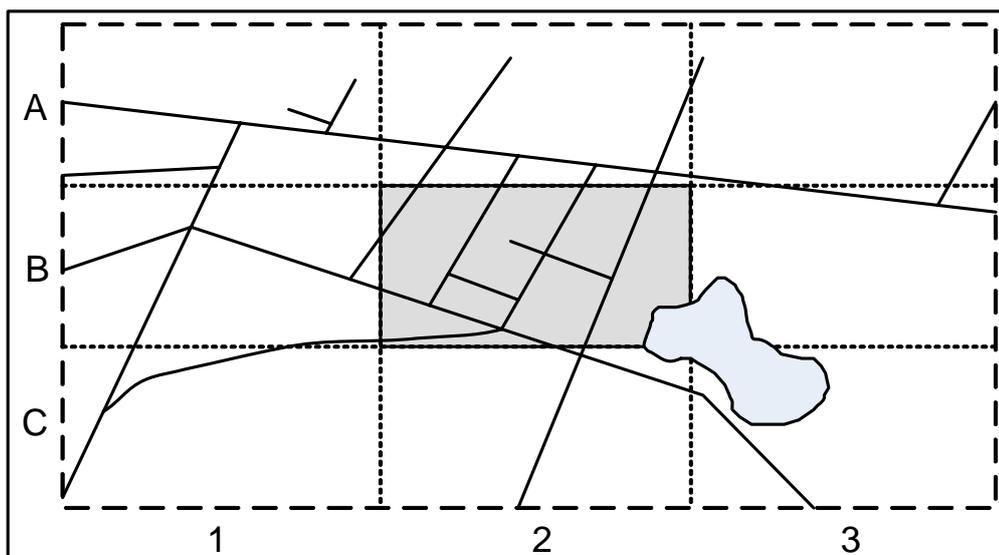


Figura H.1 — Lote seleccionado aleatoriamente de la base de datos completa y unidad de muestreo seleccionada aleatoriamente (rectángulo sombreado)

H.2.2 Especificación de producto

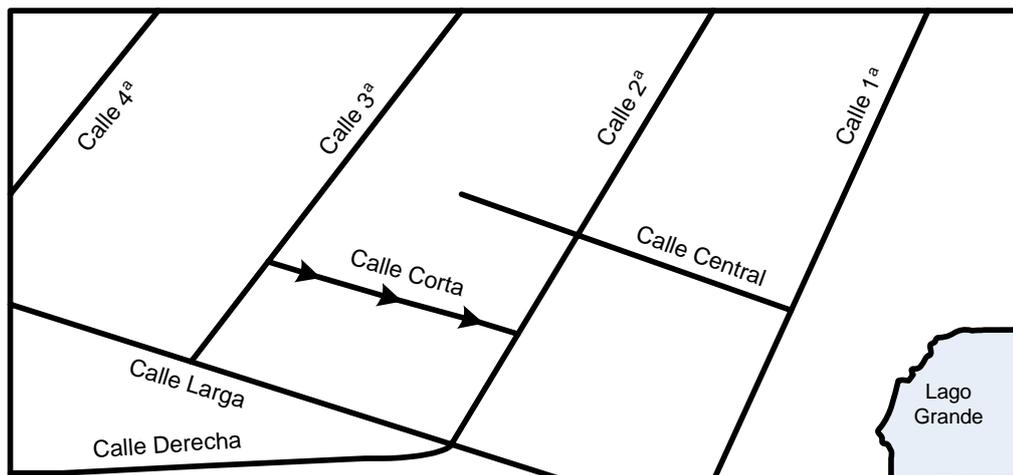
Aunque abreviada para el propósito de este ejemplo, la especificación de producto que define el universo de discurso se muestra en la figura H.2. La especificación describe aquellas reglas que son consideradas importantes para el producto.

Reglas para la especificación de producto
— Todas las vías deberían estar incluidas.
— Todas las vías deberían poseer su nombre.
— La dirección de la circulación de todas las vías de sentido único debe estar indicada.
— Todos los objetos geográficos hidrográficos deberían estar incluidos.

Figura H.2 — Ejemplo de especificación de producto

H.3 Universo de discurso

El universo de discurso se representa en la figura H.3. Para este ejemplo, proporciona una referencia gráfica de la realidad respecto a la cual se comparará el contenido del conjunto de datos.



NOTA - Las flechas indican el sentido de la circulación del tráfico; la no existencia de flechas indica que la circulación del tráfico es en ambos sentidos.

Figura H.3 — Representación gráfica del universo de discurso

H.4 Conjunto de datos

El contenido del conjunto de datos está representado en la figura H.4. Las líneas de puntos indican los lugares donde se detectaron errores, es decir, donde el conjunto de datos no se correspondía con la realidad. Se han apuntado varios tipos de errores. La tabla H.1 identifica los errores y sus tipos.

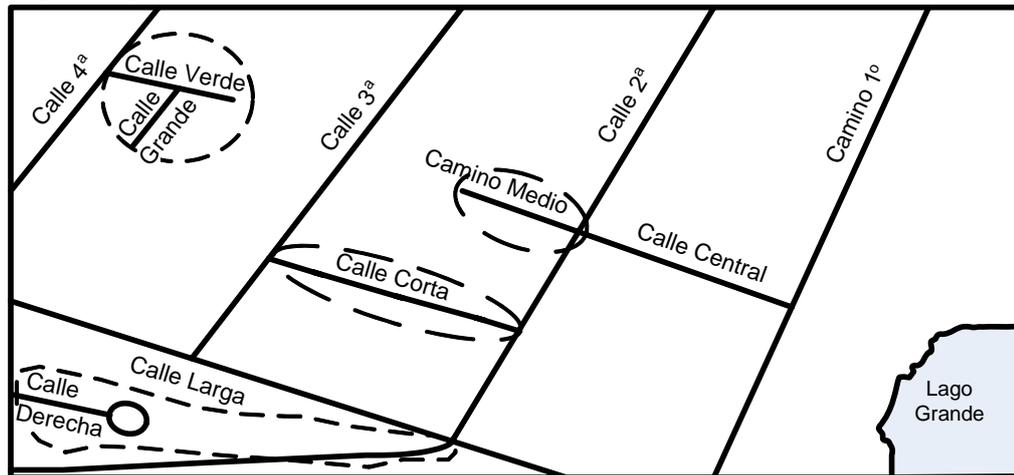


Figura. H.4 — Representación gráfica del contenido del conjunto de datos

Tabla H.1 — Tipos de errores detectados y subelementos de calidad típicos con los que se puede informar sobre la calidad

Tipos de errores detectados	Elemento	Subelemento con el que se informa sobre el error
Calles que no existen, por ejemplo, Calle Verde	Compleitud	Comisión
Nombres de calles incorrectos, por ejemplo, 1 ^{er} Camino	Exactitud temática	Corrección de atributos cualitativos
Ausencia de parte de una calle, por ejemplo, Calle Derecha	Consistencia lógica	Consistencia topológica
Ausencia de atributos, por ejemplo, las flechas de flujo de la Calle Corta	Exactitud temática	Corrección de atributos cualitativos ^a

^a Si las reglas para el conjunto de datos dadas en la especificación de producto requieren que el campo para la dirección del tráfico siempre tenga una entrada, tal como único o doble sentido, el error se mide como una omisión. Sin embargo, si sólo se requiere una entrada, se mide como corrección temática.

H.5 Agregación de los resultados de la evaluación e informe

Se prepara una tabla para mostrar el número de errores encontrados y cómo se han clasificado de acuerdo a los procedimientos típicos empleados en la industria de las bases de datos de carreteras. En este ejemplo se asignan pesos a cada tipo de error. La suma de los pesos es igual al 100%. Se considera que el valor ponderado resultante es la calidad del conjunto de datos. La tabla H.2 muestra un ejemplo de cálculo de un resultado agregado. Se define un ítem

como un segmento de calle limitado con otras calles por puntos de intersección o por los límites de la unidad de muestreo.

Tabla H.2 — Ejemplo de cálculo de un resultado agregado de evaluación de la calidad

Objeto geográfico	Número ítems en el lote	Número de ítems no conformes	Proporción no conforme	Proporción exactitud (definición como relación 1)	Pesos	Valor ponderado (proporción de exactitud * peso)
Segmento de camino	19					
Incorrecto		1				
Omisión		0	4/19	0.79	50%	0.39
Exceso		3				
Nombre de la calle						
Nombre base	19	5	5/19	0.74	15%	0.11
Dirección de la vía	19	1	1/19	0.95	25%	0.23
Hidrografía	1	0	0/1	1.00	10%	0.10
Resultado de calidad de datos agregados (definido como la suma de la proporción exactitud ponderada por 100)						84%

ANEXO I (Normativo)

PRESENTACIÓN DE LA INFORMACIÓN SOBRE LA CALIDAD MEDIANTE UN INFORME DE EVALUACIÓN DE LA CALIDAD

I.1 Introducción

Este anexo describe el contenido de un informe detallado de evaluación cuantitativa de la calidad. El informe de evaluación de la calidad proporciona más detalle sobre los resultados de la calidad y los procedimientos utilizados para calcularlos que los metadatos. La tabla I.1 proporciona un gráfico sobre las relaciones anidadas del contenido del informe.

I.2 Componentes del informe de evaluación de la calidad

Las cabeceras de las columnas y los códigos de la tabla I.1 son los siguientes:

Número de línea: Da una referencia para cada ítem de la tabla y se utiliza en la columna del dominio para mostrar el rango de los componentes de este ítem en la tabla.

Nombre: Nombre del elemento en el informe.

Definición/contenido: Define el ítem o describe su contenido.

Obligación/condición: Ofrece los requisitos para informar acerca del ítem o las condiciones bajo las que se requiere el ítem. Hay tres códigos de obligación:

- **Obligatorio** (Mandatory, **M**): Indica que la entrada es necesaria.
- **Condicional** (Conditional, **C**): Indica que la entrada es necesaria cuando se cumple la condición.
- **Opcional** (Optional, **O**): Indica que la entrada es optativa.

Máxima ocurrencia (máx. ocur): Máximo número de veces que este ítem puede aparecer en un dominio superior. Un número entero indica el número de veces, y N indica que puede aparecer tantas veces como se quiera.

Tipo de dato: Puede ser una sección de un informe, un texto, una entidad o un guión “-” cuando no se aplica.

Dominio: Para cada elemento del informe, el dominio especifica los valores permitidos o el uso de texto libre. La opción de texto libre indica que no hay restricciones en el contenido del campo. Los códigos basados en números enteros deben ser usados para representar valores en dominios restringidos (cerrados).

Tabla I.1 — Componentes de un informe de evaluación de la calidad

N° de línea	Nombre	Definición/contenido	Obligación/condición	Máx. ocur	Tipo de datos	Dominio
1	addQualityReport	Informe de evaluación de la calidad	C/ apartado 9.2	1	Sección de informe	Líneas 2 - 40
2	reportIdentification	Información de identificación del informe	M	1	Cadena de caracteres	Texto libre
3	reportScope	Ámbito del conjunto de datos evaluados en este informe (Norma ISO 19113)	O	1	Cadena de caracteres	MD_MetadataScope <<CodeList>>
4	compQuantDesc	Descripción complementaria de la evaluación cuantitativa, tal como los valores de las medidas de la calidad y sus límites de confianza	M	1	Sección de informe	Líneas 5 - 14
5	dataQualMeasure	Información acerca de la definición y valor de las medidas de la calidad correspondientes a un ámbito de calidad de datos de objeto	M	1	Sección de informe	Líneas 6 - 10
6	mathDesc	Descripción matemática de la medida de calidad	M	1	Cadena de caracteres	Texto libre
7	compMeasValue	Valores de la medida de la calidad aplicada	M	1	Cadena de caracteres	Texto libre
8	valType	Unidad en la que se registra el valor de la medida de la calidad	M	1	Cadena de caracteres	Texto libre
9	reabilityValue	Valores límite de fiabilidad o confianza del valor calculado o estimado de la medida de la calidad	O	1	Cadena de caracteres	Texto libre
10	reabilityValueUnits	Unidades en las cuales se registran los valores de fiabilidad	O	1	Cadena de caracteres	Texto libre
11	conformConfidence	Confianza en la conformidad	O	1	Sección de informe	Líneas 12 - 14
12	conformConfValue	Confianza en el resultado de conformidad NOTA - La confianza en la conformidad puede ser del tipo ALTA, BAJA, NINGUNA, o 95%, etcétera	M	1	Cadena de caracteres	Texto libre
13	conformConfValDesc	Unidad o tipo de valor con los cuales se registra la confianza en la conformidad	M	1	Cadena de caracteres	Unidad de valor o tipo de valor
14	referenceDoc	Información sobre los documentos usados como referencia para el desarrollo y aplicación de los métodos de evaluación de la calidad	O	N	Clase	CI_Citation
15	dqeMethodTypeInfo	Información detallada sobre la aplicación del método de evaluación de la calidad	M	1	Sección del informe	Líneas 16 - 37
16	dqeMethodType	Clase del método de evaluación de la calidad	M	1	Cadena de caracteres	1-directo-externo 2-directo-interno 3-indirecto
17	dqeSamplingApplied	Información sobre la estrategia de muestreo empleada	M	1	Cadena de caracteres	1-muestreo aplicado 2-inspección completa 3-no aplicable

Tabla I.1 (Continuación)

N° de línea	Nombre	Definición/contenido	Obligación/condición	Máx. ocur	Tipo de datos	Dominio
18	dqeMethodInfo	Información sobre el método de evaluación de la calidad	M	1	Sección de informe	Líneas 19 - 37
19	dqeAssumptions	Información sobre suposiciones subyacentes para el desarrollo y aplicación del método de evaluación de la calidad	O	1	Cadena de caracteres	Texto libre
21	dqeProcAlgorithm	Información sobre cómo se procesan los datos para determinar el resultado de la calidad	M	1	Cadena de caracteres	Texto libre (si se ha empleado un algoritmo o comando de computadora específico se debe incluir su nombre)
22	dqeParamInfo	Información sobre los parámetros usados en el método de evaluación de la calidad de datos	O	N	Sección de informe	Líneas 23 - 37
23	dqeParamDefinition	Información sobre la definición del parámetro empleado	M	1	Cadena de caracteres	Texto libre, por ejemplo, valor del peso de cada medida agregada de calidad
24	dqeParamValues	Valor del parámetro usado en el método de evaluación de la calidad	M	1	Cadena de caracteres	Texto libre
25	dqeParamDomain	Unidad en la que se registra el valor del parámetro	M	1	Cadena de caracteres	Texto libre
26	dqeFullInspectionMethod	Información sobre el método de inspección completa	C/ se ha aplicado inspección completa	1	Sección de informe	Líneas 27 - 29
27	dqeFullInspectionType	Información sobre el tipo de inspección completa y descripción del procedimiento	M	1	Cadena de caracteres	Texto libre
28	dqeItemDescription	Información sobre cómo se han definido los ítems	M	1	Cadena de caracteres	Texto libre
29	referenceDoc	Información sobre los documentos usados como referencia para el desarrollo y aplicación del método de evaluación de la calidad	O	N	Clase	CI_Citation
30	dqeSampleMethod	Información sobre el método de muestreo	C/ se ha realizado muestreo	1	Sección de informe	Líneas 31 - 37
31	dqeSamplingScheme	Información sobre el tipo de plan de muestreo y descripción del procedimiento de muestreo	M	1	Cadena de caracteres	Texto libre, por ejemplo, muestreo aleatorio simple: los ítems se muestrean en cada lote
32	dqeItemDescription	Información sobre cómo se definen los ítems	M	1	Cadena de caracteres	Texto libre
33	dqeLotDescription	Información sobre cómo se definen los lotes	C/ si se emplean lotes	1	Cadena de caracteres	Texto libre

Tabla I.1 (continuación)

N° de línea	Nombre	Definición/contenido	Obligación/condición	Máx. ocur	Tipo de datos	Dominio
34	dqeSamplingRatio	Información sobre el número medio de muestras que se extraen para la inspección de cada lote o población	M	1	Cadena de caracteres	Texto libre
35	dqeDeductiveSource	Información sobre qué datos son utilizados como fuentes para el método de evaluación deductiva	C/ se aplica el método deductivo	1	Cadena de caracteres	Texto libre, por ejemplo, linaje y uso del ámbito de la calidad
36	dqeDeductRefDocs	Identificación de los documentos fuente empleados como base para la deducción	M	N	Cadena de caracteres	Texto libre
37	referenceDoc	Información de los documentos usados como referencia para el desarrollo y aplicación del método de evaluación de la calidad	O	N	Clase	CI_Citation
38	aggSourceValues	Información sobre qué componentes del conjunto de datos son empleados y qué medidas de la calidad son agregadas para determinar el valor y conformidad	C/ se calcula un resultado agregado	N	Sección de informe	Líneas 39 - 44
39	aggResult	Descripción del valor como un resultado cuantitativo	M	1	Sección de informe	Líneas 40 - 44
40	aggValueDomain	Unidad con la que se registra el valor cuantitativo	M	1	Cadena de caracteres	Texto libre, por ejemplo, metros, kilómetros
41	aggMeasureValue	Valor de la medida	M	1	Cadena de caracteres	Texto libre
42	aggErrorStat	Tipo de estadístico	M	1	Cadena de caracteres	Texto libre, por ejemplo, EMC
43	dateTime	Fecha y hora en la cual se calculó el valor	O	1	Fecha y hora	Norma ISO 19108
44	aggQEPreport	Una referencia a un informe de evaluación de la calidad	O	1	Clase	CI_Citation
45	qepOtherDesc	Información adicional, incluidos resultados intermedios, que se consideren importantes para estimar los valores de la calidad y para determinar su conformidad	O	N	Cadena de caracteres	Texto libre

ANEXO J (Informativo)

AGREGACIÓN DE RESULTADOS DE LA CALIDAD

J.1 Introducción

La calidad de un conjunto de datos puede ser representada por uno o más resultados agregados de la calidad (Aggregated Data Quality Result, ADQR; o Resultado Agregado de la Calidad de los Datos, RACD). Un RACD combina los resultados de la calidad a partir de evaluaciones de la calidad basadas en diferentes ítems, subelementos y/o ámbitos de la calidad.

Los siguientes apartados son ejemplos de métodos que pueden usarse para producir un RACD. Aunque los ejemplos muestran cálculos con valores Booleanos, éstos no tienen por qué serlo. Un resultado de la calidad de datos puede ser cuantitativo o cualitativo, y puede representarse por un valor numérico o Booleano. Se puede considerar que un conjunto de datos tiene una calidad agregada aceptable a pesar de que uno o más de los resultados individuales de calidad no sean aceptables. En cualquier caso, el significado del resultado agregado debería quedar claro.

Dado que el RACD puede ser difícil de ser comprendido completamente, su significado debería comprenderse antes de extraer conclusiones basadas en él acerca de la calidad de un conjunto de datos.

El capítulo 8 describe los requisitos para informar sobre resultados agregados de la calidad.

J.2 Conforme / no conforme al 100%

A cada resultado de la calidad implicado en el cálculo se le da un valor Booleano v de uno (1) si ha pasado la prueba y de cero (0) si ha fallado. La calidad agregada se determina por la ecuación,

$$\text{RACD} = v_1 \times v_2 \times v_3 \times \dots \times v_n$$

Donde: n es el número de marcos de medida de la calidad bajo consideración.

Si $\text{RACD} = 1$, entonces la calidad global del conjunto de datos se considera completamente aceptable, por lo tanto pasa.

Si $\text{RACD} = 0$, entonces se considera no conforme, por lo tanto falla. Esta técnica no ofrece un resultado que indique la localización o magnitud de la no conformidad.

J.3 Conforme / no conforme ponderado

A cada resultado de la calidad implicado en el cálculo se le asigna un valor Booleano v de uno (1) si ha pasado y de cero (0) si ha fallado. A cada uno de ellos se le asigna un peso w entre 0,0 y 1,0, ambos inclusive, basándose en su importancia para el propósito del producto. La suma de todos los pesos debería ser igual a 1,0. La elección de pesos es una decisión subjetiva realizada por el productor de datos o el usuario. La decisión del productor de los datos debería estar razonada y formar parte del resultado. La calidad agregada se determina mediante la siguiente ecuación

$$\text{RACD} = v_1 \times w_1 + v_2 \times w_2 + v_3 \times w_3 + \dots + v_n \times w_n$$

Donde: n es el número de marcos de medida de la calidad bajo consideración.

Esta técnica proporciona un valor de magnitud que indica la proximidad de un conjunto de datos a su conformidad global. La técnica no proporciona un valor cuantitativo que indique donde ocurre o no la conformidad.

J.4 Subconjunto de resultados suficientes para los propósitos del producto

Esta técnica es una modificación de los métodos de Conforme / no conforme al 100% y ponderado. A partir de los resultados producidos durante la evaluación global de la calidad, se selecciona un subconjunto de resultados implicados en el cálculo. Este subconjunto representa los resultados considerados significativos para el propósito del producto. Esta técnica puede emplearse cuando se han medido más ítems de calidad que los que son necesarios para cubrir la especificación de producto y/o su propósito.

La calidad agregada se determina mediante la aplicación al subconjunto de resultados de las medidas de la calidad de las técnicas de Conforme / no conforme al 100% y ponderado, o cualquier otra técnica de evaluación por agregación.

Cuando se aplica esta técnica, se deberían documentar las identidades de los marcos de medidas de la calidad seleccionados como componentes del subconjunto.

J.5 Valor máximo / mínimo

A cada resultado de calidad se le asigna un valor v basado en su importancia para el propósito del producto. La decisión del productor de los datos debería estar razonada y formar parte del resultado. La calidad agregada se determina por cualquiera de estas dos ecuaciones,

$$\text{RACD} = \text{máx. } (v_i, i = 1 \dots n)$$

o

$$\text{RACD} = \text{mín. } (v_i, i = 1 \dots n)$$

Donde: n es el número de marcos de medida de la calidad.

La técnica proporciona un valor que indica la proximidad de un conjunto de datos a su conformidad global, pero sólo respecto al marco de medidas representado por el máximo o mínimo. La técnica proporciona un valor cuantitativo que indica dónde se produce la conformidad o la no conformidad cuando el marco de medidas seleccionado se expresa junto con el RACD. Sin embargo, este tipo de RACD dice poco acerca de la magnitud de los otros resultados de la calidad.

Bibliografía

- [1] NC ISO 2859 (todas las partes) — Procedimientos de muestreo para la inspección por atributos.
- [2] NC ISO 9001:2008— Sistemas de gestión de la calidad. Requisitos.
- [3] ISO 3951-1 — ²⁾ — Procedimientos de muestreo para la inspección por variables. Parte 1: Especificación de planes de muestreo simple clasificados por límite de calidad aceptable (LCA) para la inspección lote a lote de una característica y LCA individuales.
- [4] ISO 8601:2000 — Ítems de datos y formatos de intercambio. Intercambio de información. Representación de la fecha y de la hora.
- [5] ISO 3534-2: — ¹⁾— Estadística. Vocabulario y símbolos. Parte 2: estadística aplicada.
- [6] ISO 11404:1996 — Tecnología de la información. Programación de lenguajes, sus entornos y sus interfaces en el sistema operativo. Tipos de datos independientes del lenguaje.
- [7] ISO 19108:2002— Información geográfica. Modelo temporal.

1) Por publicar. (Revisión de la Norma ISO 3534-2:1993)

2) Por publicar. (Revisión de la Norma ISO 3951:1989)

ANEXO NACIONAL

(Informativo)

Las normas internacionales que se relacionan a continuación, citadas en esta norma, han sido incorporadas al cuerpo Normativo NC con los códigos siguientes:

Norma Internacional	Norma UNE
ISO 19113	NC-ISO 19113
ISO 19115	NC-ISO 19115 ¹⁾
ISO 9001:2008	NC-ISO 9001:2008

¹⁾ Pendiente de publicación