GUÍA DE PRODUCCIÓN CARTOGRÁFICA

Generación de cartografía topográfica a escala 1:25,000

INSTITUTO GEOGRÁFICO NACIONAL JOSÉ JOAQUÍN HUNGRÍA MORELL
FACILIDAD REGIONAL PARA LA RESILIENCIA DEL CARIBE
GFDRR
GRUPO BANCO MUNDIAL
GOBIERNO DE ESPAÑA
MINISTERIO DE TRANSPORTES, MOVILIDAD Y AGENDA URBANA
INSTITUTO GEOGRÁFICO NACIONAL

Prólogo

Esta Guía de Producción Cartográfica ha sido generada por personal técnico de la Subdirección General de Cartografía y Observación del Territorio del Instituto Geográfico Nacional de España (IGN), en el marco de colaboración establecido por el Grupo del Banco Mundial, para la "Asistencia técnica al desarrollo de la Infraestructura de Datos Espaciales de la República Dominicana", y en concreto, para el apoyo técnico al Instituto Geográfico Nacional "José Joaquín Hungría Morell" (IGN -JJHM) en esta materia.

Se trata de un extenso documento elaborado en base a la experiencia del IGN de España en cuanto a la producción de cartografía topográfica, adquirida durante más un siglo en el ejercicio de sus funciones.

En el marco de este apoyo técnico, los contenidos aportados en esta guía han sido complementados por una capacitación de 100 horas, donde los conceptos explicados se han puesto en práctica sobre ejemplos de datos geográficos, por medio de la aplicación del software al que se hace referencia en el contenido de la guía.

INSTITUTO GEOGRÁFICO NACIONAL DE ESPAÑA

Listado de autores

Alicia González
Gonzalo Moreno
Gema Martín-Asín
Lorenzo Camón
Francisco Moreno
Valeriano Martín
Julián Taboada
Cristina Calvo
Pablo de la Presa
Felisa Quesada Bustos
Alfonso Boluda
Ana Maldonado
Adolfo Pérez
Santiago Prieto
Alfonso de Tomás
Daniel J. Rolanía
Francisco Mariño
Antonio Moreno

Listado de tutores de la capacitación

Además de los autores mencionados:

Mª Paz Navas
Pablo Roces
Verónica Martínez

INSTITUTO GEOGRÁFICO NACIONAL JOSÉ JOAQUÍN HUNGRÍA MORELL REPÚBLICA DOMINICANA

Listado de revisores

Wanda Lisselote Binet
Dominique Feliz
Yoenny Urbáez
Saulimar Rodríguez

Listado de alumnos de la capacitación

Además de los revisores mencionados:

Rhaymar Matos
Daily Gómez
Juan Rafael Rijo
Julio Cesar Reyes
Saderis Carmona
Nólan Durán
Jean Carlos Ramírez

ÍNDICE

CAPÍTULO 1: Flujo general de producción cartográfica……..**1**

CAPÍTULO 2: Adquisición de información geométrica……..**14**

CAPÍTULO 3: Integración de información semántica……..**114**

CAPÍTULO 4: Calidad de la información geográfica……..**162**

CAPÍTULO 5: Producción del Mapa I. Generación de los datos cartográficos….**217**

CAPÍTULO 6: Producción del Mapa II. Visualización final del producto……**.298**

CAPÍTULO 1: FLUJO GENERAL DE PRODUCCIÓN CARTOGRÁFICA

Generación de cartografía topográfica a escala 1:25,000

Descripción breve

Capítulo introductorio donde se expone el flujo general de producción, considerando como punto de partida la existencia de la información auxiliar con calidad posicional adecuada y como punto final la generación de cartografía 1:25,000

Índice de contenidos

1.1. Introducción 1
1.2. Flujos habituales en la actualidad 2
1.3. Resumen de los contenidos de la guía 6
Capítulo 2: Adquisición de información geométrica 7
Capítulo 3: Integración de información semántica 8
Capítulo 4: Calidad de la información geográfica 9
Capítulo 5: Producción del Mapa I. Generación de los datos cartográficos 10
Capítulo 6: Producción del Mapa II. Visualización final del producto 11

Índice de ilustraciones

Ilustración 1- 1: Flujo habitual para la elaboración de cartografía en la actualidad 3

Índice de tablas

No se encuentran elementos de tabla de ilustraciones.

1.1. Introducción

No cabe duda de que la tecnología, la forma de llevar a cabo los proyectos y el mundo en general han sufrido importantísimos avances y revoluciones durante el último siglo y medio. La elaboración de mapas no es ajena a todos estos cambios. Así, si se echa la vista atrás entre 100 y 150 años, podemos observar varias evoluciones tecnológicas que han revolucionado la forma de elaborar la cartografía de un territorio.

En aquellos tiempos era fundamental establecer, como primer paso, las redes geodésicas y las redes topográficas de densificación en las que apoyarse para poder obtener coordenadas de los elementos que definen el territorio con una exactitud aceptable, dentro de las metodologías e instrumentación empleadas. Estos levantamientos topográficos tenían que realizarse en campo, a través de la toma de diferentes medidas, del dibujo de los croquis de los elementos presentes en el terreno y de la recopilación de la toponimia más relevante del lugar, obtenida generalmente a partir de entablar conversaciones con los habitantes de la zona. Con todo ello se configuraban las minutas de campo con las que posteriormente se elaboraban los mapas.

Se puede decir que la gran revolución del siglo XX en la elaboración de cartografía es la fotogrametría. La utilización de vuelos fotogramétricos permite la obtención de información del territorio de forma masiva, reduciendo enormemente la necesidad de realizar levantamientos en campo. Mediante la restitución fotogramétrica se pueden conseguir coordenadas tridimensionales de cualquier punto visible en dos fotogramas, lo que la convierte en la principal herramienta para conseguir información topográfica del territorio. Se trata de una disciplina que sigue en constante evolución, no solo desde la perspectiva de la mejora de los sensores de captura, sino también en relación a las diferentes plataformas que pueden ser empleadas para la colocación de estos sensores. Así, actualmente, además de los clásicos aviones utilizados en la fotogrametría aérea, pueden emplearse otras alternativas como son, por un lado, los vehículos aéreos no tripulados (UAV) o drones, y, por otro, los satélites artificiales, que llevan ya varias décadas proporcionando información a través de diferentes programas de observación de la tierra. La elección de un tipo de plataforma u otro depende de los requisitos técnicos del producto a generar, así como de otros condicionantes generales del proyecto (viabilidad económica, etc.).

La otra gran revolución que tiene lugar principalmente en la segunda mitad del siglo XX es el desarrollo de la geodesia espacial, gracias al empleo de satélites artificiales. Con estas técnicas se pueden calcular distancias entre puntos muy alejados con tasas de error muy pequeñas si las comparamos con las técnicas de la geodesia clásica. Gracias a esta disciplina comenzaron a definirse los sistemas de referencia globales y se puso de manifiesto los errores que contenían los marcos de referencia clásicos o locales. Así, hoy en día disponemos de estas redes geodésicas de gran exactitud definidas por técnicas espaciales y, en muchos casos, de redes activas compuestas por estaciones permanentes GNSS (Global Navigation Satellite System).

En un plano más general, la informatización y el abaratamiento de costes de equipamiento, tanto informático como de sensores de captura de datos, ocurrida en los últimos 50 años ha permitido reducir tiempos y costes de producción, a la vez que ha posibilitado mitigar la concatenación de determinados errores del flujo productivo, mejorando en consecuencia la precisión del resultado. En este sentido cabe destacar la revolución que ha supuesto el desarrollo de la fotogrametría digital, donde el empleo de técnicas de análisis digital de

imágenes permite un alto nivel de automatización de todo el proceso fotogramétrico, y que se estima que incrementará notablemente con la introducción de técnicas de inteligencia artificial.

La última revolución se puede decir que llega con el auge de Internet, al multiplicar las posibilidades tanto de explotación como de producción colaborativa de la información. Así, el espectro de productores y diseminadores de información geográfica no se limita a las tradicionales agencias cartográficas, fundamentalmente de carácter público, sino que un amplio abanico de organizaciones se hace consciente de las ventajas de introducir la componente geográfica en su modelo de negocio. De hecho, suele decirse que aproximadamente el 80% de los datos pueden referirse a alguna localización espacial. Disponer de forma explícita de esta localización multiplica las capacidades de explotación de los datos, por lo que, en las últimas décadas, organizaciones de diversa índole han invertido en la geolocalización de su información de interés. Por su parte, en el plano de la administración, la colaboración entre instituciones se convierte casi de facto en una obligación, ya que los mismos elementos del mundo real son de interés para una multiplicidad de organismos con distinto ámbito competencial, tanto desde el punto de vista del ámbito territorial administrado (todo un estado o ámbitos locales), como desde el punto de vista temático.

Por otro lado, ante el auge en el número y la variedad de productos y servicios que proporcionan información geoespacial, cobra una gran importancia su normalización. Así, surgen proyectos normativos como la familia de normas ISO 19100 de información geográfica.

Por último, se debe destacar que el propósito final de la recopilación de información geográfica no se limita hoy en día a la obtención de mapas a escalas fijas, sino que prima el poder disponer de ella en bases de datos espaciales de cara a su explotación de múltiples formas mediante el uso de Sistemas de Información Geográfica.

1.2. Flujos habituales en la actualidad

En primer lugar, cabe reseñar el hecho de que en la mayoría de los países existen programas de elaboración de cartografía de su territorio a escalas intermedias, generalmente 1:25,000, 1:50,000, 1:100,000 o, incluso, 1:250,000 en algunas zonas. La decisión de emplear una escala u otra suele depender de la viabilidad económica y técnica que se estime para el proyecto, tanto en la elaboración de su primera edición como en la de su mantenimiento. Esta viabilidad está muy condicionada a su vez por la extensión del territorio a cartografiar. Por ejemplo, países con grandes extensiones de territorio poco habitado, como Brasil o Argentina, pueden optar por una escala 1:250,000 para cubrir todo el territorio y emplear una escala más grande sólo en las zonas más habitadas. En cambio, otros países cuya extensión es mucho más ajustada, como Costa Rica o la propia República Dominicana, pueden abordar la cartografía de todo su territorio con una escala homogénea y más grande, como 1:25,000 o 1:50,000.

En la siguiente infografía se muestra de forma esquemática el flujo habitual para la elaboración de esta cartografía en la actualidad, que se describirse a continuación.

Ilustración 1- 1: Flujo habitual para la elaboración de cartografía en la actualidad

Como se ha comentado en la introducción, el problema de conseguir coordenadas con una exactitud suficiente para estas escalas está prácticamente superado en la actualidad. Las técnicas de geodesia espacial y el uso generalizado de GNSS han permitido la definición de marcos de referencia muy precisos apoyados en sistemas de referencia globales, lo que facilita enormemente la obtención de coordenadas de cualquier punto del territorio y, en última instancia, la integración de datos de diversa procedencia.

Una vez establecido el marco de referencia en el que apoyarse, el siguiente aspecto que suele afrontar el productor de cartografía es identificar el conjunto de imágenes métricas que puede emplear para interpretar y extraer la información geométrica de los objetos geográficos que están presentes en el territorio. ¿Debe realizar un vuelo fotogramétrico exprofeso? ¿Puede disponer de otros vuelos fotogramétricos? ¿Puede emplear imágenes de satélite de alta resolución como alternativa al uso de fotogrametría aérea? Se debe tener en cuenta que las especificaciones de estos vuelos, o de los programas de teledetección en su caso, y su ejecución cumpliendo los requisitos de calidad exigidos tienen un impacto directo sobre las posibilidades de explotación de las imágenes que proporcionan. Por ejemplo:

- Los problemas con la métrica de las imágenes, con la aerotriangulación del vuelo o con los parámetros orbitales del satélite, tienen una repercusión directa sobre la exactitud posicional que se puede obtener.
- La presencia de nubes en las imágenes puede ocultar instancias de objetos geográficos e imposibilitar su captura.
- La temporalidad en la captura de las imágenes condiciona el grado de actualización que se puede conseguir en la información extraída.

Actualmente muchos países realizan vuelos fotogramétricos periódicos con objeto de satisfacer múltiples necesidades, como la monitorización de cambios que suceden en el territorio (deforestación, cambios en cultivos, aparición de nuevas construcciones, etc.) o para la propia elaboración de cartografía. En su defecto, existen diferentes programas de observación de la Tierra desde satélites que obtienen imágenes de alta resolución espacial que pueden ser empleadas con estos mismos propósitos y suponen una alternativa muy interesante frente a los vuelos fotogramétricos.

Así pues, una vez que se dispone de un conjunto de imágenes métricas que han sido apoyadas en el marco de referencia deseado, se está en disposición de realizar la captura masiva de los objetos geográficos. En primer lugar, se debe destacar la importancia de tener un universo de discurso, que no es más que la vista del mundo real o hipotético que incluye todo aquello que es de interés [ISO19101-1:2014], definido en un modelo conceptual. Este universo de discurso está totalmente condicionado por el producto cartográfico que se desea producir, y siempre incluye los elementos que deben ser modelados así como sus características y las relaciones existentes entre todos los componentes. Por ejemplo, si hablamos de cartografía topográfica a escala 1:25,000 generalmente existe la necesidad de modelar temáticas esenciales como son los transportes, la hidrografía, la orografía, las edificaciones o las divisiones administrativas, entre otras. Emana de aquí la importancia de la definición de unas normas de captura y de las reglas que deben cumplir por sí mismos los objetos geográficos recogidos, además de las implicadas en sus relaciones con los otros objetos con los que interactúan. Una forma normalizada de documentar esta información es a través del desarrollo de las Especificaciones del Producto de Datos [ISO19131:2007], donde particularmente adquieren especial relevancia los diagramas del modelo de datos y el correspondiente Catálogo de objetos geográficos

[ISO19110:2016], pues en ellos se recoge la definición detallada de los elementos componentes del producto.

De cara a la propia captura de la información de las imágenes, se emplean básicamente dos alternativas:

- La restitución fotogramétrica. La gran ventaja que presenta es que la visión estereoscópica del terreno proporciona una capacidad de fotointerpretación muy elevada. Por ejemplo, las vaguadas son claramente identificables y la interpretación de elementos construidos que sobresalen del terreno es relativamente sencilla. Además, la altitud de los elementos se obtiene en la propia captura de datos. Como contrapartida, se requiere hardware y software especializado, y operadores entrenados capaces de interpretar correctamente el terreno.
- La captura sobre ortofotografía. Su principal ventaja es la facilidad de uso. No requiere ningún equipamiento especial, se puede realizar en cualquier software de SIG y no precisa de una capacitación especial por parte de los operadores, si bien la experiencia en la fotointerpretación es importante. La altura de los elementos se asigna habitualmente utilizando un Modelo Digital del Terreno (MDT). Como contrapartida, las capacidades de fotointerpretación se ven limitadas al tener una visión desde un único punto de vista. Además, en ocasiones, el propio proceso de ortorrectificación puede haber alterado y degradado las imágenes.

Se debe destacar que, actualmente, las técnicas de inteligencia artificial están avanzando de forma muy significativa en la extracción automática de los elementos. Si bien todavía cometen errores en la interpretación y vectorización de los objetos geográficos, se han postulado como herramientas de indiscutible interés en la detección de cambios sucedidos en el territorio y se pueden llegar a posicionar como componentes de los flujos semiautomáticos de captura.

A pesar de las enormes ventajas que supone el uso de imágenes aéreas y de satélite en la elaboración de cartografía, hay una parte importante de la información que no se puede extraer de ellas, (ej. por encontrarse oculta por arbolado, etc.), pero especialmente por tratarse de información no detectable por un sensor, como son los nombres geográficos. Las revisiones de campo han sido siempre la herramienta fundamental para suplir esta carencia y conseguir toda la información necesaria para elaborar una buena cartografía. Sin embargo, su elevado coste de ejecución, tanto económico como en tiempo, choca con los requisitos actuales de inmediatez y economía, por lo que en muchas ocasiones no se realizan.

No obstante, hoy en día existe una multitud de fuentes de información que pueden utilizarse para completar todo el contenido definido por el modelo conceptual. Así, la integración de datos de otras fuentes resulta una fase fundamental en la elaboración actual de cartografía. Estas fuentes pueden ser de diferentes tipos. Por ejemplo, listados alfanuméricos, con algún tipo de identificador que permita relacionarlos a una posición; datos GPS tomados en campo por algún grupo de interés, o información geoespacial recopilada por otro organismo. Además, puede presentar diversas formas de acceso, desde servicios web interoperables hasta documentación en papel. Por tanto, una labor muy importante de las agencias cartográficas es la identificación de estos conjuntos de datos, el establecimiento de relaciones con los organismos que los proporcionan, el desarrollo de procesos que permitan la incorporación de la información que se requiere y el mantenimiento en el tiempo de estas interacciones.

Otro aspecto importante a destacar en la producción actual de la cartografía, es que la información suele almacenarse en bases de datos espaciales. Así, se habla de bases de datos

de información geográfica o de información geoespacial, cuyo propósito va más allá de la elaboración de cartografía. Entre las ventajas que proporciona el uso de estas bases de datos se debe destacar que permiten automatizar determinados procesos como, por ejemplo, el aseguramiento de la propia consistencia lógica de los datos; facilitan el establecimiento de relaciones con otros conjuntos de datos; posibilitan realizar la gestión del ciclo de vida de los elementos y permiten múltiples formas de explotar la información.

Por otro lado, las bases de datos geoespaciales no presentan ninguna limitación en cuanto a la cantidad de información que pueden albergar (más allá de las propias limitaciones de los sistemas gestores de bases de datos y de la capacidad de almacenamiento físico). Es decir, no existe la limitación espacial que tiene el mapa a la hora de contener datos. Este hecho hace que, en la práctica, se tienda a almacenar en la base de datos información de mayor detalle que la requerida para la producción de cartografía a una escala específica.

Por tanto, la generación de mapas a partir de la base de datos requiere la realización de varios procesos sobre los objetos geográficos contenidos en ella, como la selección priorizada de elementos, colapso, simplificación de geometrías, alineamiento, desplazamiento, etc., y la generación de las etiquetas con los textos de los nombres geográficos que se deben mostrar, todo ello con fines de permitir una buena legibilidad del mapa resultante, de ahí que se les suela denominar como procesos de redacción cartográfica. Van seguidos de una simbolización de los elementos, que debe ser acorde a la redacción cartográfica realizada y cuidadosamente elegida para garantizar que el mapa resultante transmita de forma eficaz la información que se desea.

Este conjunto de redacción cartográfica más simbolización es quizás una de las fases más complejas a la hora de elaborar una buena cartografía, puesto que se debe tener en cuenta las capacidades de percepción e interpretación del ojo humano. No obstante, a pesar de su complejidad, se han conseguido avances muy importantes en la automatización de los procesos, hasta el punto de que algunas organizaciones publican mapas generados de forma 100% automática. A pesar de ello, sigue habiendo una pequeña diferencia a favor de la redacción cartográfica realizada de forma manual por un experto frente a la que se consigue de forma automática. Así pues, una política habitual en las agencias cartográficas es la combinación de un proceso automático encadenado con otro de revisión y edición manual para la generación de sus series cartográficas.

1.3. Resumen de los contenidos de la guía

Los contenidos de esta Guía de Producción Cartográfica han sido elaborados en base a la experiencia y conocimiento del Instituto Geográfico Nacional de España en materia de generación de cartografía topográfica a escala 1:25,000, iniciado en los años 70’ con el comienzo de la producción del Mapa Topográfico Nacional 1:25,000 (MTN25), y continuado hasta nuestros días conforme a la evolución que ha sufrido la tecnología sobre la que se sustentan y optimizan los procesos de producción.

Además de este capítulo introductorio, la guía se compone de cinco capítulos que abordan las fases fundamentales del proceso, ilustradas con aquellos ejemplos de buenas prácticas que se han considerado relevantes para facilitar la comprensión de los conceptos técnicos que se describen en ellos. A continuación, se describe brevemente el contenido de cada uno de estos capítulos.

Capítulo 2: Adquisición de información geométrica

Principales fuentes de datos, técnicas y herramientas

En este capítulo se describe la primera fase del proyecto cartográfico, que comienza con el desarrollo del **modelado de datos** generado a partir de la abstracción de la realidad que se precisa representar y para cuya documentación se recomienda emplear la familia de normas ISO 19100 como referencia.

A continuación, se describen los principales **tipos de fuentes** de información a partir de las cuales se pueden capturar los datos con que poblar el modelo físico implementado en un entorno de base de datos: fuentes imagen, que proporcionarán fundamentalmente datos relacionados con la localización, y fuentes de información cartográficas o alfanuméricas existentes, donde la componente temática de los datos suele ser de gran riqueza.

En relación a la captura de datos de las **fuentes imagen**, se aportan **recomendaciones de captura** tanto para planimetría como altimetría además de criterios específicos para los objetos que se describen y que componen cada una de las **temáticas** habitualmente consideradas en este tipo de cartografía: orografía y paisaje, hidrografía, servicios e instalaciones y, edificios y construcciones. Esta sección del capítulo es extensa por el gran contenido de objetos geográficos que describe y los criterios de su captura que se aporta.

Respecto de las **fuentes de información cartográficas o alfanuméricas existentes**, se describen los formatos más habituales en los que se suelen encontrar: papel, raster, vectorial y tablas, aportando las principales características de cada uno de ellos.

Una vez descritas las principales características de los distintos formatos de las fuentes de referencia, el capítulo contempla una sección donde se explican las operaciones o **transformaciones** más habituales que se suelen realizar sobre las fuentes a la hora de proceder a integrar o utilizar de referencia sus datos en el proceso de generación de la información que compondrá el mapa topográfico. En el contenido se describen las transformaciones : 1) **de formatos**, es decir, procesos de impresión (de vector/ráster a papel), escaneado (de papel a ráster), vectorización (de ráster a vector), digitalización manual (de papel a vector), rasterización (de vector a ráster); 2) **de coordenadas y georreferenciación**, donde además de recopilar las más comunes, se da una serie de recomendaciones sobre su puesta en práctica, incidiendo con mayor detalle en la operación de remuestreo que se realiza en caso de la transformación de imágenes; 3) **de modelos de datos**, señalando las características principales de los modelos (geometría, atributos a conservar, tipo de dato, valores de dominio, cardinalidad y relaciones entre los componentes del modelo) que generalmente se ven afectadas cuando es necesario realizar esta operación en la fase de integración de fuentes de referencia.

Por último, se aborda una última tipología de información donde los datos de localización se encuentran implícitos en su componente temática: las **direcciones postales**. Sobre este conjunto de datos, de especial interés para determinar la localización precisa de topónimos relevantes en ámbitos urbanos, se aplican unas transformaciones propias que responden a la naturaleza de este dato: **normalización y geocodificación** tanto directa como inversa, que son descritas e ilustradas con referencias a herramientas existentes que permiten su aplicación.

Capítulo 3: Integración de información semántica

Ejemplos de fuentes de datos y técnicas

Este capítulo aborda el proceso de **integración de datos** tan frecuente como variable según lo sea el cómputo de los distintos posibles tipos de fuentes de datos de partida y los diferentes requisitos que se precisan considerar en la cartografía específica a generar en cada ocasión.

Así, dada su amplia variabilidad, este proceso se aborda desde una perspectiva global y común a todos los casos en base a la doble componente que caracteriza al dato geográfico: la **semántica y la geométrica**.

En relación a la primera, entre otros aspectos, se remarca la relevancia del mantenimiento de los **identificadores únicos** de los objetos en las fuentes de referencia para la optimización de la metodología de actualización de la cartografía. Respecto a la geometría, se profundiza con mayor alcance en los **procesos de conflación**, distinguiendo entre las particularidades que presentan las geometrías **superficiales** y las geometrías **lineales**, y aportando una serie de recomendaciones sobre los pasos a seguir en cada proceso (selección de candidatos, medida de similitud, etc.). Además, para el caso de geometrías lineales se aportan dos **casos de uso** basado en la experiencia desarrollada por el IGN de España y que se consideran relevantes por ilustrar el contenido y fundamentalmente en el caso de uso de detección de cambios en redes de transporte a partir de datos de **Open Street Map (OSM),** porque puede ser extrapolado a otros ámbitos geográficos.

El input de este proceso de integración y la accesibilidad a él es clave para la viabilidad de su aplicación. Por este motivo, la segunda mitad del capítulo contiene la descripción de distintas **técnicas de acceso a las fuentes de datos existentes en la red** que son de especial relevancia para las agencias cartográficas a la hora de realizar los procesos de actualización de sus conjuntos de datos.

Así, en primer lugar, se describe cómo la información de las webs que se consideren de referencia puede ser rastreada por **bots** diseñados para la búsqueda de palabras claves evocadoras de cambios en los objetos geográficos (ej. en el ámbito de las carreteras podrían ser: cesión, inauguración, trazado, etc.).

En segundo lugar, dada su globalidad y su naturaleza de acceso libre, se dedica un amplio apartado a explicar las diversas formas de consulta, descarga y tratamiento de datos descargados procedentes del conjunto de información que compone la cartografía colaborativa de **OSM**.

En tercer lugar, el capítulo se centra en la utilidad de los servicios web estándar de descarga **Web Feature Service** (**WFS)**, a partir de un **ejemplo de buenas prácticas** sobre un sistema creado para la consulta y el consumo de los datos de origen catastral que se facilitan a través de este sistema.

Capítulo 4: Calidad de la información geográfica

Norma ISO 19157 y estrategias de Gestión de la Calidad en los datos geográficos

El objeto de este capítulo es describir los aspectos fundamentales a considerar en la gestión de la calidad de la Información Geográfica.

Comienza con unas **consideraciones preliminares** donde se aporta la descripción de los conceptos básicos en materia de **calidad** y de la **metodología** a aplicar para su gestión tanto en los procesos de producción como sobre el resultado final del producto.

El capítulo incluye una sección específica sobre la **norma ISO 19157** que establece los principios para describir la calidad de los datos geográficos, los procesos de su evaluación y el reporte de los resultados, motivada por el carácter global de esta familia de normas y, por tanto, por la viabilidad de ser implementada en cualquier ámbito geográfico. Tras describir la **estructura de la norma**, el capítulo aporta la relación de los **componentes** que conforman la calidad de datos: unidad y elementos de calidad, descriptores de los elementos (medida, método de evaluación y resultado) y metacalidad. Por último, aporta un epígrafe a título de **guía práctica** para la aplicación de la norma, que contiene una serie de recomendaciones a considerar a la hora de implementar las fases del proceso de evaluación de la calidad de la información geográfica, ilustradas con referencias a ejemplos de buenas prácticas.

El proceso de evaluación de la calidad en un producto cartográfico que contempla gran diversidad de objetos geográficos por cada una de las temáticas consideradas, debe ser abordado desde la perspectiva de la **automatización de los controles de calidad.** El capítulo contiene un epígrafe entero dedicado a estos conceptos, donde se explica la viabilidad de implementar un proceso automático según el elemento de calidad a evaluar y se aporta una relación de **herramientas** ampliamente utilizadas en este sector que pueden ser consideradas de referencia a la hora de aplicar los controles.

Desde esta perspectiva más práctica, el capítulo diferencia dos escenarios de aplicación de controles de calidad:

1) Durante el <u>proceso de producción</u>: se aporta una serie de **recomendaciones y estrategias** a seguir **en la aplicación de los controles de calidad que evalúan cada elemento de calidad** de los datos durante el proceso de producción. Éstas se ilustran con imágenes y referencias a **funcionalidades de las herramientas** nombradas anteriormente, y que permiten implementar dichas recomendaciones.
2) Sobre el <u>producto final</u>: contempla **estrategias** para los procesos de inspección o **validación de los datos por cada elemento de calidad a evaluar**, y para la **corrección de las incidencias** detectadas, ilustrando el contenido con ejemplos de casos reales y aportando, según el elemento de calidad que se trate, una serie de consideraciones a tener en cuenta.

Por último, dada la relevancia de la componente espacial en este tipo de información y lo voluminosos que suelen ser los conjuntos de datos geográficos, el capítulo tiene unas últimas secciones donde se describen, por un lado, los **aspectos generales de la calidad posicional**, y por otro, las **estrategias de muestreo** que pueden emplearse a la hora de evaluar grandes conjuntos de datos donde es inviable la revisión exhaustiva de todos y cada uno de los componentes del producto cartográfico.

Capítulo 5: Producción del Mapa I. Generación de los datos cartográficos

Métodos de producción: proceso automático y Edición simplificada

Tras haber visto los aspectos fundamentales de las fuentes de datos, las metodologías para su captura e integración y las estrategias de evaluación de la calidad de la información geográfica, a partir de los capítulos 5 y 6, la guía se centra en los contenidos propios de la generación de cartografía a partir de los datos validados.

El capítulo 5 aborda la primera fase cuyo objetivo es **transformar los datos geográficos en datos cartográficos**, por aplicación de las tareas de redacción y formación cartográfica que requiere la producción de cartografía topográfica a la escala 1:25,000.

El capítulo comienza con una **parte introductoria sobre la cartografía**, a partir de la descripción de conceptos básicos generales a considerar en la representación de la realidad a través de un mapa (sistema geodésico de referencia, proyección cartográfica, escala, etc.), y más particulares para el caso de la **cartografía topográfica**. En esta segunda parte se listan las temáticas que más frecuentemente componen este tipo de cartografía, ilustrando el contenido que se contempla en el Mapa Topográfico Nacional 1:25,000 de España.

A continuación, se describe la **metodología de producción automática** de este producto, señalando como preámbulo la motivación que justifica el paso de la producción tradicional a la automática. Esta extensa sección recopila la **experiencia del IGN de España** en cuanto a la evolución del proceso de producción cartográfico del MTN25, desarrollado a partir de la documentación correspondiente a la metodología de generación tradicional del mapa (normas de edición y normas de toponimia), y la implementación de los automatismos que simulan dichas pautas. Contiene epígrafes dedicados a describir los componentes principales de esta metodología: 1) las **fuentes de datos origen** y el tratamiento de correspondencias hacia el modelo de datos cartográficos; 2) los **procesos automáticos de redacción cartográfica**, donde se describen las bases en las que se fundamenta esta automatización así como los operadores más importantes utilizados (filtrado, simplificación, suavizado, etc.); 3) el **etiquetado automático**, que es, sin duda, uno de los desafíos más complejos de todo el proceso donde se aborda la selección y correcta ubicación de la toponimia en forma de rótulos; 4) **almacenamiento** de los datos cartográficos generados **en base de datos**. Por último, se incorpora la sección dedicada a la explicación sobre la **implementación del flujo de automatización del proceso de redacción cartográfica** explicado anteriormente, siguiendo las directrices que ahí se plantean, ilustrada con el caso de buenas prácticas del MTN25.

El tercer y último gran apartado de este capítulo está dedicado a explicar la **metodología** que se aplica cuando la salida de la producción cartográfica requiere de una mayor calidad, pues tiene por destino convertirse en una publicación impresa y, por tanto, precisa de una **edición adicional** sobre los resultados obtenidos del automatismo. En esta sección se describe la solución adoptada por el IGN de España, que es la metodología denominada "**MTN25 Edición Simplificada**", y que consiste en aplicar un protocolo de revisión y corrección de datos, compuesto por los ítems más sensibles y susceptibles de incorporar incidencias debidas al automatismo.

Capítulo 6: Producción del Mapa II. Visualización final del producto

Técnicas de simbolización, sombreado y rasterizado.

El último capítulo aborda la fase final de la producción cartográfica en la que, sobre los datos cartográficos generados según la metodología explicada en el capítulo anterior, se aplican las técnicas de **simbolización, incorporación del sombreado del relieve y rasterización** para la generación del producto en el formato digital final.

El capítulo contiene una primera parte en la que se incluyen los conceptos básicos relativos a la **simbolización.** Comienza con una parte introductoria acerca de conceptos básicos pero imprescindibles para la correcta aplicación de dichas técnicas, ligados a la **teoría del color** y su impacto en la impresión offset. A continuación, explica las **características** de la simbología, tratada como un elemento fundamental del diseño gráfico para la correcta trasmisión del mensaje, en base a la tipología de símbolos, las características visuales y el empleo de rótulos como signos convencionales. Por último, esta sección incluye un apartado donde se lista el **software de referencia** a considerar en la producción de este tipo de cartografía.

A continuación, se aborda la técnica del **sombreado** o representación bidimensional del modelo digital del terreno, cuya integración en el mapa facilita enormemente la interpretación del relieve de la zona representada y realza y embellece el aspecto final del producto. Tras una sección introductoria sobre los modelos digitales del terreno, el capítulo describe, en base al ejemplo de buenas prácticas sobre cómo se realiza en el IGN de España, la técnica de **obtención del sombreado** (proceso y software empleado), su agrupación en mosaicos de grandes ámbitos del territorio, y la **exportación final** de éste por unidades u hojas que componen la serie cartográfica.

El tercer bloque del capítulo tiene por objeto la explicación de los **procesos de simbolización e integración de sombreado**, diferenciando entre procesos que se realizan de forma manual y procesos en los que esas fases se encuentran automatizadas. Todo el contenido se describe ligado al software de referencia que se emplea en cada metodología.

Las dos últimas secciones del capítulo abordan el tema de las diferentes **salidas** que se pueden generar a la hora de publicar los resultados del proceso de producción del mapa, en función del caso de uso que se trate: formato **impreso**, archivos **vectoriales** y archivos **ráster**. En último lugar, se incluye un apartado dedicado a las connotaciones y diferencias que existen entre la producción cartográfica destinada a ser impresa, y por tanto constreñida a la imagen estática del territorio que determina su escala, y la publicación de información geográfica demandada para su consulta a través de **visualizadores**, donde la escala ya no es un factor limitante. Se aporta la experiencia de este campo del IGN de España a través de la descripción del denominado **mapa de Alta Resolución**.

CAPÍTULO 2: ADQUISICIÓN DE INFORMACIÓN GEOMÉTRICA

Principales fuentes de datos, técnicas y herramientas

Descripción breve

Exposición de las principales fuentes y metodologías empleadas en la captura de la información, principalmente geométrica y adicionalmente semántica, necesaria para la elaboración de la cartografía.

Índice de contenidos

2.1. Introducción 1

Modelo de datos 1

Datos geográficos 2

Principales fuentes de datos geográficos 2

2.2. Captura de datos a partir de imágenes 4

Modos de captura 4

Captura planimétrica en modo continuo 5

Captura planimétrica en modo discreto 5

Captura planimétrica puntual 6

Captura altimétrica 6

Criterios de Selección 7

Normas de captura para objetos identificables sobre una imagen 7

Orografía y paisaje 8

Espacio natural protegido 8

Curva de nivel 8

Punto de cota 10

Lugar 11

Cadena montañosa 12

Vegetación 13

Cortafuegos 14

Hidrografía 14

Curso natural 14

Curso artificial 21

Aguas marinas 25

Costa 25

Laguna-lago 27

Humedal 29

Embalse 29

Punto hidrográfico 32

Isla 33

Presa 34

Servicios e instalaciones ... 37
Espacio dotacional ... 37
Transporte de energía ... 42
Estructura ... 44
Edificios y construcciones ... 46
Población o Mancha urbanizada ... 46
Edificio ... 47
Construcción de almacenaje ... 51
Monumento ... 53
Cerramiento ... 55
Infraestructura de cruce ... 56
Pista deportiva ... 56
Construcción portuaria ... 58
Auxiliar construido ... 60
Transportes ... 60
Vía de comunicación ... 60
Transporte por rieles ... 63
Teleférico ... 64
Infraestructura de transporte ... 64
2.3. Integración de cartografía e información geográfica existente ... 67
2.3.1. Formato del dato de entrada ... 67
2.3.1.1. Formato papel ... 67
2.3.1.2. Formato ráster ... 67
Definición ... 67
Formatos de datos ráster habituales ... 68
2.3.1.3. Formato vectorial ... 71
Definición ... 71
Estructura de datos en el formato vectorial ... 71
Formatos vectoriales habituales soportados ... 73
2.3.1.4. Tablas y datos estructurados ... 74
Archivos de texto ... 74
Archivos semiestructurados o de intercambio de datos ... 75
Bases de Datos ... 76
2.3.2. Transformaciones ... 77
2.3.2.1. De formatos ... 77
Impresión (ráster/vectorial -> papel) ... 78

Escaneado (papel -> ráster) 78
Vectorización (Ráster-> Vectorial) 79
Digitalización manual (Papel-> Vectorial) 80
Rasterización (Vectorial -> Ráster) 81
2.3.2.2. Transformación de coordenadas y georreferenciación 82
2.3.2.3. Transformación de modelo de datos 84
2.4. Información sin coordenadas explícitas: geocodificación de direcciones 87
2.4.1. Normalización de las direcciones postales 88
2.4.2. Geocodificación: callejeros y geocodificadores 90
Ejemplo de funcionamiento de un servicio de geocodificación de direcciones 91
Servicios de Geocodificación existentes 93

Índice de ilustraciones

Ilustración 2 - 1: Toma de imágenes en un vuelo fotogramétrico 4
Ilustración 2- 2: Método de captura en modo continuo 5
Ilustración 2- 3: Método de captura en modo discreto 6
Ilustración 2- 4: Restitución 3D con modelos estereoscópicos 6
Ilustración 2- 5: Detalle de curvas de nivel cerca de un embalse 10
Ilustración 2- 6: Detalle de captura playas, salientes costeros y bahías 12
Ilustración 2- 7: Captura de cadena montañosa (destacada en azul) por su eje representativo principal 13
Ilustración 2- 8: Captura de curso natural solo como eje (ancho <5m) 17
Ilustración 2- 9: Curso natural canalizado de ancho <5m 17
Ilustración 2- 10: Conexión en un sólo brazo de isla fluvial 18
Ilustración 2- 11: Eje dentro de embalse 18
Ilustración 2- 12: Conexión dentro de tramo superficial (ancho >5m) 18
Ilustración 2- 13: Tramo de río cubierto al atravesar una presa 19
Ilustración 2- 14: Tramo de río cubierto al atravesar una población 19
Ilustración 2- 15: Tramos de río cubierto al atravesar carreteras 19
Ilustración 2- 16: Río con representación superficial que desemboca en el mar 20
Ilustración 2- 17: Agua aparente (lleva agua en el momento de la captura) 20
Ilustración 2- 18: Cauce habitual aparente (no lleva agua en el momento de la captura y las crecidas son identificables en el terreno) 20
Ilustración 2- 19: Canal en superficie 23
Ilustración 2- 20: Acequia en superficie 23
Ilustración 2- 21: Tubería elevado 24
Ilustración 2- 22: Tramo subterráneo que cruza una vía 24
Ilustración 2- 23: Tramo subterráneo no capturado por ser < 3m 24
Ilustración 2- 24: Cauce artificial no conectado con embalse – Captura incorrecta 25

Ilustración 2- 25: Captura de un polígono de aguas marinas coincidente con la línea de costa 25
Ilustración 2- 26: Línea de costa natural 26
Ilustración 2- 27: Línea de costa artificial en un puerto 26
Ilustración 2- 28: Curva de cota 0 obtenida con MDT (morado) y línea de costa en cota 0 resultante (azul) 27
Ilustración 2- 29: Línea de costa en cota 0, desplazada de la línea tierra-mar visible en la ortofoto 27
Ilustración 2- 30: Presa no ajustada a laguna y masas de agua desconectadas - Captura incorrecta 28
Ilustración 2- 31: Embalse con los ríos lineales conectados en su interior 31
Ilustración 2- 32: Geometrías coincidentes en los contornos de embalse, presa y río 31
Ilustración 2- 33: Sumideros en ríos en los que se pierde el trazado en una zona muy llana 33
Ilustración 2- 34: Captura de islas marítimas 34
Ilustración 2- 35: Captura de islas fluviales (aguas interiores) 34
Ilustración 2- 36: Presa de materiales sueltos capturada por su contorno 35
Ilustración 2- 37: Presa de hormigón capturada por su contorno 36
Ilustración 2- 38: Dique de contención capturado como presa 36
Ilustración 2- 39: Represa en un cauce natural 36
Ilustración 2- 40: Ejemplo de captura para una mina a cielo abierto 39
Ilustración 2- 41: Ejemplo de captura para una salina 40
Ilustración 2- 42: Ejemplo de captura para una central hidroeléctrica 40
Ilustración 2- 43: Ejemplo de captura para una zona industrial 40
Ilustración 2- 44: Captura de un área recreativa con una alambrada como cerramiento 41
Ilustración 2- 45: Ejemplo de captura para una depuradora 41
Ilustración 2- 46: Ejemplo de captura para un observatorio espacial 41
Ilustración 2- 47: Ejemplo de captura para un parque-jardín 42
Ilustración 2- 48: Central termosolar 42
Ilustración 2- 49: Captura correcta de dos líneas eléctricas 44
Ilustración 2- 50: Torre eléctrica - Captura correcta en la base de la estructura 45
Ilustración 2- 51: Torre eléctrica - Captura incorrecta en la parte superior de la estructura 45
Ilustración 2- 52: Torre eléctrica - Captura incorrecta en la sombra de la estructura 46
Ilustración 2- 53: Ejemplo de captura de una población 46
Ilustración 2- 54: Plaza de toros - Captura correcta 49
Ilustración 2- 55: Plaza de toros - Captura incorrecta 49
Ilustración 2- 56: Faro - Captura correcta 50
Ilustración 2- 57: Faro - Captura incorrecta 50
Ilustración 2- 58: Edificios ligeros en área de servicio - Captura correcta (cubiertas de protección surtidores) 51
Ilustración 2- 59: Edificio ligero en playa - Captura incorrecta (se trata de un elemento no estable) 51
Ilustración 2- 60: Captura de silos 53
Ilustración 2- 61: Captura de depósito de agua 53
Ilustración 2- 62: Fuente ornamental que debe capturarse como monumento 54
Ilustración 2- 63: Fuente que permite coger agua que debe capturarse como punto hidrográfico 54
Ilustración 2- 64: Pistas de tenis y campo de fútbol ubicados en una zona deportiva 57
Ilustración 2- 65: Campo de fútbol en el interior de una pista de atletismo 58
Ilustración 2- 66: Puerto con muelle (rojo), pantalán (naranja) y rompeolas (negro) 59

Ilustración 2- 67: Pantalán capturado por eje (ancho<5m) y por contorno (ancho>5m) 59
Ilustración 2- 68: Rompeolas fuera de un puerto, en este caso para proteger la costa 60
Ilustración 2-69: Aparcamiento junto a una playa 66
Ilustración 2-70: Estructura del formato TIFF 70
Ilustración 2-71: Ejemplo de documento KML (fuente: https://es.wikipedia.org/wiki/KML) 75
Ilustración 2-72: Ejemplo de documento GeoJSON (fuente: https://open.canada.ca/data/en/dataset?collection=fgp&_res_format_limit=0&topic_category=geoscientific_information&keywords=Government+information&res_format=GEOJSON&res_format=GML) 76
Ilustración 2-73: Esquema de transformaciones entre formas de entrada gráfica 77
Ilustración 2-74: Ejemplo de diferentes nomenclaturas de calles 87
Ilustración 2-75: Tabla ejemplo de direcciones sin normalizar de farmacias en Cáceres..... 88
Ilustración 2-76: Tabla ejemplo de direcciones normalizadas por campos 89
Ilustración 2-77: Ejemplo de variantes de representación de una dirección: arriba, según la estructura y el orden de los componentes, y abajo, según el contenido de estos. 90
Ilustración 2-78: Ejemplo de datos completos de un registro de dirección. Fuente Catastro (España) 91
Ilustración 2-79: Estructuración de componentes de la dirección 92
Ilustración 2-80: Definición de la lista de valores para cada componentes de la dirección 92
Ilustración 2-81: Ejemplos de variantes de representación de una dirección 92

2.1. Introducción

Modelo de datos

El primer paso para elaborar un producto cartográfico es determinar qué información se precisa disponer y representar para garantizar que la cartografía generada transmite correctamente el mensaje completo que motivó su producción, que en el caso que nos ocupa es la descripción topográfica del territorio a escala 1:25,000.

Dicho de otro modo, lo primero que debe hacer el cartógrafo es determinar las temáticas que compondrán el mapa, así como los objetos geográficos incluidos en cada una, junto con sus atributos, restricciones y relaciones.

Por tanto, es necesario realizar un ejercicio de "abstracción" de la realidad para identificar un universo de discurso, que no es más que la *vista del mundo real o hipotético que incluye todo aquello que es de interés*. El universo de discurso incluye componentes u objetos geográficos, como, por ejemplo, cursos de agua, edificios, carreteras, coberturas del suelo, límites de propiedad, etc.; sus características (propiedades, funciones), y las relaciones existentes entre ellos [2].

Los conceptos del universo de discurso se definen en un **modelo conceptual**, que se describe formalmente en un esquema conceptual, utilizando un lenguaje de esquemas conceptuales. El lenguaje habitualmente utilizado es el Lenguaje Unificado de Modelado (UML). Los objetos geográficos se definen mediante clases, en las que principalmente se indica su denominación, atributos y operaciones, y, a través de una simbolización de líneas y flechas, se expresa la tipología de relaciones que existe entre ellos, así como su cardinalidad.

Un esquema conceptual que define cómo se describe un universo de discurso desde la perspectiva de los datos se llama **esquema de aplicación**, y es independiente de las tecnologías y plataformas de implementación física [2]. Así pues, *un esquema de aplicación provee la descripción formal de la estructura de datos y el contenido requerido por una o más aplicaciones*. La norma ISO 19109:2015 [3] proporciona las reglas para la elaboración de esquemas de aplicación.

La documentación del modelo se completa a través del catálogo de objetos geográficos, que define los conceptos utilizados en el esquema de aplicación. Así, según la ISO 19101-1:2014 [2], el catálogo de objetos geográficos contiene las definiciones y descripciones de los tipos de objeto geográfico, de sus atributos y de las relaciones que se dan entre objetos geográficos de uno o más conjuntos de datos, junto con las operaciones que se les puedan aplicar. La norma ISO 19110:2016 [4] proporciona la metodología para la catalogación de objetos geográficos.

Según la ISO 19131:2007 [5], de Especificaciones de producto de datos, el esquema de aplicación y el catálogo de objetos geográficos deben ser referidos, junto con una descripción narrativa, desde las especificaciones del producto de datos. Además de esta información, las especificaciones deben cubrir también otros aspectos del producto, como su identificación, el campo de aplicación, su distribución, el sistema de referencia, etc.

Una vez finalizada la definición completa del modelo y sus componentes, el siguiente paso es materializarlo en un modelo físico o estructura "tangible", generalmente en el entorno de bases de datos donde los objetos geográficos se almacenan en tablas, y sus propiedades, en los atributos que las componen.

Datos geográficos

Después de realizar la definición conceptual del proyecto cartográfico y que quede plasmado en el modelo físico, el siguiente paso es disponer de los datos con los que poblar dicho modelo. Antes de comenzar la descripción de las posibles fuentes de datos y las metodologías de captura conviene recordar brevemente qué implica el término de datos geográfico.

La norma ISO 19101-1:2014 [2] define los Datos Geográficos como "*datos acerca de objetos geográficos asociados implícita o explícitamente con una localización relativa a la Tierra*". De esta definición se deduce que la información geográfica se estructura en dos grandes componentes: una geoespacial o de localización respecto de un sistema de referencia, y otra temática o descriptiva de las características del objeto geográfico a cartografiar.

A su vez, son frecuentes otras clasificaciones de los datos basadas en la dimensión espacial del objeto geográfico, (puntual, lineal, superficial [polígono]), y si ésta viene expresada en un marco planimétrico (2D) o tridimensional (3D). Por último, atendiendo al contenido temático y a la completitud del producto, es frecuente utilizar la terminología de *división vertical* de la información, considerando cada temática como una capa de modo que la suma consistente de todas conformará el contenido completo del mapa, y de *división horizontal* o estructuración del producto completo en "hojas" o "celdas" de una cuadrícula con la que, el administrador de la cartografía decide dividir la totalidad del producto (serie cartográfica) para garantizar su correcta gestión cuando se trata de grandes extensiones del territorio (p.e.: Países).

Principales fuentes de datos geográficos

Como se ha señalado en el capítulo 1, en la actualizad existe una multitud de fuentes de información que puede utilizarse para completar todo el contenido que se necesita para elaborar la cartografía. La principal fuente de información que se suele emplear para conformar todo el esqueleto geométrico son las imágenes aéreas o de satélite, que permiten fotointerpretar la mayor parte de los objetos geográficos que forman parte del universo de discurso y capturarlos con una calidad métrica que generalmente es adecuada.

También es frecuente que se disponga de otra información de interés para el propósito que se persigue, bien porque ya existe o bien porque se ha elaborado exprofeso. Por ejemplo, datos de campo tomados con GPS (o con cualquier otra técnica) con la localización de elementos concretos o con tracks de elementos lineales; minutas de revisión de campo; cartografía existente elaborada con anterioridad o por otras organizaciones, por ejemplo, mapas forestales, cartografía catastral, etc.

A menudo también se dispone de otros datos geográficos que, si bien no tienen una referencia explícita a su localización, si lo hacen implícitamente, por ejemplo, mediante un identificador que permite relacionarlos con datos georreferenciados, o porque contienen una dirección postal que, a través de procesos de geocodificación directa, se traduce en una localización concreta.

Las fuentes de información se encuentran en formas muy variadas, por ejemplo, en papel, en imágenes digitales (ráster) o en cualquiera de los múltiples formatos existentes para almacenar información vectorial y alfanumérica. A su vez, dentro de lo que sería información digital, hay también diversidad de formas de acceso, desde archivos, que se pueden tener almacenados en disco o descargarse vía internet, hasta servicios interoperables que facilitan su consulta y acceso.

A lo largo de este capítulo y del siguiente, se desarrollan conceptos y técnicas que pueden resultar útiles para la incorporación de todo este tipo de datos geográficos. No obstante, antes de incorporar cualquier información es preciso evaluar la calidad de la misma en todas sus facetas (completitud, posicional, temática, temporal). Así, se debe determinar si es válida para el producto que se está elaborando y qué precauciones se deben considerar.

2.2. Captura de datos a partir de imágenes

El empleo de imágenes -aéreas o satelitales- como base para la captura y actualización de los objetos geográficos de una base de datos topográfica, se ha convertido, hoy en día, en una de las técnicas más ampliamente utilizadas. Frente a los clásicos trabajos de campo, la captura sobre imagen abarata notablemente los costes, cubre una gran superficie de territorio de manera sencilla y rápida y, dependiendo de la resolución, ofrece un nivel de detalle muy notable. Aunque es cierto que para determinados objetos o características el uso de la imagen no es suficiente y se necesitan de otras fuentes de datos para completar su registro, en general, su uso es sencillo, asequible y solo requiere de un aprendizaje en técnicas de fotointerpretación.

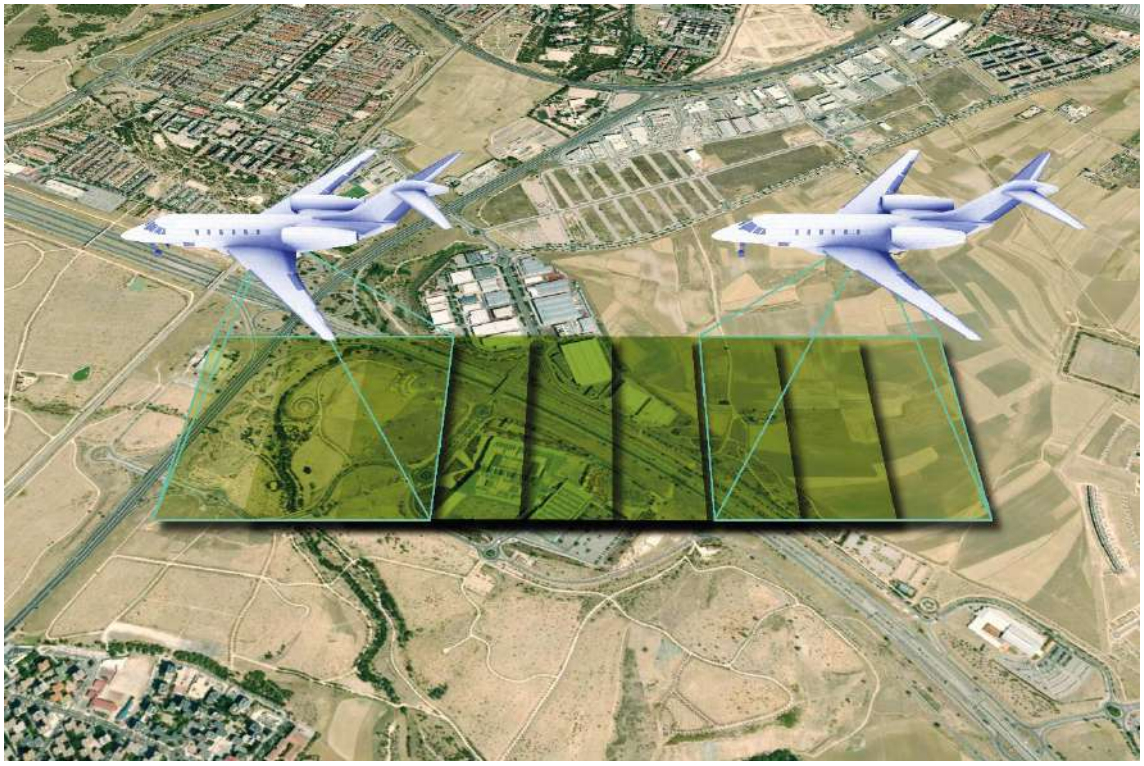


Ilustración 2 - 1: Toma de imágenes en un vuelo fotogramétrico

El proceso de captura de datos por fotointerpretación de imágenes consiste en el registro de las coordenadas X, Y, H, que definen la geometría de aquellos objetos geográficos definidos en el catálogo, acorde a los criterios impuestos en el mismo, y que sean reconocibles en los pares estereoscópicos o en las ortofotografías aéreas.

El presente apartado recoge las normas, características principales y nivel de detalle que hay que tener en cuenta para la captura de los objetos geográficos que habitualmente se toman cuando la escala de referencia es 1:25,000.

Modos de captura

La captura de los elementos se puede realizar de distintos modos, principalmente, modo continuo y modo discreto, y con distintas técnicas, básicamente, restitución fotogramétrica y captura sobre ortofoto más MDT. A continuación, se detallan algunas características propias de cada modo de captura.

Captura planimétrica en modo continuo

Se trata de un registro continuo de coordenadas en el que se sigue con el cursor el contorno del elemento. Se usa principalmente para el registro de elementos lineales (curvas de nivel, carreteras, cursos de agua, etc.).

Para elementos capturados en modo continuo, los parámetros de registro que deben establecerse son los siguientes:

- A partir del primer vértice del elemento debe insertarse un punto cada 5 m (distancia mínima entre puntos muestreados); si el ángulo entre las alineaciones de los vértices 'n'-'n+1' y 'n+1'-'n+2' es menor de 10 grados, o la superficie encerrada es menor de 300 m^2, se ampliará la distancia entre puntos registrados a 20m. Y así sucesivamente con el resto de los vértices de la línea.

En el gráfico siguiente, los círculos representan puntos muestreados y los triángulos vértices registrados, siendo 5 m la distancia mínima entre puntos muestreados y 20 m la distancia máxima que debe existir entre vértices registrados.

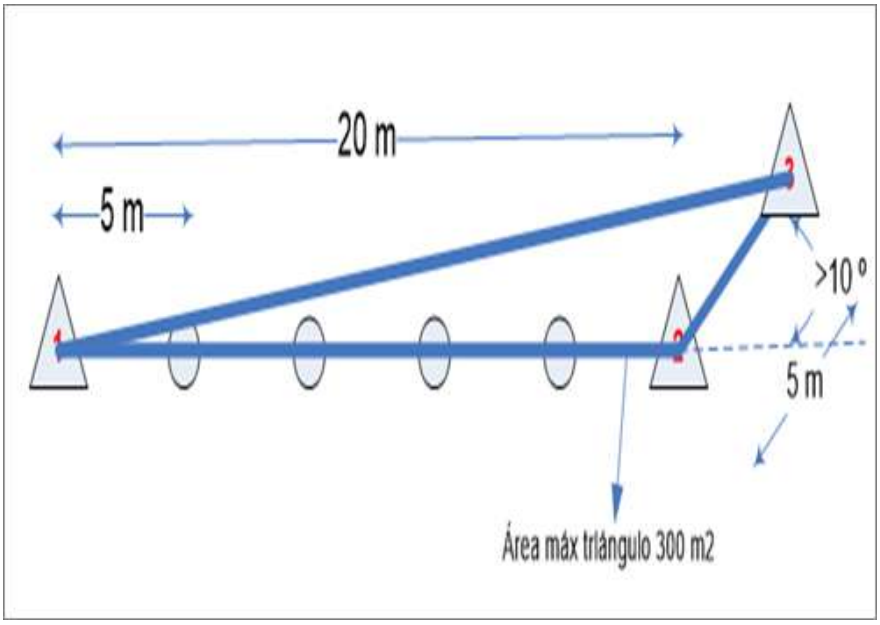


Ilustración 2- 2: Método de captura en modo continuo

Captura planimétrica en modo discreto

Se trata de un registro de coordenadas que se realiza punto a punto definiendo el contorno del elemento. Se usa principalmente para objetos que se registran por sus esquinas (casas, cercas, etc.).

Para elementos capturados en modo discreto, los parámetros de registro que deben establecerse son los siguientes:

- Para elementos superficiales, se establecerá un vértice en cada uno de los puntos que se quiera registrar, estableciendo una longitud mínima de registro de 1 m.
- Para elementos lineales, la captura debe garantizar que la flecha formada cada 3 vértices registrados de la línea sea superior a 2 m.

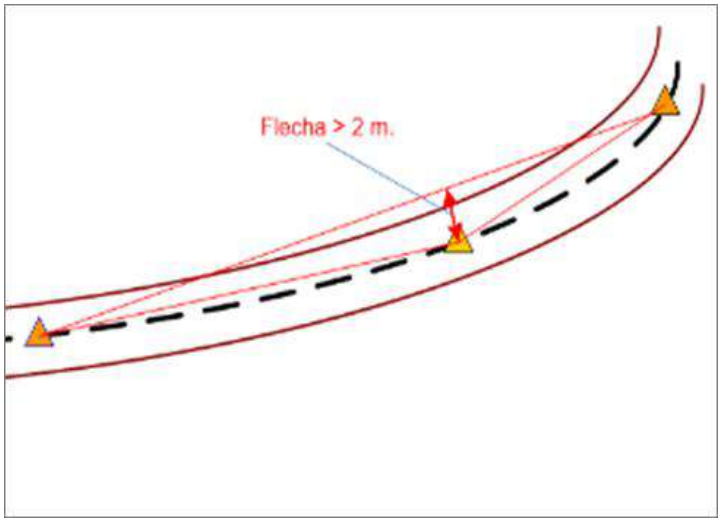


Ilustración 2- 3: Método de captura en modo discreto

Captura planimétrica puntual

Se utiliza para el registro de objetos cuya representación queda materializada mediante un punto (monumento, fuente, etc.).

Captura altimétrica

A la vez que se realiza el registro planimétrico (X,Y) sobre la imagen, es importante obtener también el valor altimétrico (H) de cada uno de los puntos capturados.

Restitución fotogramétrica. Cuando la captura se realiza mediante restitución-pares estereoscópicos el problema es trivial, ya que la cota es inherente a la identificación de los distintos puntos homólogos del objeto que se esté capturando. Este modo de captura permite una capacidad de fotointerpretación muy alta, ya que el operador puede discernir claramente las distintas profundidades de los elementos y la pendiente del terreno. Sin embargo, requiere de equipamiento más especializado y operadores más entrenados.

Ilustración 2- 4: Restitución 3D con modelos estereoscópicos

Captura sobre ortofotografía más MDT. Otra posibilidad es utilizar imágenes sin estereoscopía (ortofotografía, imágenes satelitales). En este caso, y siempre que sea posible, se recomienda el uso conjunto de la imagen con un modelo digital del terreno (MDT). Para una escala de referencia 1:25,000, se recomienda que el paso de malla del MDT sea de 5 – 10 m.

No obstante, puede ocurrir que al realizar la captura no se disponga del valor altimétrico. En este caso es conveniente asignar a la coordenada H de los puntos capturados un valor fuera de dominio, por ejemplo -99999. De esta manera será muy sencillo identificar estos elementos para poder asignarles H a posteriori, generalmente con la ayuda del MDT.

Como norma general, debe garantizarse con la mayor precisión posible la coincidencia de altitudes en los cruces altimétricos o con otros detalles planimétricos, permitiéndose un error máximo altimétrico de 2,5 m.

Criterios de Selección

Los criterios de selección pretenden establecer, en los objetos geográficos en los que sea de aplicación, parámetros objetivos de elección de la información, es decir, cuándo un objeto es o no representable, o qué intensidad del mismo ha de representarse, siendo la escala de referencia 1:25,000.

Los criterios más habituales están basados en el ámbito geográfico, la aridez, la humedad y la temperatura, la densidad y la pendiente de la zona de trabajo, que se detallarán en cada uno de los objetos geográficos definidos en el siguiente apartado dedicado a las normas de captura sobre imagen. Por ejemplo, un cauce natural cuya longitud sea inferior a 500 m no debe tener representación en zonas húmedas, o el área mínima que delimita una curva de nivel para ser capturada debe ser de 25x25m^2.

En definitiva, los criterios de selección tienen por objetivo establecer cuándo un elemento es notable en un entorno y debe ser representado, y cuándo el mismo elemento deja de ser notable en otro entorno y no debe capturarse. Ese carácter singular hace que la entidad sea representada a pesar de tener un tamaño menor que el establecido para su captura, o que no sea necesario su registro, aunque cumpla los criterios de tamaño o longitud mínimos.

Las medidas mínimas establecidas para la representación de los objetos aquí definidos son aproximadas, y, en cada caso, deberán estar sujetas a valoración, ya que su captura dependerá de la densidad e importancia de dicho objeto en la zona. Por ejemplo, un estanque de superficie inferior a la establecida en las normas de captura deberá tener representación en zonas de escasez de elementos hidrográficos.

Normas de captura para objetos identificables sobre una imagen

En este apartado se explican las normas de captura y se especifica el contenido y el nivel de detalle que deben tener los elementos que, habitualmente, se toman del territorio para la generación de una base topográfica a escala 1:25,000. Se han incluido los objetos geográficos -y sus atributos descriptivos- que se consideran básicos para conformar una base de datos topográfica de carácter general. No obstante, puede ocurrir que algún objeto, atributo o valor no sea de aplicación en el territorio considerado, por lo que simplemente se omitirá. También puede producirse el caso contrario, es decir, que no se haya incluido algún objeto, atributo o valor debido a su carácter local.

Igualmente se ha intentado utilizar una terminología de carácter general para nombrar a los objetos geográficos y sus atributos, pudiendo ser necesaria su adaptación a idiomas o características locales.

También es importante reseñar que los objetos y los atributos que se exponen pueden ser identificados y discriminados sobre una imagen. Sin embargo, se ha incluido alguno -por considerarse importante en una base topográfica o en su aplicación cartográfica posterior- que no cumple la afirmación anterior, lo cual quedará convenientemente explicado.

Con carácter general, deberá tenerse en cuenta lo siguiente:

- Mínima distancia teórica de captura de 5 m = 0,2 mm (tamaño mínimo que discrimina el ojo humano) x 25,000 (escala de referencia).
- Todas las entidades superficiales que se capturen deberán ir orientadas en sentido horario (sentido de las agujas del reloj).
- La unidad de medida utilizada para el registro espacial de los datos es el decímetro.
- La escala de visualización utilizada como escala de trabajo para la captura o actualización de elementos debe estar comprendida entre 1:2,000 - 1:7,000.

Los objetos geográficos que se exponen a continuación se organizan en un catálogo estructurado en los siguientes temas: Orografía y paisaje; Hidrografía; Edificios y construcciones; Servicios e instalaciones; Transporte. Dentro de estos grandes grupos o temas los objetos geográficos se han agrupado en función de dos factores: 1) tener similitudes en cuanto a su naturaleza o características, y 2) presentar normas de captura iguales o muy parecidas. Por tanto, no existe necesariamente una relación directa entre los objetos geográficos aquí presentados y las capas del modelo de datos físico que se diseñe particularmente en cada caso.

Orografía y paisaje

Espacio natural protegido

Definición: Zona que por su alto valor natural y cultural o carácter representativo merece una conservación preferente.

Geometría: Polígono.

Método de captura:

Los límites de las zonas naturales protegidas, cuando no coinciden con accidentes geográficos naturales del terreno (divisorias, ríos, línea de costa, etc.) o con divisiones administrativas, no es posible obtenerlos mediante la captura sobre una imagen. Por tanto, la información relativa a los límites de las zonas protegidas debe recabarse de los organismos oficiales competentes.

Curva de nivel

Definición: Línea imaginaria de altitud constante que sirve para describir la forma tridimensional de la superficie terrestre, incluyendo aquellas zonas sumergidas por masas de agua.

Geometría: Línea.

Método de Captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Normal: Caso genérico de curva de nivel.
- De depresión: Curva situada en una zona donde el terreno está hundido respecto a su entorno.
- Auxiliar: Curva con menor equidistancia de la establecida con carácter general.
- Batimétrica: La curva se encuentra en el fondo marino.

Igualmente es importante definir el origen o método de obtención del elemento.

- Restitución: obtenido mediante restitución fotogramétrica.
- MDT: obtenido mediante interpolación de un modelo digital del terreno.
- Hidrográfico: obtenido mediante técnicas batimétricas.

El registro de este objeto se realiza a partir de restitución fotogramétrica o interpolación MDT. Su representación es a través de elementos lineales continuos, sin cortes ni intersecciones entre ellas, salvo en el caso de curvas auxiliares que sí pueden quedar cortadas. Las curvas de nivel deberán ser consecutivas a la equidistancia debida. Para una escala 1:25,000 la equidistancia adecuada se establece en 10 m.

Si la actualización planimétrica de una hoja ya restituida se ha hecho mediante el uso de ortofotografías digitales, la actualización altimétrica de las zonas afectadas se calcula utilizando el modelo digital de elevaciones asociado a las mismas.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Población:
    - Densidad o altura de edificación baja: las curvas se trazan de forma continua, atravesando con especial cuidado en el paso por calles.
    - Densidad y altura de edificación elevada: para conservar la continuidad de las curvas será necesario imaginar el trazado de estas tomando las calles como referencia.
- Zonas de relieve escarpado (viseras, cortados) o sin visibilidad (tapadas por nubes): las curvas se capturan de modo continuo imaginando el trazado en estas zonas.
- Curvas auxiliares: cuando el terreno es muy llano y las curvas de nivel están separadas entre sí por distancias mayores a 100 m, se traza entre las curvas de nivel otras curvas auxiliares interpoladas (mitad de la equidistancia=5m). Si se cree conveniente pueden capturarse además puntos acotados.
- Curvas batimétricas:
    - No son equidistantes entre sí, aportando menor información batimétrica a medida que aumenta la profundidad.
    - Pueden obtenerse digitalizando a partir de cartas náuticas suministradas por el organismo oficial competente.
    - No tienen representación en el interior de las zonas emergidas.

Criterio de selección:

El área mínima que delimita una curva de nivel para ser capturada debe ser de 25x25m^2.

Ilustración 2- 5: Detalle de curvas de nivel cerca de un embalse

Punto de cota

Definición: Punto de altitud conocida que facilita la interpretación morfológica del terreno.

Geometría: Punto.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías que depende de la morfología del terreno donde se encuentre el punto acotado:

- Genérico: Punto situado en un terreno que no presenta ninguna particularidad morfológica.
- Cima: Punto de cota máxima en su entorno, rodeado de curvas de nivel de cota inferior. En una línea divisoria de aguas, las cimas representan sus puntos de cota máxima relativa.
- Collado-Puerto/Paso de montaña: Punto situado en una zona de doble curvatura, cóncava y convexa, con curvas superiores en una dirección e inferiores en la perpendicular (aproximadamente), por el que puede tener paso una vía de comunicación (puerto). En una línea divisoria de aguas, los collados representan sus puntos de cota mínima relativa.
- Hoya-depresión: Punto de cota mínima en su entorno, rodeado de curvas de nivel de cota superior.

Las entidades que se recogen para este objeto geográfico son elementos puntuales que representan puntos sobre el terreno que se consideran claves para la definición correcta de la

superficie terrestre, no debiendo existir puntos de cota en el interior de aguas continentales ni oceánicas.

En general, deben capturarse todos los puntos singulares del terreno: cimas, collados, hoyas, puntos distribuidos en los escarpados, entre las curvas cuando el terreno no sea uniforme, etc.

Debe ponerse especial cuidado en que la altitud del punto acotado sea acorde a la de las curvas de nivel en las que se encuentra. Es decir, la altitud del punto acotado que se capture debe estar comprobada de modo que su diferencia con la curva de nivel consecutiva no exceda en más o menos 10 m en todos los casos.

En las zonas de máximas depresiones del terreno (punto más bajo en curvas de depresión, fondo de lagunas no permanentes, etc.) debe materializarse su punto significativo mediante un punto de Hoya.

Criterio de selección:

La densidad de puntos acotados que deben recogerse varía según el tipo de terreno y su pendiente. Existe una gran cantidad de clasificaciones en función de la pendiente, muchas de ellas con un enfoque diferente, dependiendo del tipo de uso que se le va a dar al territorio en estudio.

Para la determinación de la pendiente puede utilizarse un mapa de pendientes generado a partir de los MDT con resolución de 5 m. Los intervalos de clasificación establecidos en este documento para efectuar la selección son los siguientes:

- Pendiente suave: <3%
- Pendiente media: 3-20%
- Pendiente acusada: >20%

En zonas llanas o de escasa pendiente, donde las curvas de nivel pueden ser poco definitorias del terreno, es conveniente aportar un mayor número de puntos de cota para definir mejor el relieve de la zona.

En zonas de crestas, donde el número de cerros y collados es grande, deben seleccionarse las altitudes máximas y collados destacables por los que circule una vía de comunicación.

Lugar

Definición: Sitio o elemento de interés natural o paisajístico.

Geometría: Punto.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Sitio: Nombre de lugar no habitado, tierras, bosques, depresiones, valles, llanos, etc.
- Bahía: Entrada natural del mar en la costa formando una concavidad o golfo.
- Saliente costero: Porción de tierra que penetra en el mar, cabo, punta.
- Playa: Ribera del mar o de un río grande, formada de arenales u otras partículas sueltas de mayor tamaño, en general de poca pendiente.

- Cueva: Cavidad subterránea que representa la entrada a una caverna sobre el terreno.
- Árbol: Planta que merece ser destacada por su tamaño o notoriedad.
- Mirador: Emplazamiento estratégico que cuenta con vistas a un paisaje o zona determinada.

Las entidades que se registran para este objeto geográfico son elementos puntuales que representan la posición más distintiva del elemento sobre el terreno.

La captura de algunos de estos elementos, como sitios, bahías, cuevas o árboles, es difícil de obtener empleando solamente la fotointerpretación sobre una imagen, por lo que debe realizarse apoyándose en otras fuentes de datos, por ejemplo, cartografía histórica.

Ilustración 2- 6: Detalle de captura playas, salientes costeros y bahías

Cadena montañosa

Definición: Conjunto de montañas que pueden forman una sierra, un macizo o sistema montañoso.

Geometría: Línea o polígono, según la información disponible.

Método de captura:

Las entidades que se recogen representan el eje principal de la entidad a la que representen. En el caso de sierras o cadenas montañosas el eje principal es la línea central de cumbres, en la que generalmente se encuentran los picos más elevados. En los casos de macizos y sistemas montañosos, el eje principal es una línea virtual simplificada que abarca los relieves que conforman el macizo o el sistema montañoso.

La captura de este objeto es difícil de obtener solo mediante fotointerpretación sobre una imagen, por lo que debe realizarse apoyándose en otras fuentes de datos, por ejemplo, cartografía histórica.

Ilustración 2- 7: Captura de cadena montañosa (destacada en azul) por su eje representativo principal

Vegetación

Definición: Masa de cultivos o vegetación natural.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías en función del tipo de cubierta vegetal:

- Cultivo
- Bosque
- Matorral
- Terreno despejado
- No se especifican tipos más concretos ya que dependen de la zona de trabajo y tienen un carácter marcadamente local.

La captura de este objeto entra en el campo de la fotointerpretación. No obstante, el operador deberá auxiliarse de un examen previo de la cartografía existente y de información exógena.

En esta capa pueden recogerse todos los usos de suelo característicos de un lugar. Se deben delimitar zonas claramente diferentes desde el punto de vista de uso del suelo.

Su registro se realizará mediante una superficie de modo que ésta quede claramente calificada con un tipo determinado, con las siguientes particularidades:

- Si la superficie de cultivo se encuentra delimitada por un detalle planimétrico (río, carretera, cerca, etc.) se harán coincidir las dos líneas.
- Los recintos de cultivos creados no deben superponerse ni cruzarse unos con otros.
- Si dos o más zonas del mismo cultivo se encuentran separadas entre sí una distancia menor de 12 m, se agruparán para obtener una zona común.

Criterio de selección:

Una zona se considera representable cuando su superficie sea mayor de 4 hectáreas o bien cuando sea de gran importancia resaltarla, aunque no sea representable a escala.

Cortafuegos

Definición: Vereda o zanja ancha que se deja sin vegetación en los montes para que no se propaguen los incendios.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías definidas según su anchura:

- Estrecho: Eje de cortafuegos de anchura inferior a 15 m.
- Ancho: Eje de cortafuegos de anchura superior a 15 m.

Las entidades que se capturan para este objeto geográfico son elementos lineales continuos representados por su eje.

Cuando el trazado sigue una senda, camino o línea eléctrica, los ejes representativos de ambas entidades deben de ser coincidentes.

Hidrografía

Curso natural

Definición: Corriente de agua natural más o menos caudalosa, que recoge los aportes de una cuenca fluvial y desemboca en otra corriente, en un lago o en el mar.

Geometría: Línea y polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías o características:

- Persistencia (aplica a elementos lineales): Régimen de aguas según la presencia de agua durante el año.
    - Permanente: El tramo lleva agua todo el año o la mayor parte del mismo.
    - Estacional o temporal: El tramo lleva agua en momentos determinados del año de modo regular asociado a las estaciones.
    - Esporádico: El tramo lleva agua raramente y de modo irregular a lo largo del año.
- Componente (aplica a elementos lineales): Naturaleza del tramo del cauce.
    - Genérico: tramo de cauce que discurre por el terreno sin ninguna característica reseñable.
    - Canalizado: tramo en el que se han realizado obras importantes para encauzar o desviar la corriente de agua. No se considera canalizado las obras de fábrica que se hacen para cruzar vías de comunicación.

    - Cubierto: Tramo cubierto al atravesar una edificación, una población, una presa o una vía de comunicación.

- Nivel de agua (aplica a elementos poligonales): Cauce aparente del tramo de curso natural.
    - Agua aparente: Nivel de agua en el instante de la captura.
    - Cauce habitual aparente: Cauce más habitual identificado por la huella que deja el agua en el terreno. Se corresponde con tramos de crecidas o avenidas claramente identificables, mediante fotointerpretación, en el terreno.
- Identificador: Código que identifica de manera única cada curso natural de agua desde su nacimiento hasta su desembocadura. Puede obtenerse de un organismo oficial competente o ser definido dentro de la base de datos con unas reglas establecidas a priori.

Las entidades que se capturan para este objeto geográfico son elementos lineales que representan el eje de un tramo de corriente de agua natural. Cuando el ancho del cauce sea superior a 5 m, además del eje, se ha de capturar un polígono que represente la anchura del cauce en ese tramo. La suma de las distintas masas superficiales de un río, que no tienen que ser contiguas, junto con su representación lineal, posibilita la creación del elemento único para los cursos fluviales. Tiene las siguientes características básicas:

- Continuo desde el nacimiento hasta la desembocadura en el mar o en otro curso de orden superior.
- El sentido de registro es aguas abajo.

Para asegurar la continuidad y poder generar un modelo de red hidrográfica válido para el tratamiento hidrológico y geográfico se debe garantizar:

- La conectividad de todos los ejes que forman un mismo cauce natural, aunque discurran por el interior de masas de agua como embalses, lagunas o polígonos de cauces naturales o artificiales.
- La conectividad de los polígonos con la red lineal o con otras masas de agua.
- La conectividad de un cauce con otro -cuando la desembocadura es en otro río-, o con la línea de costa -cuando la desembocadura es en el mar-. Esta condición es aplicable tanto a los ejes como a los polígonos.
- La asignación del mismo identificador a todos los tramos -ejes y polígonos- que forman un mismo cauce natural.
- Excepcionalmente, el cauce natural puede quedar interrumpido por ser endorreico o por desconocerse su trazado por diversos motivos. En estos casos se coloca un sumidero (ver objeto *Punto hidrográfico. Sumidero*) en el vértice extremo del eje del río cuando este desaparece.

La captura ha de ser conforme con la descripción altimétrica del terreno. En este sentido, debe discurrir por la concavidad formada por las curvas de nivel con una diferencia máxima de altitud de 2,5 m en su cruce con éstas, no debiendo tener subidas ni bajadas significativas de cota (H) que no sean las reales del terreno (cascada, salto, etc.).

Como ayuda para determinar la persistencia o régimen de aguas, es conveniente consultar la fecha de vuelo y la cartografía existente. Si el cauce aparece seco en época de lluvias es indicativo de un régimen estacional o esporádico, y si aparece seco en temporada seca, conviene consultar la cartografía existente histórica. En general, los cauces estacionales y esporádicos son cursos de agua de caudal escaso y cauce pequeño, y suelen desembocar en otro río.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Tramos cubiertos: No tienen representación superficial, solo lineal. Se han de intentar capturar siempre, aunque no sea de forma precisa, para dar continuidad al cauce. Tanto si la captura se hace mediante restitución o mediante ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo.
- Presas o diques (ver objeto *Presa*) a lo largo de su recorrido: Deben representarse siempre.
- Islas (ver objeto *Isla*) dentro de un tramo de río con representación superficial: El eje se llevará por un solo brazo de la isla.
- Masas de agua: Los ejes de los cauces deben discurrir por su interior y conectarse al cauce principal, no debiendo existir más de un cruce entre un cauce y lámina de agua (excepto el principal que tiene dos, al entrar y salir de la lámina de agua).
- Quebradas: se representan como cursos de agua no permanentes.

Criterio de selección:

Como norma general al seleccionar la hidrografía, hay que poner especial cuidado en no dar un efecto erróneo de mucha hidrografía en zonas donde hay numerosos arroyos y generalizar de acuerdo con la zona de trabajo en la que nos encontremos.

Para valorar este factor pueden emplearse criterios de aridez, que permiten medir el déficit pluviométrico del territorio (húmedo, semiárido y árido). Esta clasificación permite seleccionar determinados objetos en función del régimen térmico y de humedad de cada país y de la región dónde se encuentre la zona en la que se esté trabajando.

- Zonas húmedas: Deben omitirse aquellos cursos de agua cuya longitud sea inferior a 500 m.
- Zonas áridas: Deben capturarse todos los cursos de agua con nombre propio. Si el terreno tiene muchos cursos no permanentes debidos al efecto de erosión de las aguas, solo se registrarán los que tengan longitud superior a 2000 m, para evitar el efecto erróneo de hidrografía abundante en zonas donde apenas existe.

Hay disponibles varios productos a nivel mundial, como el Atlas Mundial de Desertificación, que puede ser de utilidad, aunque es conveniente consultar esta información a las autoridades oficiales competentes.

Hay casos particulares en los que se puede atender, además, a los siguientes criterios de selección:

- Tramos cubiertos al atravesar la red de transporte: Solo se capturan cuando son > 3 m, los < 3 m se capturan continuos sin cortar.

- Vaguadas: Son ondulaciones sin hendiduras en el terreno, producidas por la lluvia y no se representan ya que vienen indicadas por las curvas de nivel.
- Quebradas: Solo se capturan si desembocan en otro río y su longitud es > 500 m.
- En casos en que se dude si representar o no un curso de agua puede recurrirse a la cartografía existente y registrar solo los que aparezcan en él.

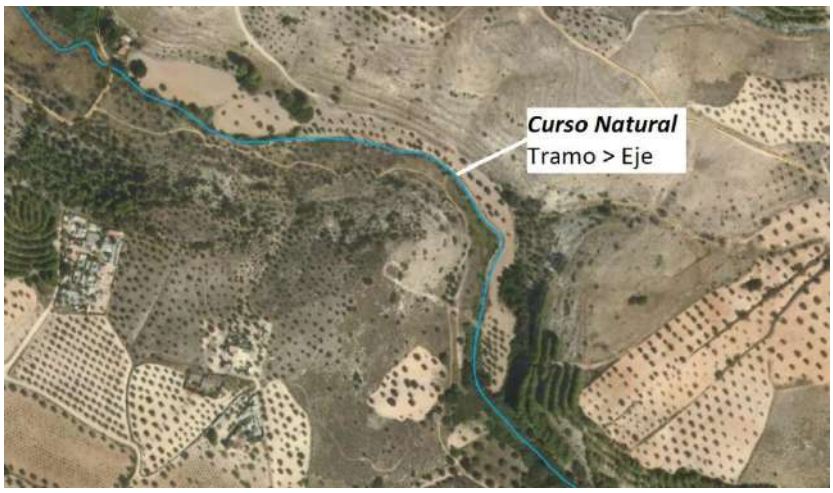


Ilustración 2- 8: Captura de curso natural solo como eje (ancho <5m)

Ilustración 2- 9: Curso natural canalizado de ancho <5m

Ilustración 2- 10: Conexión en un sólo brazo de isla fluvial

Ilustración 2- 11: Eje dentro de embalse

Ilustración 2- 12: Conexión dentro de tramo superficial (ancho >5m)

Ilustración 2- 13: Tramo de río cubierto al atravesar una presa

Ilustración 2- 14: Tramo de río cubierto al atravesar una población

Ilustración 2- 15: Tramos de río cubierto al atravesar carreteras

Ilustración 2- 16: Río con representación superficial que desemboca en el mar

Ilustración 2- 17: Agua aparente (lleva agua en el momento de la captura)

Ilustración 2- 18: Cauce habitual aparente (no lleva agua en el momento de la captura y las crecidas son identificables en el terreno)

Curso artificial

Definición: Obra realizada para conducir una corriente de agua para distintos usos.

Geometría: Línea y polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías de cauce artificial:

- Canal: Tramo abierto que conduce agua con fines de abastecimiento, riego, navegación, trasvase, etc.
- Acequia: Tramo abierto principalmente destinado a la conducción de agua para riego.
- Tubería: Tramo de cauce cerrado en una tubo u otro tipo de obra de fábrica o conducto.

Igualmente son importantes las siguientes características del objeto:

- Estado físico: Estado de uso en el que se encuentra el elemento.
    - En uso: Se encuentra utilizándose o en condiciones aparentes de buen estado.
    - En construcción: Se encuentra en obra.
    - En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.
- Situación: Localización respecto de la superficie terrestre.
    - En superficie: Se encuentra sobre la superficie terrestre.
    - Elevado: Se encuentra elevado respecto a la superficie terrestre.
    - Subterráneo: Se encuentra bajo la superficie terrestre.
- Identificador: Código que identifica de manera única cada curso artificial de agua. Puede obtenerse de un organismo oficial competente o ser definido dentro de la base de datos con unas reglas establecidas a priori.

Las entidades que se capturan para este objeto geográfico son elementos lineales que representan el eje de un tramo de corriente de agua artificial. Cuando el ancho del cauce sea superior a 5 m, además del eje, se ha de capturar un polígono que represente la anchura del cauce artificial en ese tramo. La suma de las masas superficiales de un cauce artificial junto con su representación lineal posibilita la creación del elemento único para los cauces artificiales. Tiene las siguientes características básicas:

- Continuo, aunque puede resultar mucho más complicado que en el caso de los cauces naturales.
- Sentido de registro aguas abajo. Generalmente se puede determinar a pesar de ser sensiblemente horizontales, con ayuda de la información hidrográfica cercana.
- No debe tener subidas ni bajadas significativas de cota (Altura - H) que no sean las reales del terreno.

Para asegurar la continuidad y poder generar un modelo de red hidrográfica válido para el tratamiento hidrológico y geográfico se debe garantizar:

- La conectividad de todos los ejes que forman un mismo cauce artificial, aunque discurran por el interior de masas de agua como embalses, lagunas o polígonos de cauces naturales o artificiales.
- La conectividad de los polígonos con la red lineal o con otras masas de agua.
- La conectividad de un cauce artificial con otro cauce artificial o natural o con la línea de costa. Esta condición es aplicable tanto a los ejes como a los polígonos.
- La asignación del mismo identificador a todos los tramos -ejes y polígonos- que forman un mismo cauce artificial.
- Como se ha comentado anteriormente, es frecuente que el trazado pueda quedar interrumpido por la dificultad de seguir el trazado del curso artificial. En estos casos se coloca un sumidero (ver objeto *Punto hidrográfico. Sumidero*) en el vértice extremo del eje del cauce cuando este desaparece.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Canales: Como norma general, se pueden clasificar como canales los cauces artificiales de anchura superior a 2 m. Se reconocen también por su pequeña pendiente y trazado suave y rectilíneo, y porque parten siempre de ríos, embalses o lagos. Si en la imagen aparecen secos, es conveniente marcarlos para determinar a posteriori su posible abandono.
- Acequias: De menor categoría que los canales; como norma general, se pueden clasificar como acequias los cauces artificiales de anchura inferior a 2 m.
- Tramos elevados: Si la captura se hace mediante restitución, la altura (H) queda perfectamente definida ya que el tramo es visible. Si el registro se realiza con ortofotografía+MDT, debe interpolarse entre la altura de los vértices inicial y final de dicho tramo.
- Tramos subterráneos: No tienen representación superficial, solo lineal. Se han de intentar capturar siempre, aunque no sea de forma precisa, para dar continuidad al cauce. Si no se puede dar esta continuidad, se coloca un sumidero (ver objeto *Punto hidrográfico. Sumidero*) en el punto donde se pierda su trazado. Tanto si el registro se hace mediante restitución como por captura sobre ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo.

Criterio de selección:

En los siguientes casos se atenderá a su criterio específico:

- Canales: Se capturan todos.
- Acequias, drenajes:
    - Zona de alta densidad: Los que tengan una longitud menor de 1 Km no se representan, a excepción de los que comiencen, terminen o se hallen a una distancia máxima de 20 m de un detalle planimétrico de carácter hidrográfico o de construcción hidrográfica (estanque, alberca, molino, sifón, etc.).
    - Zona de baja densidad: Tendrán representación todos los visibles.

- Tramos subterráneos: Solo se capturan cuando su longitud es mayor de 3 m; los menores de 3 m se capturan continuos como cauce en superficie.

Ilustración 2- 19: Canal en superficie

Ilustración 2- 20: Acequia en superficie

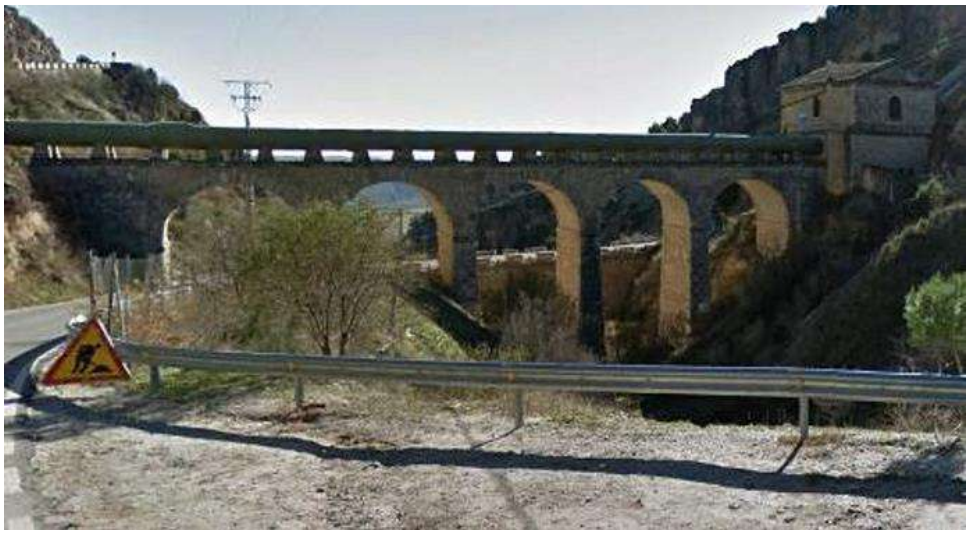


Ilustración 2- 21: Tubería elevado

Ilustración 2- 22: Tramo subterráneo que cruza una vía

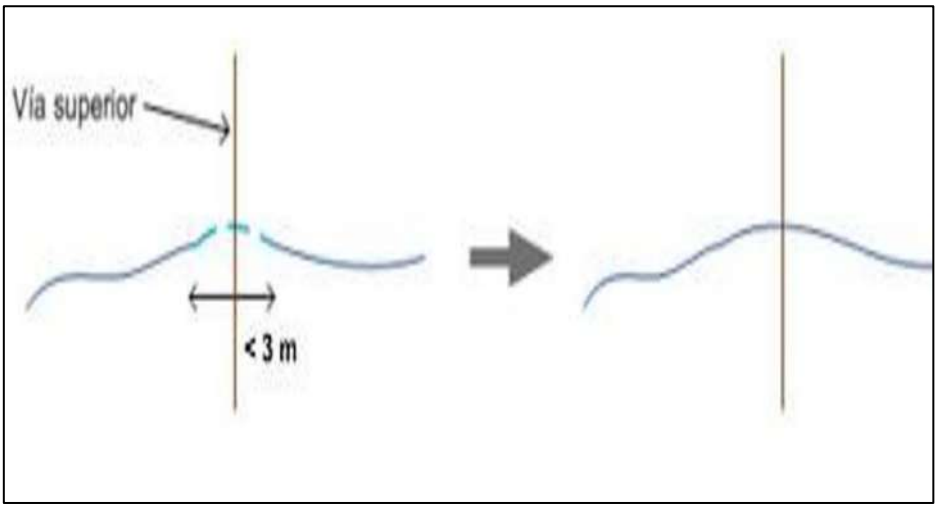


Ilustración 2- 23: Tramo subterráneo no capturado por ser < 3m

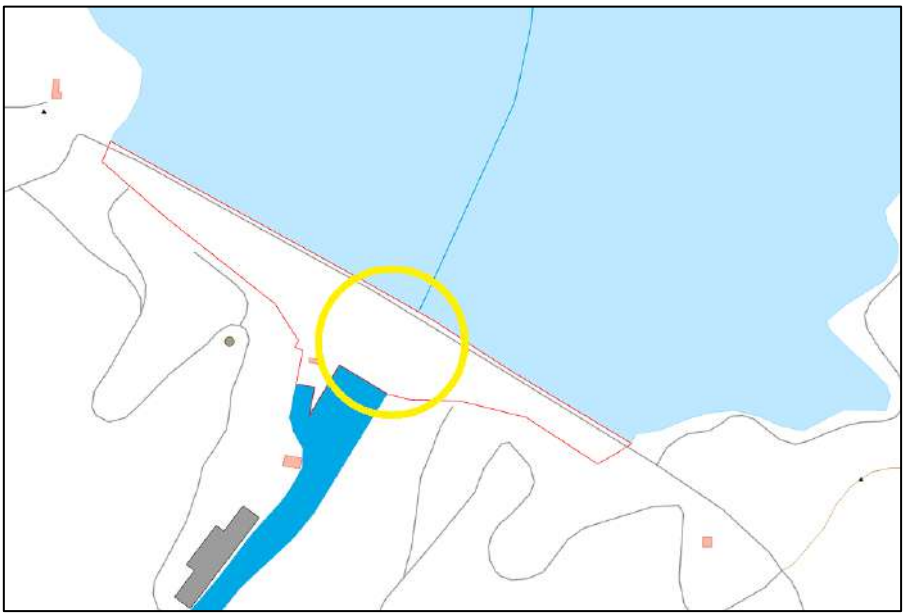


Ilustración 2- 24: Cauce artificial no conectado con embalse – Captura incorrecta

Aguas marinas

Definición: Masa de agua exterior a la línea de costa.

Geometría: Polígono.

Método de captura:

Masa de agua superficial capturada por su contorno, que será coincidente con la línea de costa -en la parte que limita con la isla- y aproximado en los límites con otras masas de agua.

Ilustración 2- 25: Captura de un polígono de aguas marinas coincidente con la línea de costa

Costa

Definición: Límite entre la zona terrestre y la zona marítima que coincide con la curva de nivel de cota cero.

Geometría: Línea.

Método de captura:

La línea de costa es un elemento lineal continuo y debe registrarse sin interrupciones. Si existen acantilados o playas, las líneas que delimitan estos elementos no interceptarán la línea de costa, llegarán hasta ella sin cortarla. Los cursos de agua que desembocan en el mar, representados por sus márgenes o por el eje, deben de coincidir geométricamente con los puntos de la línea de costa, sin cortarla.

Generalmente en las condiciones técnicas de vuelo debe especificarse que las tomas se hagan en bajamar para facilitar la captura de la línea de costa coincidente con la curva de nivel cero en las zonas naturales. Si el vuelo o la imagen satelital fue tomada en pleamar y no puede trazarse la curva de nivel de cota cero, debe registrarse la línea tierra-agua, con una nota para que se corrija en fases posteriores.

En zonas artificiales (muelles, , espigones, rompeolas, terreno ganado al mar, etc.) la línea de costa se captura coincidente con la infraestructura.

Si la línea de costa se determina a partir de un MDT, puede ofrecer cierta incertidumbre y ser necesaria su generalización. También puede haber incertidumbre por falta de coincidencia entre la línea de costa y la línea tierra-mar visible en la ortofoto. Para determinar la curva de cota cero se utilizan fuentes auxiliares de información como líneas de pleamar-bajamar (de organismos oficiales competentes), restituciones anteriores y línea de cota cero interpolada del MDT.

Ilustración 2- 26: Línea de costa natural

Ilustración 2- 27: Línea de costa artificial en un puerto

Ilustración 2- 28: Curva de cota 0 obtenida con MDT (morado) y línea de costa en cota 0 resultante (azul)

Ilustración 2- 29: Línea de costa en cota 0, desplazada de la línea tierra-mar visible en la ortofoto

Laguna-lago

Definición: Masa de agua acumulada en una depresión natural del terreno.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes características en función del régimen de aguas existente durante el año:

- Permanente: La lámina presenta agua todo el año o la mayor parte del mismo.

- Estacional: La lámina presenta agua en momentos determinados del año de modo regular asociado a las estaciones.
- Esporádico: La lámina de agua se manifiesta raramente a lo largo del año.

El registro se realizará mediante una superficie siguiendo el contorno del nivel de agua, observado en la imagen, en su máxima capacidad o de vegetación fija, y de altura constante.

Si el registro se hace por restitución se establece la altura (H) probable de la laguna y se captura manteniendo la altura constante. Si se hace mediante ortofotografía+MDT, se registra planimétricamente y se impone una altura (H) probable fija a los puntos capturados.

Como ayuda para determinar la persistencia o régimen de aguas, es conveniente consultar la fecha del vuelo y la cartografía existente en caso de duda.

Cuando la laguna está seca debe darse un punto de hoya en el punto más bajo, no siendo necesario si la laguna tiene agua.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Ejes de cauces: Deben atravesar las lagunas para dar continuidad al elemento hidrográfico (ver *Curso natural* y *Curso artificial*).
- Presas o diques: Si existen, se deben ajustar geométricamente al contorno de la laguna.
- Balsas de riego: Pueden capturarse como dos tipos de objeto geográfico.
    - Como balsa-alberca-estanque: Cuando es de construcción totalmente artificial.
    - Como laguna-lago (ver *Laguna-lago*): Cuando es una depresión natural cerrada por una presa o dique.

Ilustración 2- 30: Presa no ajustada a laguna y masas de agua desconectadas - Captura incorrecta

Criterio de selección:

Para seleccionar los objetos a capturar hay que generalizar de acuerdo a la zona en la que se encuentre la zona de trabajo, según el criterio de aridez establecido (ver objeto *Curso natural*).

- Zonas húmedas: Deben capturarse todas las lagunas (permanentes o no) que tengan nombre propio y omitir aquellas cuya superficie sea menor de 100 m^2 y no tengan nombre propio.
- Zonas áridas: Deben capturarse todas las lagunas permanentes y aquellas no permanentes cuya superficie sea mayor de 100 m^2 o menor de 100 m^2 con nombre propio.

Humedal

Definición: Terreno bajo y pantanoso que se inunda por las aguas del mar durante las mareas o por las de los ríos en sus crecidas. Son zonas inundables, de naturaleza irregular, cambiante y dinámica, y de una gran diversidad.

Geometría: Polígono.

Método de captura:

El registro se realizará mediante una superficie que delimite el contorno del humedal. En su interior deben representarse todos los detalles planimétricos que aparezcan.

Embalse

Definición: Obra hidráulica que genera una masa de agua acumulada artificialmente en el transcurso de un cauce mediante una presa, para su aprovechamiento en regadíos, producción de energía, abastecimiento, etc.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes características en función del nivel de la lámina de agua del embalse:

- Máxima elevación: Nivel máximo alcanzable por el agua en condiciones normales de aprovechamiento, para aquellos embalses cuya presa esté en explotación.
- En carga: Nivel máximo alcanzable por el agua en condiciones normales de explotación, para aquellos embalses cuya presa esté en construcción o en puesta en carga.
- Máximo aparente: Nivel máximo aparente obtenido por fotointerpretación sobre las imágenes de los últimos 10 años, para aquellos embalses cuya presa esté en explotación.

Las entidades que se registran para este objeto geográfico son elementos superficiales, de altura constante, que representan el contorno del embalse hasta llegar a su cierre en la presa.

Si el registro se hace por restitución, se establece la altura (H) del embalse y se captura manteniendo la altura constante, como si fuera una curva de nivel con dicha cota. Si se hace mediante ortofotografía+MDT, se obtiene por la generación de la línea de dicha altura del MDT.

Los embalses se capturan siempre a nivel de máxima elevación o cota de la lámina de agua máxima en las condiciones normales de explotación, es decir, la máxima elevación es independientemente del nivel del agua que en ese momento aparezca en la imagen. El valor de máxima elevación es recomendable obtenerlo de la autoridad competente. Si no se dispusiera de esta información, a la altitud de la restitución o MDT en la coronación de la presa se le restan 2 m para obtener la línea de máximo embalse.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Máximo aparente: En aquellos embalses en los que el nivel máximo proyectado entre en conflicto con infraestructuras o viales, además de la superficie de máxima elevación, puede capturarse también una superficie al nivel máximo aparente de agua. Esta superficie se captura en base a la fotointerpretación sobre imágenes de los últimos 10 años de la marca más elevada dejada por el agua en el terreno. La superficie de máximo aparente siempre va contenida en la superficie de máximo embalse.
- Elementos lineales de red hidrográfica: Para garantizar la consistencia entre el embalse y los ejes de los cauces, éstos últimos deben encontrarse dentro de sus límites y conectados al eje del cauce principal, no debiendo existir más de un cruce entre un cauce y la superficie embalsada (excepto el principal que tiene dos cruces, uno de entrada y otro de salida).
- Presas y otras masas de agua: Deben ser coincidentes geométricamente con el cierre del embalse.

Criterio de selección:

Deben tener representación todos los embalses independientemente del volumen de agua embalsada.

En los siguientes casos se atenderá además a los siguientes criterios de selección:

- Detalles planimétricos por debajo de la línea de máximo embalse: Por lo general, no deben capturarse, salvo que se encuentren en la franja de terreno entre el máximo aparente y la máxima elevación.
- Detalles altimétricos: No se representan por debajo de la línea de máximo embalse, y por encima la primera curva representable será la de altitud múltiplo de 10 más próxima y superior a ese valor.
- Islas: Solo deben tener representación aquellas cuya altitud sea superior a la cota de máximo embalse y tengan una superficie mínima de 100 m^2.

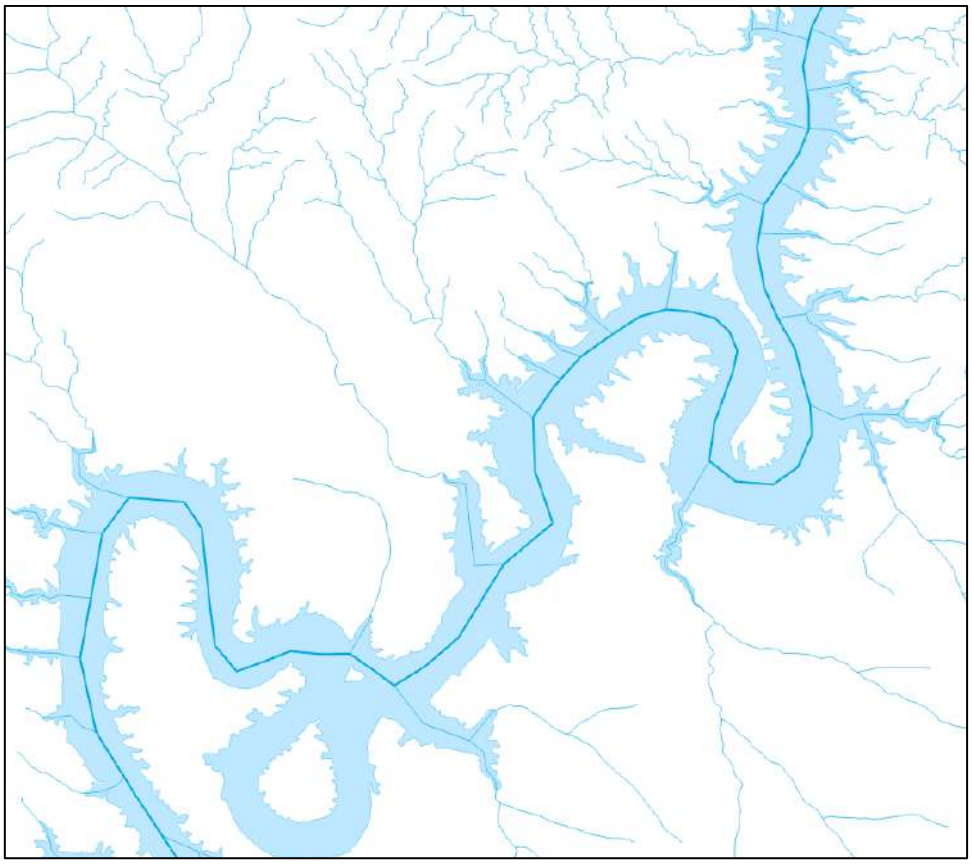


Ilustración 2- 31: Embalse con los ríos lineales conectados en su interior

Ilustración 2- 32: Geometrías coincidentes en los contornos de embalse, presa y río

Punto hidrográfico

Definición: Lugar donde existe un punto de interés hidrográfico, natural o artificial.

Geometría: Punto.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías según su naturaleza o uso:

- Fuente: Lugar natural o artificial de agua (potable o no) acondicionado con una obra, cuya finalidad es la poder coger agua para distintos usos.
- Manantial: Lugar donde nace o brota de forma natural el agua.
- Pozo: Perforación en el terreno para extraer agua subterránea que puede venir acompañada de una obra de fábrica en superficie.
- Cascada-Salto/poza: Caída del agua desde cierta altura / concavidad de cierta profundidad originada por dicha caída de agua en el cauce.
- Sumidero: Lugar de entrada de un cauce natural o artificial de agua en el terreno cuando éste pasa a ser subterráneo y no se conoce su trazado ni dónde vuelve a surgir. En general, lugar en que se pierde el trazado de un cauce natural o artificial.

Como norma general las entidades que se recogen para este objeto geográfico se capturan como elementos puntuales indicativos del lugar donde está ubicado el punto hidrográfico, que son visibles en la imagen o con ayuda de la fotointerpretación y de la cartografía existente.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Cascada-salto-poza: Debe estar siempre situada sobre una corriente de agua, formando parte de ella, coincidiendo exactamente con un punto del cauce.
- Sumidero: Debe situarse en el vértice extremo donde se pierde el trazado del cauce, por tratarse de un tramo subterráneo que no es posible capturar, aunque sea de forma aproximada, o por cualquier otro motivo. En los tramos subterráneos en los que sea posible dar continuidad al cauce, no se registrarán sumideros en los puntos de entrada o salida. Las razones principales por las que no se puede dar continuidad a un cauce son:
    - Se trata de un cauce natural endorreico.
    - No se puede capturar el tramo subterráneo porque no se localiza el punto de entrada o salida y, por tanto, no es posible dar continuidad al cauce.
    - Se pierde el trazado en una zona muy llana, generalmente en zonas de cultivo.

Ilustración 2- 33: Sumideros en ríos en los que se pierde el trazado en una zona muy llana

Isla

Definición: Porción de tierra rodeada enteramente de agua, por el mar o por un lago o río.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías según su ubicación:

- Marítima: La isla se encuentra situada en el mar.
- Aguas interiores: La isla se encuentra situada en un recinto de agua situado tierra adentro.

Las entidades que se recogen para este objeto geográfico son elementos superficiales capturados siguiendo el contorno definido por la línea tierra-agua.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Islas en embalses o lagunas: La altura (H) de los puntos de su contorno debe ser constante y acorde a la máxima elevación del embalse o laguna.
- Islas marítimas: Deben coincidir exactamente con la línea de costa y con la curva de nivel de cota cero.

Criterio de selección:

Con carácter general, deben capturarse todas las islas, tanto interiores como marítimas, siempre que tengan una superficie superior a 100 m^2.

Ilustración 2- 34: Captura de islas marítimas

Ilustración 2- 35: Captura de islas fluviales (aguas interiores)

Presa

Definición: Construcción interpuesta transversalmente en un punto de un curso fluvial o cauce artificial con el objeto de almacenar el agua del mismo, derivarla o regular su curso fuera del cauce.

Geometría: Línea o Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías según sus características constructivas:

- Presa de hormigón: Material empleado para la construcción de las presas de tipo gravedad, arco y contrafuerte.
- Presa de materiales sueltos: Técnica empleada en la construcción de presas en la que se emplean rocas o tierras sueltas sin mortero.

Las entidades que se recogen para este objeto geográfico son elementos lineales registrados por el eje longitudinal de la construcción cuando su anchura es inferior a 10 m o como un elemento superficial cuando la anchura (teniendo en cuenta la caída) es igual o mayor de 10m. En este último caso, su representación debe realizarse por la línea de su contorno, que se corresponde con la intersección de la obra con el terreno, debiendo quedar materializada claramente su estructura.

Cuando la captura se hace mediante ortofotografía+MDT, la determinación de la altura (H) estará condicionada a la naturaleza del MDT utilizado, por lo que habrá que tener en cuenta las indeterminaciones que causan las interpolaciones de H entre dos puntos de la malla que compone el MDT y será necesario una cierta interpretación de la misma.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Vías de comunicación que atraviesan la presa: No deben quedar interrumpidas, debe continuarse su trazado por encima de la presa.
- Masas de agua: Deben ser coincidentes geométricamente con el contorno de la presa.
- Pequeñas Barreras y azudes transversales a la corriente de un río cuya labor sea la de retener, ralentizar o embalsar agua: Deben capturarse como presa, no como muro de contención (ver objeto *Cerramiento*).
- Diques de contención en balsas-albercas y lagunas: Deben capturarse como presa, no como cerramiento (ver objeto *Cerramiento*).

Criterio de selección:

Deben capturarse todas los que sean visibles.

Ilustración 2- 36: Presa de materiales sueltos capturada por su contorno

Ilustración 2- 37: Presa de hormigón capturada por su contorno

Ilustración 2- 38: Dique de contención capturado como presa

Ilustración 2- 39: Represa en un cauce natural

Servicios e instalaciones

Espacio dotacional

Definición: Recinto que se destina a usos, instalaciones o servicios determinados, ya sean públicos o privados, y que repercuten en los ciudadanos.

Geometría: Punto o Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías. No obstante, la inclusión de unos tipos u otros de espacios dotacionales depende del universo de discurso planteado:

- Explotación minera: Lugar destinado a la extracción y aprovechamiento de algún tipo de mineral o recurso, subterráneo o a cielo abierto. Puede resultar interesante recoger la sustancia que se extrae y el estado de actividad en el que se encuentra la explotación minera (activo, sin explotación, abandonado, histórico, etc.).
- Zona agropecuaria: Zona en la que se lleva a cabo la producción de productos vegetales y animales para fines de alimentación, de venta, de autoconsumo o industriales.
- Salina: Lugar o instalaciones donde se obtiene sal por evaporación de aguas salinas.
- Piscifactoría: Instalación para la cría artificial de peces y mariscos.
- Central térmica: Instalación para obtener energía eléctrica a partir de la combustión de distintas sustancias, como combustibles fósiles, biomasa, etc.
- Central nuclear: Instalación para obtener energía eléctrica a partir de la fisión nuclear de elementos radiactivos.
- Central hidroeléctrica: Instalación para obtener energía eléctrica a partir de la fuerza cinética de una corriente de agua.
- Central solar: Instalación para obtener energía eléctrica a partir de la radiación solar.
- Parque eólico: Instalación donde se lleva a cabo la producción de energía eléctrica mediante aerogeneradores, muchas veces en agrupaciones de gran tamaño.
- Subestación eléctrica: Instalación provista de un conjunto equipos dedicados a reducir o aumentar el voltaje de la corriente eléctrica.
- Instalación de hidrocarburos: Instalación destinada a la producción, tratamiento o distribución de hidrocarburos. No incluye gasolineras u otros puntos de venta orientados al público general.
- Zona industrial: Zona urbanística donde se agrupan diferentes edificios, destinados a uso industrial, ubicados generalmente en las afueras de las poblaciones.
- Recinto de ferias y congresos: Instalaciones para la organización de ferias, salones, congresos u otro tipo de eventos orientados, principalmente, a las relaciones empresariales y al ocio.
- Zona comercial: Agrupación de edificios o instalaciones situados cerca de una población cuyo fin es la venta de diferentes productos de consumo, como mercados, centros comerciales, etc.

- Hospital: Instalaciones destinadas al diagnóstico y tratamiento de enfermos donde, a menudo, se practican la investigación y la docencia.
- Campus universitario: Conjunto de terrenos, edificios e instalaciones asociadas a un centro universitario.
- Zona militar: Recinto en el que se desarrollan actividades militares.
- Centro penitenciario: Edificio o conjunto de edificios e instalaciones destinadas a la reclusión de presos.
- Cementerio: Recinto, generalmente cercado, destinado a enterramientos.
- Observatorio espacial: Instalaciones para la observación de planetas, estrellas, galaxias y otros cuerpos celestes ya sea desde la superficie terrestre o desde el propio espacio a través de satélites.
- Castillo-fortaleza: Recinto defensivo, generalmente aislado y situado en alguna eminencia del terreno, que suele estar rodeado por una muralla.
- Yacimiento arqueológico: Recinto donde se encuentran restos arqueológicos o excavaciones de interés histórico-cultural o restos materiales indicativos de vida humana o animal ya desaparecida.
- Parque temático y ocio: Recinto con un conjunto de atracciones, espacios para el ocio, entretenimiento, educación y cultura, normalmente organizados en torno a una línea argumental que les sirve de atracción. Se incluyen aquí también los parques de atracciones, acuáticos, planetarios, zoológicos, etc.
- Parque-jardín: Recinto o área en el interior de una población destinado a prados, jardines y arbolado para recreo y ornato. Son fácilmente identificables por ser recintos con grupos de árboles, setos, césped y amplias zonas de recreo.
- Área recreativa: Zona pública de fácil acceso acondicionada para proporcionar servicios básicos para el uso recreativo de los espacios naturales. Suele contar con zona de estacionamiento, mesas y asientos, fuentes, kioscos, etc.
- Camping: Recinto al aire libre para la pernoctación segura de personas al aire libre, en tienda de campaña, autocaravana o bungaló.
- Zona deportiva: Recinto o área especialmente diseñada y construida para la práctica de deportes diversos.
- Campo de golf: Recinto al aire libre, acondicionado y equipado especialmente para practicar el deporte del golf.
- Circuito de velocidad: Zona destinada a carreras de velocidad (automóviles, motos, karts, etc.) o a enseñanza vial.
- Plantas de Tratamiento: Instalación para el tratamiento de las aguas para el consumo humano o para reducir su contaminación.
- Gestión de residuos: Instalación destinada al acopio, almacenaje, selección, tratamiento o recuperación de basura, residuos materiales de desecho o escorias de diferente tipo.

Como norma general se usará siempre la representación superficial cuando la instalación sea extensa y agrupe varias edificaciones, reservándose la puntual para entidades muy pequeñas, que se ubiquen en una sola edificación, o que no le sea aplicable la representación superficial.

Si el registro se realiza mediante un polígono, éste debe definir, de la forma más precisa posible, el borde o contorno exterior de la instalación. Si la captura se realiza mediante un punto, éste debe situarse en su lugar más representativo. En ambos casos los objetos interiores de la instalación deben registrarse siguiendo sus normas de captura específicas.

Estas instalaciones suelen estar delimitadas por algún tipo de cerramiento (cerca, muro, alambrada, muralla, etc.) que debe capturarse coincidente con el polígono que la define.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Explotaciones mineras a cielo abierto: Se dará un punto de hoya en el punto o puntos de menor cota, dejando cortados en su borde los caminos de acceso e interiores a la explotación si se aprecia que no son permanentes.
- Parques eólicos: Deben aparecer representados todos y cada uno de los aerogeneradores eólicos que contiene.

Ilustración 2- 40: Ejemplo de captura para una mina a cielo abierto

Ilustración 2- 41: Ejemplo de captura para una salina

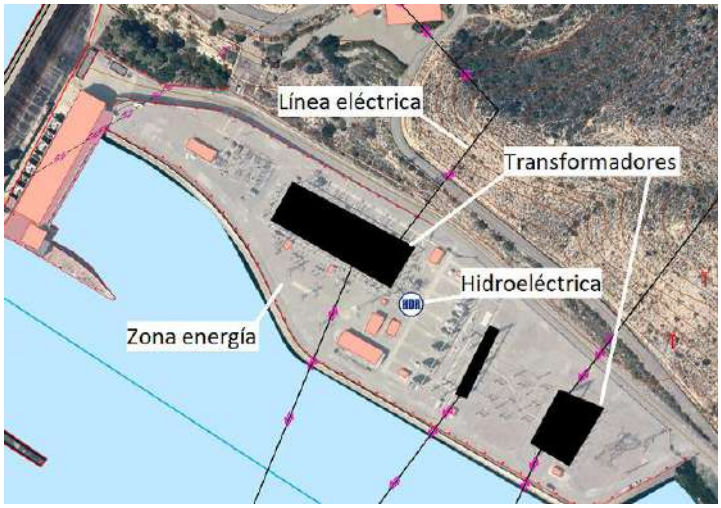


Ilustración 2- 42: Ejemplo de captura para una central hidroeléctrica

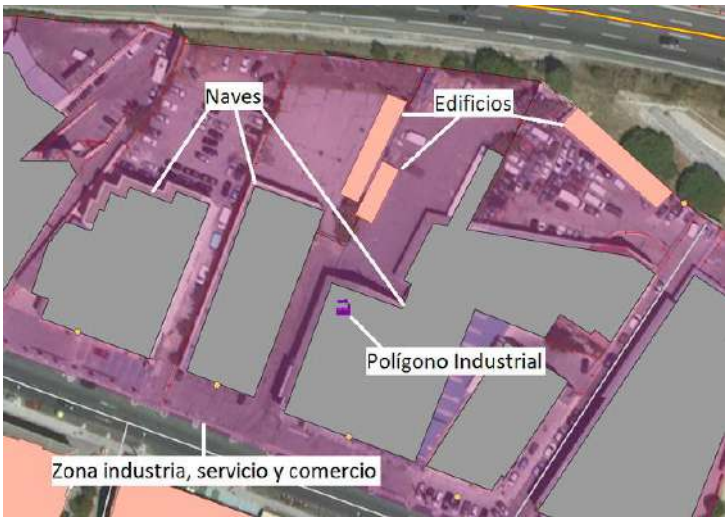


Ilustración 2- 43: Ejemplo de captura para una zona industrial

Ilustración 2- 44: Captura de un área recreativa con una alambrada como cerramiento

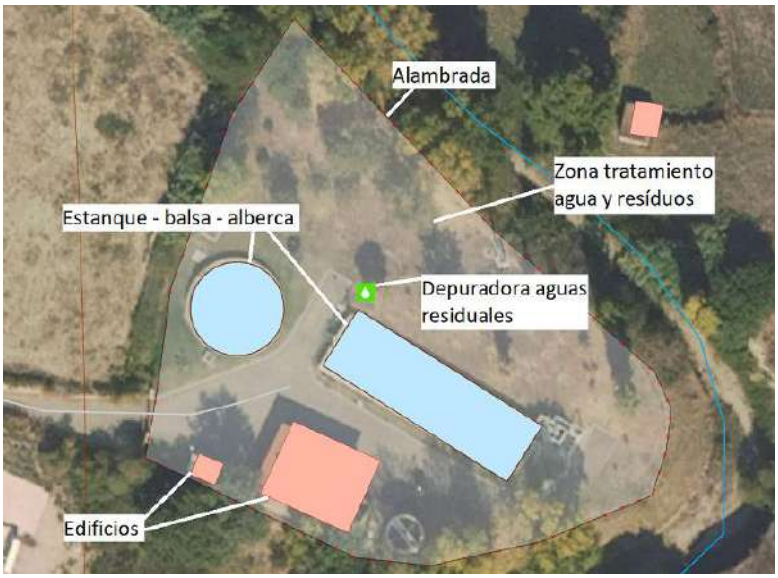


Ilustración 2- 45: Ejemplo de captura para una depuradora

Ilustración 2- 46: Ejemplo de captura para un observatorio espacial

Ilustración 2- 47: Ejemplo de captura para un parque-jardín

Ilustración 2- 48: Central termosolar

Transporte de energía

Definición: Vía de conducción energética.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Línea eléctrica: Conjunto de cables soportados por torres o postes destinados al transporte de energía eléctrica.
- Oleoducto: Conducción de combustible provista de bombas y otros aparatos para la distribución de petróleo a largas distancias.
- Gasoducto: Conducción de combustible de grueso calibre destinada a la distribución a distancia de gas.
- Tubería auxiliar: Conducto formado por tubos destinado a la conducción de alguna materia líquida o gaseosa desconocida

Igualmente es importante definir la situación del objeto respecto de la superficie terrestre.

- En superficie: Se encuentra sobre la superficie terrestre.
- Elevado: Se encuentra elevado respecto a la superficie terrestre.
- Subterráneo: Se encuentra bajo la superficie terrestre.

Las entidades que se recogen para este objeto geográfico son elementos lineales continuos que se capturan por su eje representativo.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Líneas eléctricas: Deben capturarse registrando un vértice en la base de cada una de las torres del tendido, a nivel de terreno, de manera que los únicos vértices de la línea que se registran son aquellos que coinciden exactamente con las torres eléctricas. La línea eléctrica debe terminar siempre en torre eléctrica -si el tendido se soterra o se pierde su trazado- o en central eléctrica, subestación o caseta transformadora.
- Tramos elevados: Si la captura se hace mediante restitución, la altura (H) queda perfectamente definida ya que el tramo es visible. Si el registro se realiza con ortofotografía+MDT, debe interpolarse entre la altura de los vértices inicial y final de dicho tramo.
- Tramos subterráneos: Tanto si el registro se hace mediante restitución como por captura sobre ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo. El inicio y fin del tramo de tubería subterránea se hará coincidir exactamente con el tramo de tubería que queda en superficie.

Criterio de Selección:

En los siguientes casos se atenderá a los siguientes criterios:

- Tubería auxiliar: No deben capturarse cuando su longitud sea inferior a 10 m.

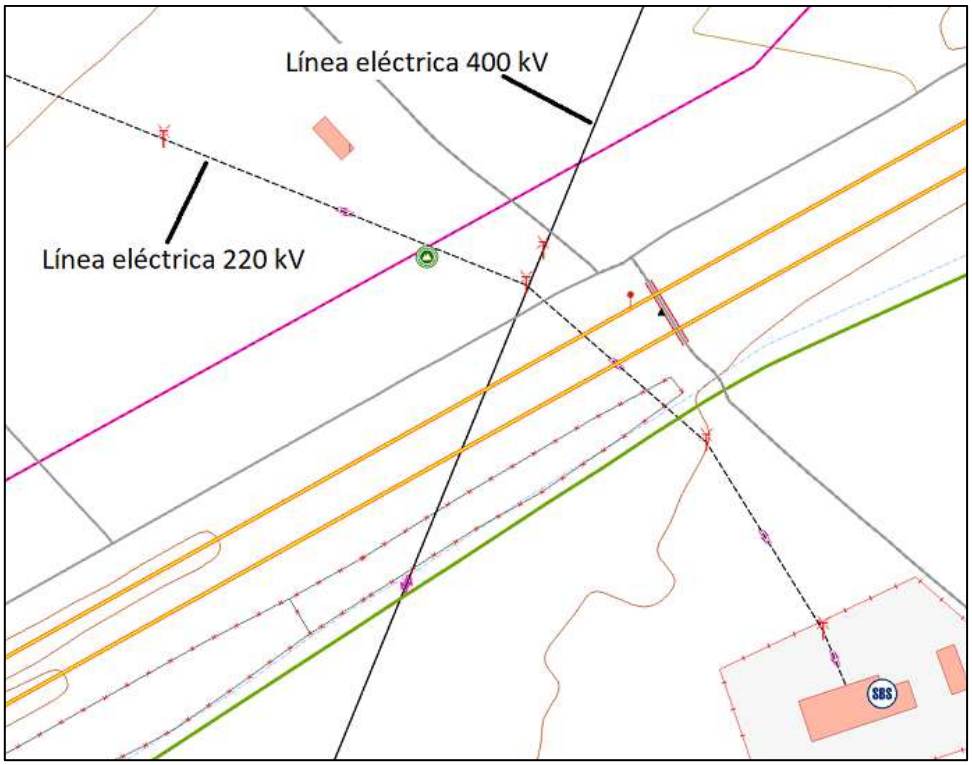


Ilustración 2- 49: Captura correcta de dos líneas eléctricas

Estructura

Definición: Estructura o armazón de cierta altura que sustenta cables de tendido eléctrico u otro tipo de dispositivos o instalaciones.

Geometría: Punto.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Aerogenerador: Torre con equipamiento adecuado para transformar energía eólica en eléctrica.
- Torre eléctrica: Estructura o armazón de cierta altura que sustenta cables de tendido eléctrico.
- Antena: Estructura o armazón de cierta altura para situar sistemas de emisión, recepción y/o amplificación de señales electromagnéticas.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento puntual en el centro de la base de la estructura.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Torre eléctrica o de telecomunicaciones:
    - Debe situarse sobre la línea de tendido y exactamente coincidente con un vértice de esta.
    - Además de las torres intermedias de un tendido, se usará también la torre cuando la línea se soterra o se pierde su trazado.

- Cuando no se vea la torre, pero se identifique un quiebro en la línea, se debe capturar una torre en el punto de quiebro de la línea.

Criterios de selección:

Deben capturarse todas los que sean visibles.

Ilustración 2- 50: Torre eléctrica - Captura correcta en la base de la estructura

Ilustración 2- 51: Torre eléctrica - Captura incorrecta en la parte superior de la estructura

Ilustración 2- 52: Torre eléctrica - Captura incorrecta en la sombra de la estructura

Edificios y construcciones

Población o Mancha urbanizada

Definición: Ciudad, pueblo, aldea, urbanización, etc.

Geometría: Polígono.

Método de captura:

El registro se realizará mediante un elemento superficial que representa el contorno exterior aproximado de la población. Se deben registrar tanto los núcleos urbanos perfectamente consolidados como los conjuntos de edificaciones fuera del casco mediante el trazado de la envolvente que contiene sus edificaciones.

Ilustración 2- 53: Ejemplo de captura de una población

Edificio

Definición: Construcción fija, hecha con materiales resistentes, para habitación humana o para otros usos.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías constructivas:

- Genérico: El edificio no presenta ninguna particularidad en especial como para pertenecer a alguno de los otros tipos. Se incluyen aquí los edificios de viviendas.
- Torre: Estructura estrecha y relativamente alta.
- Torre de vigilancia: Estructura o armazón de cierta altura donde se realizan labores de vigilancia cinegética o contra incendios del entorno que lo circunda.
- Faro: Torre alta situada en las costas y puertos que emite una luz potente a intervalos para orientar de noche a los navegantes.
- Chimenea: Construcción diseñada para dar salida a gases producto de alguna combustión.
- Molino: Edificio destinado a moler, principalmente cereal, a partir de energía hidráulica o eólica.
- Nave agroindustrial: Edificio de dimensiones y geometría apto para la producción o almacenamiento de carácter agrícola o industrial.
- Pabellón polideportivo: Edificio cubierto destinado a la práctica de varios deportes.
- Estadio: Edificio de grandes dimensiones con graderías para los espectadores, destinado a competiciones deportivas.
- Edificio ligero: Construcción estable para dotar de cierta protección a los elementos que se encuentran bajo ella. Ejemplos de este tipo son las marquesinas, naves abiertas, casetas o cobertizos.
- Invernadero: Construcción con techo, generalmente translúcido, donde se controlan factores ambientales para favorecer el cultivo de plantas.
- Plaza de Toros: Edificación cerrada, generalmente circular y descubierta, donde se celebran las corridas de toros.
- Religioso: Edificio construido para la práctica de una religión.
- Caseta transformadora: Edificio diseñado para aumentar o disminuir la tensión eléctrica.

Igualmente es importante definir el estado físico en el que se encuentra el elemento.

- En uso: Se encuentra en uso o en condiciones aparentes de buen estado.
- En construcción: Se encuentra en construcción.
- En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.

Las entidades que se recogen para este objeto geográfico son elementos superficiales que representan los límites externos de la construcción, siendo los patios agujeros en la geometría.

El registro de todas las edificaciones en general debe realizarse:

- Cuando se trabaja por restitución fotogramétrica, la captura se realiza por la parte superior de la edificación (ya que es difícil realizarla sobre el terreno, especialmente en poblaciones).
- Cuando se trabaja sobre una ortofotografía, la captura se realiza por la parte superior de la edificación y después, para sortear el abatimiento sobre todo en edificios de gran altura, se traslada la geometría a la intersección de la fachada, en alguna esquina visible de ésta, con el suelo.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Plazas de toros: Deben capturarse por los límites externos de su recinto, no debe capturarse el ruedo.
- Faros: Deben capturarse por los límites externos de su estructura, no debe capturarse exclusivamente la torre de la linterna salvo que la única estructura visible sea esta.

Criterios de selección:

El lado mínimo que debe tener este objeto geográfico para ser susceptible de ser capturado es de 2 m.

Además de la norma general, se aplicarán los siguientes criterios de densificación:

- Alta densidad o superficie de edificación por km^2 superior al 11%: Todas las edificaciones que se encuentren a menos de 5 m, unas de otras, se agrupan en una misma edificación, que conserve las características generales del conjunto. Esta situación es frecuente en el interior de las poblaciones, donde los edificios se agrupan formando manzanas.

  La agrupación de edificios debe respetar la estructura urbana, poniendo especial atención en los cascos antiguos, ensanches, suburbios, arrabales y urbanizaciones, de forma que se pueda obtener una idea de su desarrollo.

- Baja densidad o superficie de edificación por km^2 inferior al 11%: Deben representarse todas las edificaciones poniendo especial cuidado en su localización debido a su importancia cartográfica. Si existieran edificaciones en estado ruinoso deberán registrarse todas ellas, así como los caminos de acceso a las mismas.
- Alta densidad de invernaderos: En zonas de cultivo intensivo y donde exista una gran abundancia de invernaderos deberá generalizarse, agrupándolos en unidades más grandes, por ejemplo: por polígonos formados entre caminos.

Ilustración 2- 54: Plaza de toros - Captura correcta

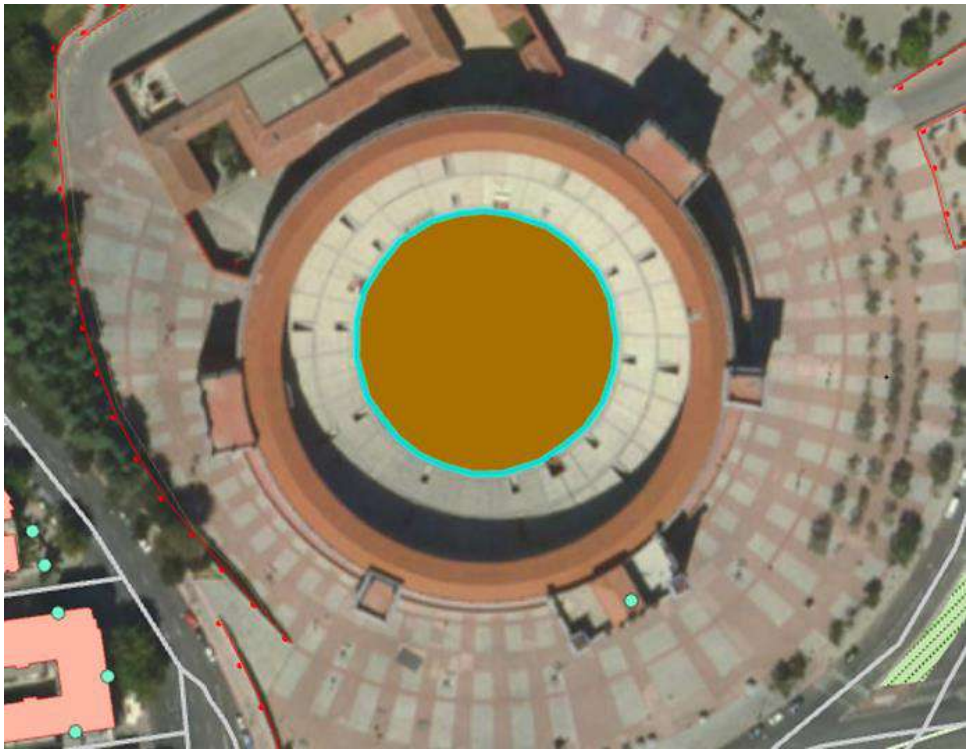


Ilustración 2- 55: Plaza de toros - Captura incorrecta

Ilustración 2- 56: Faro - Captura correcta

Ilustración 2- 57: Faro - Captura incorrecta

Ilustración 2- 58: Edificios ligeros en área de servicio - Captura correcta (cubiertas de protección surtidores)

Ilustración 2- 59: Edificio ligero en playa - Captura incorrecta (se trata de un elemento no estable)

Construcción de almacenaje

Definición: Construcciones destinadas a almacenar sustancias, materiales u objetos en su interior.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías según sus características o el uso al que se destina.

- Piscina: Recinto que almacena agua para el baño y la natación.
- Balsa-alberca-estanque: Recinto excavado en tierra o realizado con fábrica de ladrillo, tapial o mampostería para almacenar agua con fines diferentes al baño (riego, ornamental, abrevadero, etc.).
- Vaso de salina: Vaso en el que se deja evaporar agua salada de forma natural, para dejar solo la sal y poder recolectarla.
- Depósito de almacenamiento: Construcción cerrada donde se almacena algún producto, generalmente líquido, como hidrocarburos, aceites u otros (excepto el agua).
- Depósito de agua en superficie: Construcción cerrada y situada sobre el terreno donde se almacena agua.
- Depósito de agua subterráneo: Construcción cerrada y soterrada (al menos en su mayor parte) donde se almacena agua.
- Depósito de agua elevado: Construcción cerrada y elevada sobre el terreno donde se almacena agua.
- Silo: Lugar cerrado y seco para guardar trigo u otros granos, semillas o forraje.

Igualmente es importante definir el estado físico en el que se encuentra el elemento.

- En uso: Se encuentra en uso o en condiciones aparentes de buen estado.
- En construcción: Se encuentra en construcción.
- En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.

Las entidades que se recogen para este objeto geográfico son elementos superficiales que delimitan perfectamente el borde o margen de las esquinas de la construcción.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Balsas de riego: Pueden capturarse como dos tipos de objeto geográfico:
    - Como balsa-alberca-estanque: Cuando es de construcción totalmente artificial.
    - Como laguna (ver *Laguna-lago*): Cuando es una depresión natural cerrada por una presa o dique.

Criterio de selección:

Como norma general los elementos se capturan siempre que su superficie sea mayor de 10 m^2, siendo el lado mínimo de 2 m.

En los siguientes casos se atenderá además a los siguientes criterios de selección:

- Vasos de salinas: Si la superficie de los vasos o estanques es inferior a 2,500 m^2 (≈50x50) deben generalizarse, capturando un único vaso superficial exterior, sin los muros de tierra interiores.

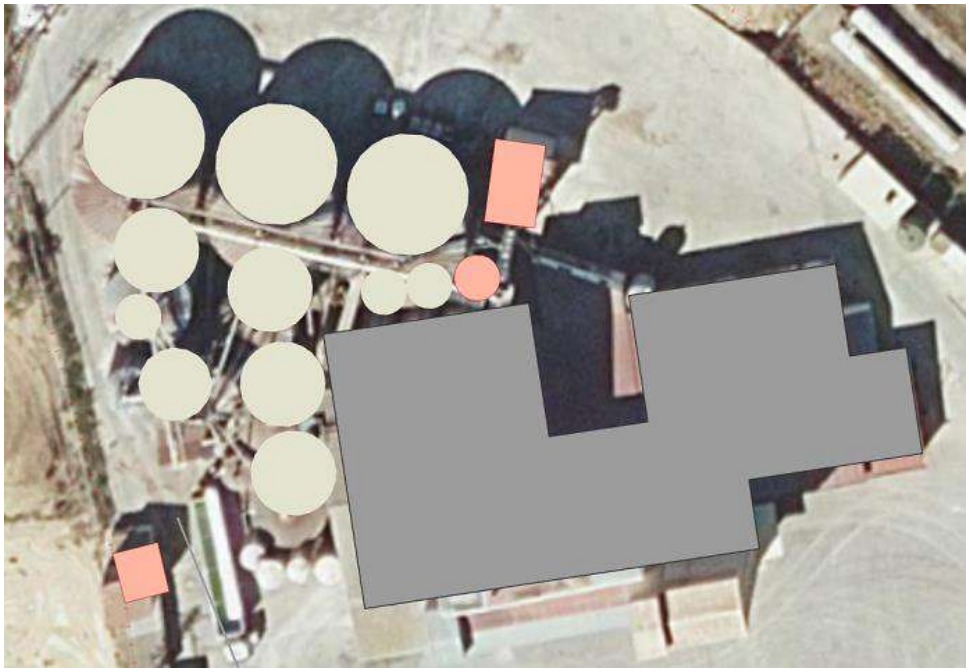


Ilustración 2- 60: Captura de silos

Ilustración 2- 61: Captura de depósito de agua

Monumento

Definición: Construcción para conmemorar algo o que tiene cierto valor histórico o artístico.

Geometría: Punto.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Genérico: Cualquier monumento que no se ajusta a los otros tipos.

- Cruz: Representación del símbolo cristiano, generalmente construida en piedra, que suele encontrarse en cruces de caminos o en lugares donde recuerda algún acontecimiento.
- Fuente ornamental: Fuente monumental, con valor artístico o histórico y que, por lo general, no permite coger agua.

Las entidades que se capturan para este objeto geográfico son elementos puntuales indicativos del lugar donde está situado.

Criterios de selección:

Deben capturarse todos los que sean visibles.

Ilustración 2- 62: Fuente ornamental que debe capturarse como monumento

Ilustración 2- 63: Fuente que permite coger agua que debe capturarse como punto hidrográfico

Cerramiento

Definición: Cercado que cierra o divide un espacio.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Cerca-muro: Pared divisoria construida con materiales de obra, vertical y resistente.
- Alambrada: Cercado de alambres afianzado en postes para cerrar un espacio o dividirlo.
- Muro de contención: Muro destinado a contener el empuje de las tierras o del agua que se encuentran a uno de sus lados, impidiendo su avance.
- Muralla histórica: Muro u obra defensiva que rodea una plaza fuerte o protegía un territorio.

Las entidades que se capturan para este objeto geográfico son elementos lineales que representan el eje longitudinal del cercado. La captura debe hacerse sobre el terreno, es decir no considerando la altura del cerramiento.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Cerramientos paralelos a vías de comunicación: Deben capturarse de modo que no intercepte el cerramiento con la vía.
- Muros de contención: Deben capturarse los muros destinados a contener el empuje de la tierra y los muros de contención en el lateral o laterales de algunos ríos, coincidiendo en estos casos el muro y los márgenes o el eje del río.
- Pequeñas barreras y azudes transversales a la corriente de un río cuya labor sea la de retener, ralentizar o embalsar agua: Deben capturarse como presa, no como muro de contención.
- Diques de contención en balsas-albercas y lagunas: Deben capturarse como presa, no como muro de contención.

Criterio de Selección:

En los siguientes casos se atenderá a los siguientes criterios de selección:

- Cercas-Muros y alambradas: Se aplicarán los siguientes criterios de densificación.
    - Alta densidad o más de 2,500 metros de cerramiento por km^2: Deben capturarse todos aquellos de longitud superior a 250 m, o aquellos de longitud inferior a 250 m siempre que sus extremos estén unidos a otros cerramientos o edificaciones. También puede generalizarse haciendo una selección de los más importantes, eliminando los interiores a polígonos ya cerrados.
    - Baja densidad o menos de 2,500 metros de cerramiento por km^2: Deben capturarse todos, aunque no superen las dimensiones mínimas, dada su importancia cartográfica.
- Murallas históricas: Deben capturarse todas las que sean visibles.

Infraestructura de cruce

Definición: Construcción que permite salvar y cruzar un obstáculo (río, barranco, montaña, etc.) o cruzar otras vías o infraestructuras situadas a otro nivel (carreteras, canales, etc.).

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Puente: Paso elevado para vehículos, ferrocarriles o peatones.
- Puente peatonal: Paso elevado para peatones y vehículos sin motor.
- Túnel: Paso subterráneo artificial para establecer una comunicación.
- Acueducto: Construcción elevada para conducir el agua a una cota adecuada salvando una depresión del terreno.
- Vado: Lugar de un río con fondo firme, llano y poco profundo, natural o artificial, por donde se puede pasar andando, cabalgando o en algún vehículo.

Igualmente es importante definir el estado físico en el que se encuentra el elemento.

- En uso: Se encuentra en uso o en condiciones aparentes de buen estado.
- En construcción: Se encuentra en construcción.
- En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento lineal que representa su eje longitudinal totalmente coincidente con la vía (excepto en los casos en los que no se aprecie vía) o cauce.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Puentes, peatonales y acueductos: Si la captura se hace mediante restitución, la altura (H) queda perfectamente definida ya que el tramo es visible. Si el registro se realiza con ortofotografía+MDT, debe interpolarse entre la de los vértices inicial y final de dicho tramo.
- Túneles: Tanto si el registro se hace mediante restitución como por captura sobre ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo.

Criterio de Selección:

En general, deben capturarse todas las que sean visibles.

En los siguientes casos se atenderá a los siguientes criterios de selección:

- Puentes sobre cauces naturales o artificiales: Deben capturarse si su longitud es mayor de 3 m.

Pista deportiva

Definición: Pista o terreno de juego destinada a la práctica de cualquier deporte.

Geometría: Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías según el tipo de deporte que se realice (no obstante, se pueden incluir otras muchas tipologías, dependiendo de la realidad del lugar y de lo que se quiera representar):

- Campo de Béisbol: Pista deportiva para la práctica del béisbol.
- Pista de atletismo: Pista deportiva para la práctica del atletismo.
- Velódromo: Pista deportiva para la práctica de carreras ciclistas.
- Campo de fútbol: Pista deportiva para la práctica del fútbol.
- Hipódromo: Pista deportiva para la práctica de carreras de caballos.
- Pista para raqueta: Pista deportiva para la práctica de deportes de raqueta.
- Cancha: Pista deportiva para la práctica del balonmano (Voleibol o Basquetbol).Otra: Pista deportiva para la práctica de otros deportes no incluidos en los tipos anteriores.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento superficial que representa el contorno exterior del terreno de juego. Las líneas que constituyen el esquema interior del terreno de juego no se representan.

No se incluyen en este contorno las edificaciones anexas, por ejemplo, las gradas.

Criterio de Selección:

Deben capturarse todas las que sean visibles de manera individual, aunque estén colindantes.

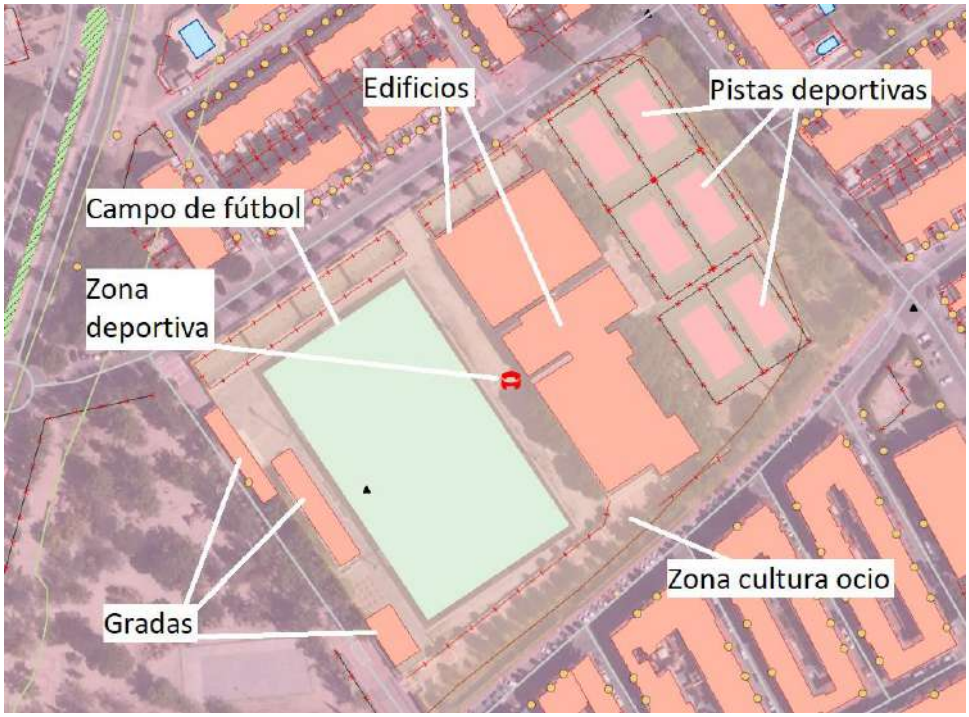


Ilustración 2- 64: Pistas de tenis y campo de fútbol ubicados en una zona deportiva

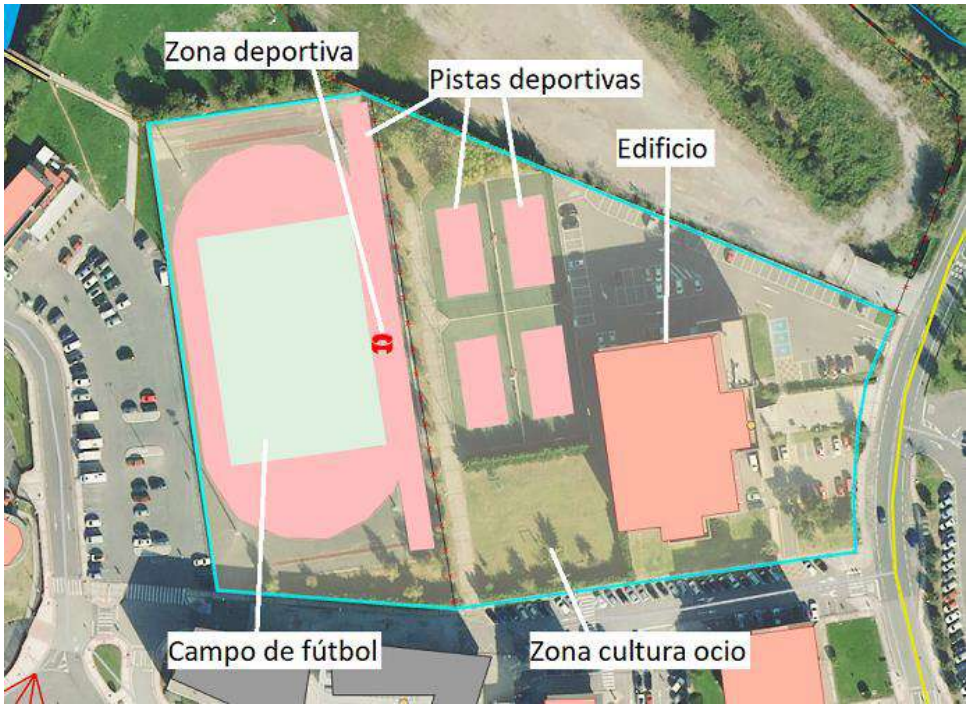


Ilustración 2- 65: Campo de fútbol en el interior de una pista de atletismo

Construcción portuaria

Definición: Elemento ubicado en la orilla de una masa de agua o curso fluvial para facilitar el embarque y desembarque de personas y mercancías, servir de refugio a las embarcaciones o defender los márgenes contra las olas.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Muelle: Malecón o muro para que los barcos atraquen a tierra.
- Rompeolas: Muro grueso o escollera construido a la orilla del mar, playa o puerto para protegerlos de la fuerza del oleaje.
- Embarcadero: Muelle o pasarela sustentada por pilares o flotando sobre el agua.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento lineal que representa el contorno exterior o eje del elemento.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Muelles: se capturan siempre coincidentes planimétricamente con la línea de agua (línea de costa, margen de embalse, margen de río o margen de laguna) pero altimétricamente con la altura (H) que les corresponda.
- Rompeolas: Se representan coincidentes planimétrica y altimétricamente con la línea de agua (línea de costa, margen de embalse, margen de río o margen de laguna).
- Embarcaderos: Al estar sobre el agua (flotantes o sobre pilares), se capturan no coincidentes con la línea de agua. Cuando su anchura es superior a 5 m se capturan por

el perímetro o contorno del pantalán creando una superficie cerrada, es decir, coincidentes el primer y el último vértice. Si su anchura es menor de 5 m se capturan con una línea por su eje representativo.

Ilustración 2- 66: Puerto con muelle (rojo), embarcadero (naranja) y rompeolas (negro)

Ilustración 2- 67: Embarcaero capturado por eje (ancho<5m) y por contorno (ancho>5m)

Ilustración 2- 68: Rompeolas fuera de un puerto, en este caso para proteger la costa

Auxiliar construido

Definición: Detalle planimétrico de algún elemento construido importante para la representación cartográfica pero que no se ajusta a ningún otro objeto del modelo de datos.

Geometría: Línea.

Método de captura:

El registro se realizará mediante un elemento lineal que representa la forma del objeto a capturar. Ejemplos: bordillos, escaleras, obras de fábrica, límites de explanadas, ejes de la pista de competición en circuitos, toboganes de un parque acuático, etc.

No se debe usar este objeto de forma masiva, solo para aquellos elementos notables que ayuden a la interpretación cartográfica. Su uso no caracteriza al elemento, solo sirve como “dibujo” o “relleno” en aquellas zonas en que sea necesario.

Transportes

Vía de comunicación

Definición: Vía para la circulación de vehículos o peatones.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Autopista-autovía: Vía para la circulación exclusiva de vehículos automóviles con calzadas separadas para ambos sentidos, accesos y salidas independientes, cruces o pasos a distinto nivel, y carece de acceso directo a las propiedades colindantes.

- Carretera: Vía para la circulación fundamentalmente de vehículos automóviles, constituida por una o varias calzadas, que no reúne las características propias de la autopista y la autovía.
- Calle: Vía en el espacio urbano o urbanizado.
- Enlace-rotonda: Vía de unión entre dos vías de comunicación, vía de acceso o cambio de sentido, o anillo por el que se circula solo en un sentido.
- Camino: Vía no pavimentada abierta al tránsito que puede presentar algún tipo de revestimiento ligero y que no tiene la categoría de carretera.
- Senda: Vía natural, sin ningún tipo de revestimiento, cuya anchura y trazado están destinados principalmente al tránsito de peatones o ganado menor.

Igualmente es aconsejable diferenciar las siguientes características:

- Firme o tipo de revestimiento:
    - Pavimentado: Existe una capa superior del firme de asfalto u hormigón o concreto.
    - Revestimiento ligero: No existe pavimento, aunque puede existir otro tipo de capa o cubierta que resguarde la superficie natural de la vía.
    - Sin revestimiento: No dispone de capa o cubierta para resguardar la superficie natural de la vía.
- Estado físico en el que se encuentra el elemento:
    - En uso: Se encuentra en uso o en condiciones aparentes de buen estado.
    - En construcción: Se encuentra en construcción.
    - En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.
- Situación respecto de la superficie terrestre:
    - En superficie: Se encuentra sobre la superficie terrestre.
    - Elevado: Se encuentra elevado respecto a la superficie terrestre.
    - Subterráneo: Se encuentra bajo la superficie terrestre.
    - En vado: Tramo que atraviesa a nivel una masa de agua.
    - Transbordador: Tramo que salva una masa de agua cruzándola mediante algún sistema de navegación.
- Identificador: Código que identifica de manera única cada carretera. Puede obtenerse de un organismo oficial competente o ser definido dentro de la base de datos con unas reglas establecidas a priori.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento lineal continuo que representa su eje representativo.

Todos los tramos (en superficie, elevados o subterráneos) que conforman la vía deben tener el mismo identificador. Para ello es necesaria la conectividad de todos sus tramos y la unión de las vías entre sí, de modo que generen una red completamente conectada, sin tramos sueltos.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Autopistas-autovías, carreteras y calles: Se capturará un eje por cada calzada, sea cual sea su anchura (2, 3, o más carriles), no debiendo nunca registrarse por los márgenes. En vías de un único sentido, el registro se hará en el sentido de la marcha, y en calzadas de doble sentido deberá ser el mismo a lo largo de todo su trazado.
- Carreteras que cruzan poblaciones o agrupaciones urbanas: Se debe intentar no interrumpir la vía, continuando su trazado por las calles principales de forma que exista continuidad del elemento a lo largo de toda su captura.
- Vías que llegan a poblaciones: Debe intentarse su conexión con el trazado urbano sin que queden tramos desconectados.
- Caminos y sendas en zonas boscosas: Debe intentarse asegurar su continuidad al máximo en las zonas donde se pierde su trazado.
- Vías de comunicación que se crucen a distinto nivel: Se capturarán los puentes y túneles correspondientes completamente coincidentes con la vía.
- Tramos elevados: Si la captura se hace mediante restitución, la altura (H) queda perfectamente definida ya que el tramo es visible. Si el registro se realiza con ortofotografía+MDT, debe interpolarse entre la de los vértices inicial y final de dicho tramo.
- Tramos subterráneos: Tanto si el registro se hace mediante restitución como por captura sobre ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo.
- Rotondas: deberán capturarse mediante un elemento lineal cerrado donde confluyen las vías de comunicación, perdiendo éstas la continuidad al llegar a ella, garantizando la unión perfecta de la vía con la glorieta.
- Cerramientos paralelos a vías de comunicación: Deben capturarse de modo que no intercepte el cerramiento con la vía.

Criterio de Selección:

En los siguientes casos se atenderá a los siguientes criterios de selección:

- Autopistas-autovías, carreteras y calles: Deben capturarse todos los tramos.
- Caminos: Deben capturarse todos, a excepción de:
    - Aquellos que no terminen en un detalle planimétrico y su longitud sea inferior a 500 m.
    - Aquellos que terminen en cultivos en zonas de gran densidad de caminos.
- Tramos de caminos y sendas desdoblados: Cuando la distancia entre las dos ramas sea inferior a 50 m solo se capturará uno de los ejes, el más representativo.
- Sendas:
    - En zonas de montaña: Deben identificarse todas las sendas visibles, ya que constituyen las únicas vías de comunicación en estos lugares. Si alguna quedara cortada, se registrará hasta donde finaliza.

- En otras zonas: Se deben capturar sólo aquellas que se intuyan como estables. Por ejemplo, se han de evitar las servidumbres agrícolas y rodadas al ser habitualmente huellas temporales.

Transporte por rieles

Definición: Vía de transporte sobre rieles.

Geometría: Línea.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Ferrocarril: Sistema de transporte de personas y mercancías guiado sobre una vía férrea.
- Funicular: Ferrocarril que recibe la tracción mediante un cable, para superar grandes pendientes.
- Tranvía: Ferrocarril urbano que transcurre por una calle coexistiendo con la circulación de automóviles y personas.
- Metro: Sistema de trenes urbanos ubicado dentro de una ciudad y su área metropolitana.

Igualmente es aconsejable diferenciar las siguientes características:

- Número de vías de la plataforma del ferrocarril:
    - Única: Tramo constituido por una única vía.
    - Doble: Tramo constituido por dos o más vías.
    - Patio de vías: Tramo de vía situado en la plataforma de maniobras de las estaciones.
- Estado físico en el que se encuentra el elemento:
    - En uso: Se encuentra en uso o en condiciones aparentes de buen estado.
    - En construcción: Se encuentra en construcción.
    - En ruinas: Se ha demolido o destruido en parte, aunque hay restos visibles.
- Situación respecto de la superficie terrestre:
    - En superficie: Se encuentra sobre la superficie terrestre.
    - Elevado: Se encuentra elevado respecto a la superficie terrestre.
    - Subterráneo: Se encuentra bajo la superficie terrestre.

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento lineal continuo por el eje representativo de la plataforma del ferrocarril, que puede contener una o varias vías.

En el caso de patio de vías deben tener siempre representación las vías exteriores, generalizando el resto de las vías interiores que lo conforman.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Cruces a distinto nivel: Se capturarán los puentes y túneles correspondientes completamente coincidentes con la vía férrea.
- Tramos elevados: Si la captura se hace mediante restitución, la altura (H) queda perfectamente definida ya que el tramo es visible. Si el registro se realiza con ortofotografía+MDT, debe interpolarse entre la de los vértices inicial y final de dicho tramo.
- Tramos subterráneos: Tanto si el registro se hace mediante restitución como por captura sobre ortofotografía+MDT, la altura (H) del tramo subterráneo debe interpolarse entre la de los vértices inicial y final de dicho tramo.

Criterio de Selección:

En general deben capturarse todos los tramos de vía férrea.

En los siguientes casos se atenderá a los siguientes criterios de selección:

- Patio de vías: Cuando son muy numerosas, pueden capturarse sólo las más representativas que den idea del conjunto y dimensiones de la instalación.

Teleférico

Definición: Cabina accionada por cables aéreos que sirve para transportar personas u objetos salvando un gran desnivel.

Geometría: Línea.

Método de captura:

Las entidades que se recogen para este objeto geográfico se registran mediante un elemento lineal continuo por el eje representativo, capturándose un vértice en cada uno de los pilares que sirven de soporte a los cables.

Criterio de Selección:

Deben capturarse todos los teleféricos.

Infraestructura de transporte

Definición: Instalaciones destinadas al tránsito de pasajeros y/o mercancías, o a prestar algún servicio auxiliar.

Geometría: Punto o Polígono.

Método de captura:

Para la captura de este objeto geográfico es aconsejable la discriminación de las siguientes tipologías:

- Aeródromo: Aeropuerto, aeroclub, campo de aviación menor, etc.
- Helipuerto: Aeródromo para uso de helicópteros.
- Hidroaeródromo: Área en el mar o en un lago destinada al amaraje de hidroaviones.
- Base aérea: Aeródromo de uso militar.
- Estación de ferrocarril: Instalaciones ferroviarias para el tránsito de pasajeros y mercancías.

- Puerto: Instalaciones portuarias para el resguardo de embarcaciones, así como el tránsito de pasajeros y mercancías.
- Área de servicio: Instalaciones para el suministro de carburantes, hospedaje, restaurante, taller, descanso y otros servicios para facilitar la comodidad y seguridad de los usuarios de la carretera.
- Estacionamiento: Área en superficie destinada habitualmente a aparcar vehículos.
- Pista de aterrizaje: Pista utilizada para el despegue y aterrizaje de aeronaves.

Como norma general se usará siempre la representación superficial cuando la instalación sea extensa y agrupe varias edificaciones, reservándose la puntual para entidades muy pequeñas, que se ubiquen en una sola edificación, o que no le sea aplicable la representación superficial.

Si el registro se realiza mediante un polígono, éste debe definir, de la forma más precisa posible, el borde o contorno exterior de la instalación. Si la captura se realiza mediante un punto, éste debe situarse en su lugar más representativo. En ambos casos los objetos interiores de la instalación deben registrarse siguiendo sus normas de captura específicas.

Estas instalaciones suelen estar delimitadas por algún tipo de cerramiento (cerca, muro, alambrada, muralla, etc.) que debe capturarse coincidente con el polígono que la define.

En los siguientes casos se atenderá además a las siguientes normas específicas de captura:

- Estacionamientos: Se capturan siempre como polígonos. Deben capturarse los grandes estacionamietos en superficie, en zonas de playa, zonas recreativas, zonas verdes, estacionamientos de grandes centros comerciales, etc., debiendo registrarse las vías de acceso.
- Pistas de aterrizaje: Se capturan siempre como polígono.
- Aeródromos: Deben contener al menos una pista de aterrizaje.

Criterio de Selección:

Deben capturarse todas las infraestructuras de transporte.

En los siguientes casos se atenderá a los siguientes criterios de selección:

- Estacionamientos:
    - Parkings subterráneos y aparcamientos en línea o batería en los laterales de las calles: No se capturan.
    - Estacionamientos en espacios naturales (zonas de montaña, playas, parques nacionales o naturales, etc.): Se representan todos dado que su localización es importante.
    - En los demás casos solo se representan los estacionamientos cuya superficie sea superior a 2,500 m^2.

Ilustración 2-69: Estacionamiento junto a una playa

2.3. Integración de cartografía e información geográfica existente

2.3.1. Formato del dato de entrada

Los formatos más habituales en los que se encuentra la cartografía e información geográfica existente se pueden resumir en los siguientes:

- Formato papel, por ejemplo, cartografía impresa, imágenes fotogramétricas analógicas.
- Modelo ráster, por ejemplo, imágenes digitales, cartografía escaneada (o rasterizada), modelos digitales de elevaciones (o de cualquier otra variable).
- Modelo vectorial, por ejemplo, cartografía digital en formato CAD, capas vectoriales de un SIG, tracks GPS.
- Datos tabulados, por ejemplo, datos estadísticos de población, listados de centros educativos.

2.3.1.1. Formato papel

A pesar en la era digital en la que nos encontramos, todavía puede haber información de interés que se encuentre en formato papel. Generalmente porque se trate de cartografía relativamente antigua o directamente sean documentos de carácter histórico. Además, es un formato muy utilizado para la difusión de los datos geográficos, ya que la utilización de formatos grandes permite una lectura de conjunto mucho más cómoda que en pantalla. A su vez, las cámaras fotogramétricas analógicas se han venido utilizando hasta hace muy poquitos años (incluso actualmente puede haber alguna en uso), por lo que puede no ser difícil encontrar fotogramas en soporte película.

El tipo de papel o de soporte físico, así como su estado de conservación influyen enormemente en varios aspectos de la calidad de la información que se puede extraer, principalmente:

- La exactitud posicional. Una mala manipulación, almacenamiento o incluso la deficiente calidad del soporte pueden ocasionar deformaciones dimensionales, con lo que cambia la escala del documento de forma no homogénea, lo que se traduce en errores posicionales si se pretende utilizar el documento como fuente directa de datos. Si estas deformaciones son homogéneas, en el proceso de georreferenciación se pueden minimizar estos errores.
- La compleción y la exactitud temática. Una mala calidad visual, bien porque la impresión fue deficiente o porque la antigüedad y las condiciones de almacenamiento han degradado el contenido, puede llevar a una mala interpretación de la información o directamente a omisiones.

2.3.1.2. Formato ráster

Definición

En el modelo ráster el espacio se divide sistemáticamente en unidades mínimas denominadas celdas (pixeles), habitualmente de forma cuadrada, formando una malla regular. Cada celda tiene su valor correspondiente.

El formato ráster suele emplearse para almacenar imágenes y también para variables continuas, como pueden ser las elevaciones o las temperaturas:

Imágenes digitales: obtenidas con distintas tecnologías (teledetección, fotogrametría, mapas escaneados, etc.), se puede decir que sus valores son proporcionales a la cantidad de energía recibida (motivada por la distinta reflectancia de cada superficie, así como por su iluminación, principalmente) en una longitud de onda concreta (o en un rango de longitudes de onda) en el momento de la captura. Varias características a destacar que caracterizan a las imágenes son las siguientes:

- Resolución espacial. En imágenes aéreas y de satélite, y en ortofotografías, se suele hablar de GSD (Ground Sample Distance), que no es más que el tamaño de píxel en el terreno. Se suele expresar en metros o en centímetros y depende de múltiples factores como el tamaño del sensor, su número de píxeles, la distancia focal y la altura de vuelo.
- Resolución radiométrica o profundidad de bits. Cantidad de bits utilizados para definir cada píxel. Cuanto mayor sea la profundidad de bits, tanto mayor será la cantidad de tonos que puedan ser representados. Por ejemplo, una imagen binaria utiliza 1 bit por píxel, de manera que pueden representar sólo dos tonos. Habitualmente se utilizan 8 bits, que permiten distinguir entre 256 tonos (2^8), si bien las cámaras fotogramétricas actuales y los sensores a bordo de satélite suelen utilizar un mayor número de bits por píxel.
- Resolución espectral o número de bandas. Puesto que el espectro de radiación puede dividirse en distintos grupos según su longitud de onda, los sensores que toman estas imágenes suelen recoger de forma diferenciada varias bandas del mismo. Así, las imágenes en color se componen de tres bandas (rojo, verde y azul) y suelen emplear 8 bits para una de ellas. Los sensores de teledetección habitualmente recogen un mayor número de bandas, muchas de ellas fuera del espectro visible, para poder estimar diferentes variables como el estado de la vegetación, el nivel de humedad, la temperatura en superficie y un largo etcétera de posibilidades.

  La mayoría de software puede leer y trabajar con imágenes de una y tres bandas (escala de grises y color). Para trabajar con un número mayor de bandas, suele requerirse software un poco más especializado.

Rásters que contienen valores de una variable continua: en este caso, el valor de cada celda representa una magnitud o una variable. Por ejemplo, la elevación del terreno, su orientación o su temperatura. Se suele requerir que el tipo de dato de los píxeles sea de números reales (habitualmente un tipo *float*), como en el caso de los modelos digitales de elevaciones (MDE). En ocasiones también se utilizan para representar variables discretas, como los usos del suelo.

Suelen ser rásters de una banda única y son de gran utilidad para multitud de tipos de análisis. De hecho, la mayoría de software SIG poseen una amplia batería de herramientas de procesamiento ráster, como cálculo de orientaciones, cuencas visuales o cuencas vertientes (que se obtienen a partir de MDE), así como métodos de interpolación o calculadoras ráster para analizar diferentes aspectos del territorio.

Formatos de datos ráster habituales

Los datos ráster son con frecuencia muy voluminosos. Un modelo de representación que minimice el tamaño de los datos, unido a un manejo óptimo de memoria, son requisitos de suma importancia para cualquier sistema informático que maneje datos ráster, máxime considerando

los grandes volúmenes de datos de este tipo que hoy en día se manejan. Piénsese en las imágenes de alta resolución que cubren grandes extensiones de territorio, en las coberturas periódicas de imágenes de satélite o en la ingente cantidad de documentación analógica que se ha ido escaneando en las últimas décadas.

Así pues, las técnicas de compresión de imágenes han cobrado muchísima importancia y muy a menudo se admite el uso de técnicas con pérdidas de información, que alteran y degradan la calidad de la imagen, pero consiguen unas tasas de compresión muy superiores. Cuando se trata de archivos ráster en los que el valor del píxel tiene un significado específico, se deben emplear técnicas de compresión sin pérdidas.

Generalmente se basan en que los valores de las celdas vecinas a una dada es más probable que sean iguales o parecidos a ella frente a que sean muy diferentes. En esto se basa, por ejemplo, la técnica denominada Run–Length Encoding. Se trata de una técnica muy sencilla que codifica una serie de n valores idénticos como un par de números, el primero de los cuales representa el propio valor, y el segundo, el número de veces que se repite.

También merece la pena comentar la estructura clásica de los *quadtrees* o árboles cuaternarios, que divide el espacio en cuadrantes sucesivamente, pero lo hace con más profundidad en aquellas zonas que así lo requieran por contener mayor número de elementos diferentes y necesitar mayor resolución.

Por otro lado, en imágenes de múltiples bandas, cabe mencionar tres formas habituales en las que las bandas se tratan dentro del archivoy el modo en que se ordenan: banda intercalada por línea (BIL), banda intercalada por píxel (BIP) y banda secuencial (BSQ).

Los principales formatos ráster utilizados son los siguientes:

- Formato TIFF (*Tagged Image File Format*): de uso común en fotogrametría y escáneres, están estructurados en tres partes: la cabecera de 8 bytes, la IFD (Image File Directory) y la imagen propiamente dicha.

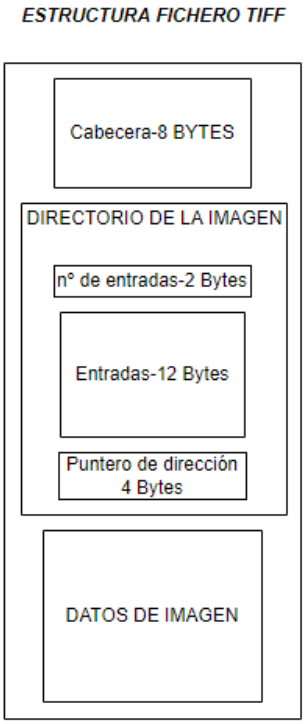


Ilustración 2-70: Estructura del formato TIFF

Las etiquetas definen, entre otros aspectos, el tipo de compresión aplicado a la imagen. También permiten la creación de extensiones, como el GeoTIFF, que introduce la componente geográfica para que la imagen quede georreferenciada.

- Formato JPEG 2000: fue creado por el *Joint Photographic Expert Group* con idea de sustituir al JPEG original. Se basa en la transformación wavelet y permite hacer compresiones con y sin pérdida de información. En el caso de compresiones con pérdida, consigue unas tasas de compresión muy elevadas, superiores a las del JPEG, por lo que resulta útil para imágenes muy voluminosas, como las que se utilizan en observación de la tierra. Sin embargo, este formato es más complejo que el anterior y existe cierta controversia sobre su condición jurídica en cuanto a las patentes de software, por lo que no ha terminado de imponerse en muchos ámbitos.
- Formato ECW (Enhanced Compression Wavelet): creado y optimizado para imágenes aéreas y de satélite. Se basa también en la transformación wavelet y permite unas tasas muy altas de compresión.
- Formato BMP (Bitmap): desarrollado por Microsoft, los archivos BMP permiten utilizar diferentes métodos de compresión, con y sin pérdidas.
- Formato GIF (Graphips Interchange Format): utiliza generalmente la compresión LZW (sin pérdida). Genera archivos relativamente pequeños y de buena calidad, pero está limitado a 256 colores/tonos colores (8 bits). Es muy utilizado para publicar imágenes en Internet. El formato ofrece varias opciones gráficas avanzadas tales como fondos transparentes, imágenes entrelazadas y animación.
- Formato PNG (Portable Network Graphics): formato gráfico basado en un algoritmo de compresión sin pérdida y no sujeto a patentes. Este formato fue desarrollado en buena

parte para solventar las deficiencias del formato GIF y permite almacenar imágenes con una mayor profundidad de color y, en la mayoría de los casos, comprime mejor.

Es ideal para imágenes con grandes áreas de color plano, o con pocas variaciones de color, como, por ejemplo, mapas escaneados o gráficos. Así pues, está diseñado para fines distintos al JPEG, más pensado para fotografías.

No admite CMYK (Cyan, Magenta, Yellow y Black), por lo que no es del todo útil para diseños de impresión.

2.3.1.3. Formato vectorial

Definición

En el formato vectorial no existen unidades fundamentales que dividen la zona recogida, sino que representa los elementos espaciales mediante la codificación explícita de sus posiciones. Modeliza el espacio geográfico mediante una serie de primitivas geométricas que contienen los elementos más destacados. Las primitivas básicas son:

- Puntos: se representan mediante el par de coordenadas (X, Y), o tres coordenadas (X, Y, H) si se incluye la altura, definiendo su posición.
- Líneas: se representan mediante la consecución de coordenadas de los vértices que se aproximan a la forma del elemento.
- Polígonos: representan elementos superficiales que se codifican mediante la consecución de coordenadas de los vértices que se aproximan al contorno del elemento, formado un recinto cerrado.

A estas geometrías se asocian una serie de valores definitorios (qué representan), que puede incluir una atribución adicional. Así, utilizando puntos, líneas y polígonos, puede modelarse el espacio geográfico.

La forma en la que el formato vectorial separa las dos componentes de la información geográfica (geométrica y semántica) es distinta. En el modelo ráster se tiene un conjunto de valores que guardan una estructura dada y por sí misma establece su disposición en el espacio. En el vectorial, por su parte, la componente espacial se recoge explícitamente según una serie de puntos y puede ser más o menos compleja, en función de la entidad a representar o el detalle con que se recoja. Este conjunto de puntos se le relaciona con una serie de valores, que definen las propiedades de la entidad.

Estructura de datos en el formato vectorial

Existen diferentes estructuras de datos que se ajustan al formato vectorial. Se puede diferenciar entre estructuras que no recogen la topología y las que permiten conocer la relación topológica de los elementos geométricos.

Conviene señalar que la topología recoge las relaciones espaciales existentes entre los distintos elementos. En un formato ráster, estas relaciones topológicas vienen implícitas en el propio formato, puesto que se conoce la vecindad entre celdas. Pero en formato vectorial no se conocen a priori estas relaciones de vecindad.

Disponer de forma explícita las relaciones topológicas de los elementos resulta de utilidad para llevar a cabo análisis espaciales. Sin embargo, no aporta beneficio a la hora de una simple representación de la información.

Las estructuras más comunes son:

- Formato vectorial sin topología (*Spaguetti*): almacena únicamente las coordenadas que definen la entidad. De cada elemento geográfico existente se registra su nombre, el número de vértices que definen su frontera y las coordenadas de cada vértice.

  Los elementos puntuales están formados por un único vértice. En los elementos superficiales el primer y el último vértice son exactamente el mismo punto, indicando que se trata de una figura poligonal cerrada.

  En el caso de dos polígonos adyacentes, la definición del lado que comparten se repite en ambos (se duplican esos vértices).

- Formato vectorial con topología básica (*Diccionario de Vértices*): mejora la anterior estructura en el sentido de que registra una sola vez las coordenadas de cada vértice, que es identificado con un nombre o etiqueta. La definición de cada objeto espacial se hace identificando los vértices que lo componen.

  De esta forma, elimina los problemas causados por la duplicidad de los vértices, pero la topología sigue sin estar suficientemente detallada.

- Formato vectorial con topología completa, *DIME* (*Dual Independent Map Encoding*): partiendo del *Diccionario de Vértices*, se identifican los segmentos lineales (sin vértices intermedios) indicando el nombre del segmento, el vértice origen, el vértice final, el polígono que tiene a la izquierda y el polígono que tiene a la derecha. La descripción de los polígonos se hace asignando un nombre e identificando los segmentos rectos que los componen.

- Formato vectorial con topología completa, ARCO/NODO: también parte de un *Diccionario de Vértices*, pero se distingue entre nodos y vértices. En los nodos conectan varias líneas y los arcos son las líneas entre dos nodos, que pueden tener vértices intermedios. La topología de arcos codifica los arcos mediante su nombre, el nodo origen, el nodo final, el polígono que se encuentra a su izquierda y el polígono que se encuentra a su derecha. La topología de polígonos que codifica los polígonos mediante su nombre y los arcos que lo componen. La topología de nodos se codifica los nodos mediante el nombre del nodo y los arcos que en él confluyen.

- Modelos de red: la topología de esta estructura se encuentra íntimamente ligada a la teoría de grafos. La teoría de grafos estudia las propiedades de los grafos, que son colecciones de objetos llamados vértices (nodos) conectados por líneas (arcos) que pueden tener orientación (dirección asignada).

- TIN (*Triangulated Irregular Network*): como se indica en [1], las particularidades del TIN hacen que existan sub–modelos para almacenar el conjunto de triángulos, distintos del habitual arco/nodo. Se pueden destacar dos sub-modelos:

  - Almacenamiento de los triángulos, con las coordenadas de sus tres vértices, un código de identificación y almacenamiento de los códigos de los triángulos adyacentes.

  - Almacenamiento de los vértices y un código para cada uno de ellos, así como los códigos de los vértices a los que se encuentra conectado, en un orden establecido (horario o anti horario).

Formatos vectoriales habituales soportados

En primer lugar, antes de enumerarlos, es conveniente clarificar las principales diferencias entre formatos que se trabajan en CAD (*Computer Aided Design*) y los que utilizan los SIG, al ser los dos modelos vectoriales.

El software CAD está enfocado al diseño, dibujo y edición con finalidades principalmente de representación geométrica, por lo que trabaja con elementos geográficos sin tener en cuenta la información semántica (atributos) que hay detrás de ellos. En cambio, incluye parámetros como el color, peso y estilo de los elementos geométricos creados con vista su representación.

En cambio, en los SIG, cobra especial relevancia la semántica de los elementos, así como sus relaciones espaciales, ya que tiene por un lado una finalidad más de inventario de objetos geográficos y, por otro, de realizar análisis a partir de los datos.

No obstante, las mayores capacidades de tratamiento geométrico del software de CAD, unido a que en la informatización de todo el proceso cartográfico la primera necesidad siempre fue mantener la generación de mapas, ha hecho que el software de CAD haya sido el ampliamente utilizado en las agencias cartográficas como soporte para la cartografía digital. En las últimas décadas, la información contenida en estos formatos CAD se ha ido transfiriendo a formatos más propios de un SIG y a la vez que se le ha ido dotando cada vez de mayor atribución.

A continuación, se señalan varios formatos vectoriales ampliamente utilizados:

- Shapefile (SHP): formato SIG creado por la compañía ESRI, que se ha convertido en estándar de facto por muy utilizado y reconocido por la gran mayoría de softwares. No obstante, se trata de un formato bastante antiguo y presenta notables carencias, como, por ejemplo, la limitación de la longitud de los nombres de los atributos a 10 caracteres o no poder almacenar la hora en un campo de fecha (limitaciones impuestas por el propio formato Dbase).

  Almacena la localización, geometría y atributos asociados, pero no almacena información topológica. Se trata de un formato multiarchivo, entre los que se pueden destacar los siguientes:

  - Mínimo de tres archivos:

    .shp: entidades geométricas.

    .shx: índices de entidades.

    .dbf: información alfanumérica en formato Dbase.

  - Opcionales:

    .prj: sistema de coordenadas.

    .sbn y .sbx: índice espacial de las entidades.

    .ain y .aih: índice de atributo de los campos activos en una tabla o una tabla de atributos del tema.

    .xml: metadatos del shapefile.

- DWG (*drawing*): formato CAD desarrollado por AutoDesk y utilizado en su software AutoCAD.

- DXF (*Drawing eXchange File*): formato de intercambio para posibilitar la interoperabilidad entre los DWG y el resto de software del mercado.
- DGN (*design*): formato CAD desarrollado por Bentley y utilizado en su software MicroStation.
- GPX (*GPS eXchange Format*): formato basado en XML pensado para transferir datos GPS entre aplicaciones. Se puede usar para describir puntos (*waypoints*), recorridos (*tracks*) y rutas (*routes*).
- OSM (*OpenStreetMap*): formato basado en XML y utilizado en Open Street Map, proyecto colaborativo para crear mapas libres y editables. Los archivos OSM son descargables para ser utilizados.
- GeoPackage (*.gpkg*): formato basado en un único archivo que trabaja con el estándar de base de datos abierta SQLite. Se trata de un formato abierto, independiente de la plataforma y definido por OGC. Puede manejar datos vectoriales y ráster.

  Cuenta con una serie de ventajas:

  - Interoperabilidad
  - Trabaja con índices espaciales
  - Almacena datos complejos y voluminosos
  - Puede utilizarse desde multitud de software, incluso en dispositivos móviles.

2.3.1.4. Tablas y datos estructurados

En este apartado se hará referencia a diversos formatos de datos con las siguientes características:

- Se almacenan de modo estructurado, bien sea mediante tablas, bien sea mediante la definición de formatos específicos que de alguna manera estructuran el contenido y permiten que sea directamente legible por máquinas.
- Uno o varios de los campos de la tabla o formato estructurado contienen la geometría del objeto geográfico o hacen referencia a un objeto geográfico.

Archivos de texto

Los formatos más simples, legibles tanto por máquinas como por humanos, y generalmente abiertos, son aquellos basados en archivos de texto. Dentro de estos, podemos destacar los siguientes, por su uso generalizado en el ámbito de la información espacial:

- **CSV (Comma Separated Values)**: archivo en el que cada fila de texto representa un registro, y los campos de cada registro están separados generalmente por comas, puntos y coma o tabulaciones. No permite definir los tipos de dato relativos a cada atributo, debido a lo cual puede ser necesario un procesamiento intermedio si debemos integrarlos con datos procedentes de otras fuentes. Es habitual que este tipo de archivos se utilicen para almacenar elementos geográficos puntuales, de manera que uno de sus campos sea la latitud y otro la longitud del elemento en cuestión.
- **Formato ráster Esri ASCII**: habitualmente se utiliza para transferir información ráster. El archivo comienza con la información que define las propiedades del ráster: tamaño de celda, número de filas y columnas, y coordenadas de origen (esquina superior izquierda

del ráster). A continuación, se incluye el valor de la celda situada en el origen, y los valores de las siguientes celdas de la primera fila del ráster, separados por un espacio. Cada fila del ráster se incluye como una fila distinta del archivo.

- Aunque no sea propiamente un formato de texto, es importante mencionar el formato **Excel de Microsoft**, muy habitualmente utilizado de manera similar a los archivos csv. Se trata de una hoja de cálculo, en la que la información se presenta en filas y columnas, y ofrece diversas ventajas respecto al formato csv: por ejemplo, la posibilidad de incorporar un número de registros mucho mayor de forma comprimida, la facilidad para definir los tipos de datos de cada columna y las opciones de filtrado y realización de muy diversas operaciones sobre los datos. La principal contrapartida es que se trata de software propietario (si bien existen alternativas de software libre que mantienen las ventajas de este formato sin la desventaja del pago por su uso).

Archivos semiestructurados o de intercambio de datos

- El Lenguaje de Marcas geográficas (GML) es un sublenguaje de XML desarrollado por el OGC (Open Geospatial Consortium) para el modelado, transferencia y almacenamiento de datos espaciales, y aprobado como norma internacional ISO 19136 en 2007. Se trata de un formato de intercambio y almacenamiento de información geográfica muy versátil, pero poco eficiente para procesar y transferir datos, ya que los archivos de este tipo son mucho más grandes que otros formatos (ya que toda la información geográfica se almacena en texto), lo que por lo general impide que puedan utilizarse en aplicaciones web. A pesar de ello, es uno de los formatos estándar en el ámbito de la Directiva INSPIRE, debido a lo cual muchos de los organismos oficiales productores de cartografía en la Unión Europea generan datos en este formato (por ejemplo, la Dirección General del Catastro de España genera la información relativa a parcelas catastrales en este formato).

- KML (Keyhole Markup Language) es un lenguaje basado en XML que se utiliza para representar datos geográficos en tres dimensiones. Fue desarrollado para ser manejado con Keyhole LT (precursor de Google Earth, que en la actualidad admite este formato). Su gramática contiene muchas similitudes con la de GML.

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
  <Placemark>
   <name>Oviedo</name>
   <description>Ciudad de Oviedo, capital de Asturias</description>
   <Point>
     <coordinates>-5.843868,43.360758,0</coordinates>
   </Point>
  </Placemark>
 </kml>
```

Ilustración 2-71: Ejemplo de documento KML (fuente: https://es.wikipedia.org/wiki/KML)

- GeoJSON: es un formato de intercambio de datos geoespaciales basado en JSON, que surgió como alternativa al GML para modelar objetos geográficos sencillos. Es muy utilizado en aplicaciones web de cartografía, ya que permiten el intercambio de datos de manera ligera y fácilmente adaptable a las necesidades de cada aplicación.

```
"type": "FeatureCollection",
"features": [
  {
    "type": "Feature",
    "geometry": {
      "type": "Point",
      "coordinates": ["-80.014725","48.153552"]
    },
    "properties": {
      "active": "Yes",
      "operated_by": "Blanche River Health (Kirkland Lake site)",
      "city": "Kirkland Lake",
      "address": "145 Government Road East",
      "postal_code": "P2N 3P4",
      "province": "ON",
      "latitude": "48.153552",
      "longitude": "-80.014725",
      "phu_id": "1",
      "phu": "Timiskaming Health Unit",
      "phu_fr": "Services de sant\u00e9 du Timiskaming",
      "phone": "705-568-2127",
      "website": "http://www.blancheriverhealth.ca/brh_website/",
      "email": null,
      "monday": "09:00-13:00",
      "tuesday": "09:00-13:00",
      "wednesday": "09:00-13:00",
      "thursday": "09:00-13:00",
      "friday": "09:00-13:00",
      "saturday": null,
      "sunday": null,
      "after_hours": "Yes",
      "appointments": "Yes",
      "online_appointments": null,
      "phone_appointments": "Yes",
      "symptomatic": "Yes",
      "drive_through": "Yes",
      "french_language_services": null,
      "accessible": "Yes",
      "asl_interpretation": null,
      "first_nations": null,
      "free_parking": null,
      "paid_parking": null,
      "public_transit": null,
      "additional_information": "Please call to schedule an appointment. After hours,
      please call the hospital in advance for instructions. No walk-ins at the
      Emergency Department please.",
    }
```


Ilustración 2-72: Ejemplo de documento GeoJSON (fuente: https://open.canada.ca/data/en/dataset?collection=fgp&_res_format_limit=0&topic_category=geoscientific_information&keywords=Government+information&res_format=GEOJSON&res_format=GML)

Bases de Datos

Es importante señalar aquí algunas de las ventajas de la utilización de Bases de Datos Espaciales respecto a otros formatos:

- Facilidad para definir permisos de acceso a los datos con un alto nivel de granularidad (esto es, de forma específica para cada uno de los distintos roles participantes en un proyecto, tanto a nivel de base de datos, como de esquema, como de tabla).
- Existencia de mecanismos específicos que agilizan el tratamiento de los datos (funciones, triggers).
- Optimización de la velocidad de acceso a los mismos gracias al indexado de las tablas.
- Garantía de integridad de los datos, basada (entre otras cosas) en la gestión de claves primarias, claves externas y dominios de los atributos.

– Capacidad para almacenar y gestionar grandes volúmenes de datos

Dentro este apartado, merece la pena mencionar PostgreSQL, una base de datos relacional de código abierto y gratuita que, gracias a diversas extensiones (la más conocida de las cuales es PostGIS), permite gestionar objetos geográficos. Permite realizar tareas de análisis espacial tanto con datos vectoriales como con datos ráster (estos últimos desde la versión 2.0), así como análisis de redes sobre grandes volúmenes de datos vectoriales gracias a la extensión pgrouting.

2.3.2. Transformaciones

2.3.2.1. De formatos

Como se ha comentado, existe una multitud de formatos, tanto vectoriales, como ráster, además del papel u otros soportes analógicos. La información disponible se puede encontrar en cualquiera de ellos, por lo que a menudo, para poder utilizarla e incorporarla al entorno de trabajo deseado es necesario hacer transformaciones, bien entre modelos de representación distintos (ej.: de ráster a vectorial), bien entre formatos de archivo dentro del mismo modelo (ej.: de *shapefile* a *geopackage*).

La siguiente figura muestra un esquema de transformaciones a realizar al cambiar entre los modos de representación papel, ráster y vectorial.

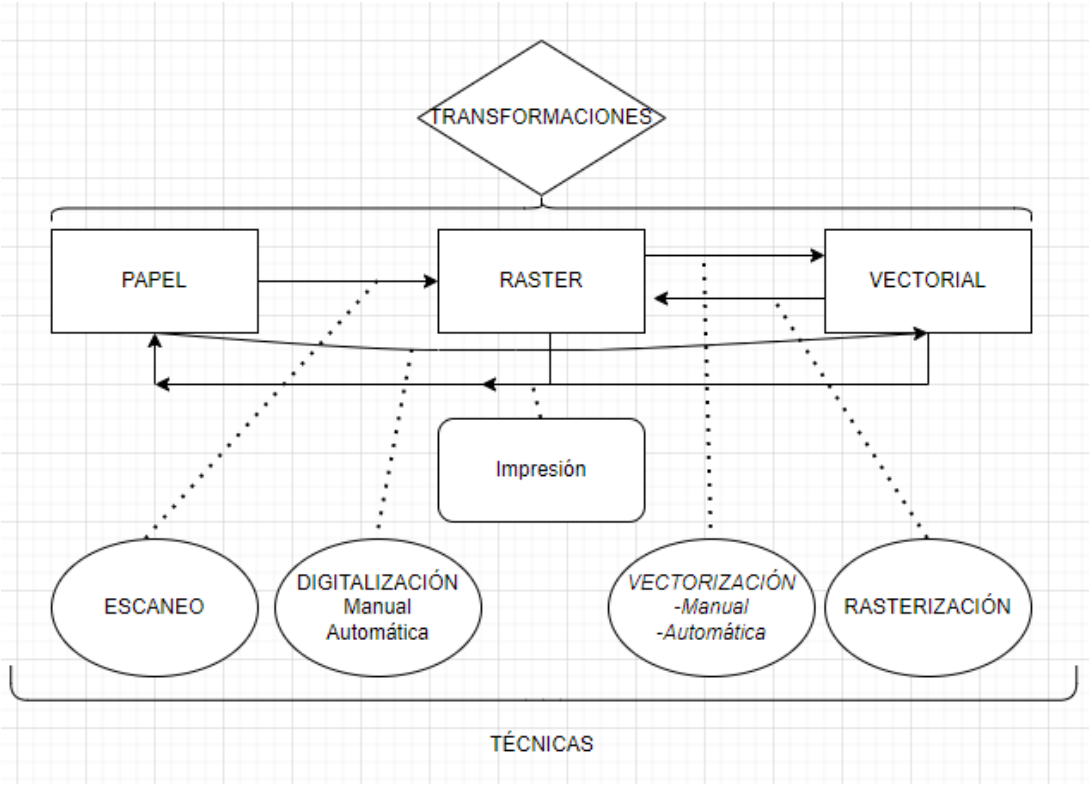


Ilustración 2-73: Esquema de transformaciones entre formas de entrada gráfica

Dentro de lo que serían los formatos digitales, el propio *software* de trabajo, sea de SIG, de CAD o especializado en ráster, suele disponer de importadores y exportadores de formatos. No obstante, sus capacidades en este sentido a veces son muy limitadas y depende mucho del programa concreto. Por otro lado, existen librerías como GDAL que contiene multitud de drivers

y de funcionalidades para la lectura/escritura de formatos y para transformaciones internas de los datos. A su vez, hay *software* especializado para este tipo de tareas, conocidos como *software* ETL (*Extract, Transform, Load*), por ejemplo, FME o GeoKettle.

A continuación, se proporcionan algunas indicaciones sobre técnicas de transformación entre modos de representación.

Impresión (ráster/vectorial -> papel)

Consiste en transformar una información o un documento, por ejemplo, un mapa, que se tiene en formato digital (vectorial o ráster) a formato analógico (papel).

Hoy en día, cualquier *software* de CAD, de SIG o de procesamiento de imágenes, es capaz de imprimir el mapa según los parámetros de impresión.

No obstante, se trata de una técnica de artes gráficas en la que se deben tener en cuenta varios aspectos para una representación óptima, según el objetivo que se persiga. Entre estos aspectos se puede citar el papel, la resolución, la escala, el tratamiento que se hace de la simbología, los colores a utilizar, etc. El capítulo 6 de la guía aborda con mayor grado de detalle estos aspectos.

Escaneado (papel -> ráster)

Básicamente sería el proceso contrario al de impresión, es decir, transformar documentación en papel (u otro soporte del estilo) en un formato digital ráster.

Es relativamente frecuente que se tenga que hacer esta operación, bien porque se disponga de fotografías analógicas o porque existan mapas en formato papel que interese transformarlos a formato digital para su mejor almacenamiento y difusión (por ejemplo, a través de servicios web), o para incluirlo en un programa informático, por ejemplo, en un SIG, de manera que pueda ser analizado y explotado en profundidad.

Para ello se utilizan escáneres, que convierten los tonos continuos del original en un archivo electrónico que representa la imagen de forma discretizada, tanto desde el punto de vista espacial como radiométrico. Es decir, divide la imagen en pixeles a los que les asigna un valor (o varios valores), dentro de un rango posible de valores, en función del tono original. Así, la calidad de la imagen final depende de múltiples factores, empezando por la calidad del documento original, los parámetros de escaneado y la propia capacidad del escáner utilizado.

Los parámetros más relevantes a tener en cuenta al escanear son los siguientes:

- Resolución (espacial) de la imagen escaneada, que se puede hablar de ella en términos de tamaño de píxel o de puntos por pulgada. No obstante, se debe tener en cuenta el grano de la imagen analógica original.
- Densidad (mínima, máxima y rango o volumen), que está relacionada con la capacidad de distinguir, sin saturarse, entre valores muy claros o entre valores muy oscuros. Esto último suele ser especialmente interesante para poder discriminar elementos en zonas de sombra.
- Resolución radiométrica, que está relacionada con la discretización que se hace de los niveles de gris. Así, se suele hablar del número de bits que se emplean para almacenar cada píxel (ej.: si se emplean 8 bits, sólo se puede distinguir entre 256 tonos).

- Error en la posición de los elementos. Se puede expresar como error cuadrático medio o como errores máximos (a un determinado porcentaje). Es decir, el grado de distorsión o de desplazamiento que se produce en los elementos de la imagen. Esto es especialmente crítico en escáneres fotogramétricos, ya que se van a tomar medidas y a obtener coordenadas a partir de las imágenes.

Hay otros parámetros a tener en cuenta, como el tamaño del propio escáner (según el formato de los documentos a escanear), el número de bandas o de canales (por ejemplo, en color o pancromático) o la velocidad de escaneo (que puede llegar a ser determinante si el número de documentos que deben ser escaneados es muy elevado).

Así pues, existe cierta variedad de tipos de escáneres y la elección de uno u otro dependerá de la finalidad para la que vaya destinado. Por ejemplo, para el escaneo de fotogramas aéreos (tomados con cámaras métricas) se utilizan escáneres fotogramétricos, que son diseñados y construidos para lograr precisiones métricas muy altas, con tamaños de píxel muy pequeños y con alta capacidad para distinguir entre los diferentes niveles de gris extremos para no perder los detalles presentes en las imágenes. Hay que tener en cuenta que el objetivo es poder medir y obtener coordenadas de los elementos que se visualizan en las imágenes, sin perder resolución espacial (aunque dependerá del grano de la imagen analógica original) ni capacidad de discriminación de elementos (especialmente en zonas de sombra, etc.). A su vez, deben tener el tamaño adecuado para escanear fotogramas de 23x23 centímetros, que es el formato estándar utilizado en fotogrametría aérea analógica.

En cambio, si el objetivo es el escaneo de mapas y documentos de carácter histórico, no es tan crítico el tamaño de píxel, la resolución radiométrica ni la precisión métrica, pero el tipo de escáner puede estar condicionado por el formato físico del original. Por ejemplo, hay documentos que se encuentran en formato libro (ej.: un atlas), que no pueden abrirse completamente para ponerse planos, o que son especialmente delicados.

Vectorización (Ráster-> Vectorial)

Consiste en transformar datos que están en un modelo de representación ráster a modelo vectorial. Este apartado se centra en los casos en los que los datos ráster son fotografías aéreas, imágenes de satélite o mapas en formato ráster, y se quiere extraer de ellos la información de interés, almacenándola en formato vectorial. No aborda, por ejemplo, las transformaciones de modelos digitales del terreno de malla regular a curvas de nivel o redes de triángulos irregulares (TIN).

El proceso lleva implícito una cierta abstracción y categorización de cuáles son los elementos de interés que están presentes en las imágenes y de qué forma se van a capturar, especialmente cuando las imágenes de partida son fotogramas aéreos o imágenes de satélite (serían de aplicación todas las indicaciones aportadas en el apartado 2.2).

Si el documento de partida es un mapa en formato ráster, la abstracción ya está implícita en el propio mapa, aunque es posible que no todos los elementos presentes en él sean de interés y haya que realizar una selección. También se deben definir ciertas normas sobre cómo categorizar y capturar las distintas líneas, símbolos y textos presentes en el mapa.

El proceso de vectorización habitualmente se realiza de forma manual. Normalmente, tanto los softwares de CAD como los de SIG, poseen funcionalidades para cargar las imágenes de las que

se va a extraer la información y herramientas diversas para capturar, en forma de puntos, líneas y polígonos, los elementos de interés.

Se debe señalar que la buena ejecución del proceso depende en gran medida de la experiencia del operador y de la calidad de las imágenes originales, especialmente cuando se trata de fotografías o de imágenes de satélite, ya que la propia fotointerpretación de los objetos geográficos presentes en ellas no siempre resulta evidente.

No obstante, a pesar de la complejidad que en ocasiones tiene la interpretación de los elementos presentes en las imágenes, a lo largo de los años se han ido desarrollando diferentes técnicas para la extracción automática de la información. Así, por un lado, se puede hablar de técnicas de seguimiento de líneas, que resultan de ayuda en la vectorización de mapas en modelo ráster. No obstante, la diversidad de formas, cruces y tipología de líneas (y de otros elementos) que habitualmente están presentes en un mapa, hace que estas técnicas no puedan utilizarse de modo 100% automático.

Por otro lado, se deben señalar los métodos de clasificación de imágenes, que las simplifican en una serie de clases (con un significado). A partir de estas clases, resulta más sencillo la determinación de polígonos y su posible colapso a líneas o puntos, según la morfología de los elementos, creando así los correspondientes objetos vectoriales.

Las técnicas que se emplean en este campo son muy variadas. En los últimos años, los avances en inteligencia artificial están revolucionando las posibilidades de interpretación automática de la información presente en imágenes. En particular, las técnicas de Deep Learning resultan muy ventajosas para la identificación de elementos en datos ráster, si bien es cierto que necesitan un volumen ingente de imágenes de entrenamiento, en las que se le indica si el objeto de interés está presente o no, para que el sistema vaya aprendiendo a discriminar cuándo ese determinado objeto se encuentra en nuevas imágenes.

Digitalización manual (Papel-> Vectorial)

En este caso, el proceso consiste en pasar directamente la información que se encuentra en un soporte físico como el papel, u otro similar, a un modelo de representación vectorial. Se van a mencionar aquí dos mecanismos que han sido muy utilizados durante bastantes años: las tabletas digitalizadoras y los restituidores fotogramétricos (analógicos y analíticos).

Las tabletas digitalizadoras pueden ser adecuadas para extraer la información de interés de un mapa en papel. Se trata de una superficie plana a modo de atril, sobre la cual se sitúa el documento cartográfico a digitalizar, y sobre este se van trazando las distintas entidades con un cursor. Este cursor registra los movimientos del operador, convirtiendo las posiciones del cursor en coordenadas reales, que son las que van a constituir la entidad digitalizada. Para ello, previamente, se señalan unos puntos de control de coordenadas conocidas, que permiten georreferenciar el documento.

Los restituidores fotogramétricos analógicos, en primer lugar, y analíticos, posteriormente, han sido ampliamente utilizados para producir cartografía. A partir de dos fotografías tomadas con cámara métrica y que capturan parcialmente la misma zona de terreno, permiten recrear la posición relativa que tenía la cámara (punto de vista y orientaciones) en el momento de cada disparo. De esta forma, para cada punto del terreno visible en los dos fotogramas, se pueden reconstruir las líneas que unen dicho punto terreno con su plasmación en ambas fotografías y

pasando por el punto de vista de cada disparo de la cámara. Esto permite determinar las coordenadas planimétricas y altimétricas de cualquier punto visible en las dos imágenes y restituir así, en modelo vectorial y en tres dimensiones, los elementos del terreno que son de interés.

Hoy en día, las capacidades de almacenamiento y de procesamiento de datos que proporcionan los sistemas informáticos han hecho que, en la transformación de formatos tipo papel a modelos vectoriales, sea preferible pasar primero por un modelo ráster, de manera que el flujo de normal pasa por escaneo y vectorización en pantalla. A pesar del inconveniente de introducir este proceso adicional, las ventajas que ofrece son múltiples. Se pueden destacar las siguientes:

- Si bien es cierto que se requiere de un escáner para conseguir el modelo ráster, en conjunto el hardware necesario es menos costoso y más fácil de mantener. Así, las tabletas digitalizadoras dejan de existir y sus funciones pueden realizarse con cualquier ordenador personal. Los restituidores analógicos y analíticos dejan de utilizarse en favor de las estaciones fotogramétricas digitales.
- Posibilidad de introducir técnicas de análisis digital de imágenes. Por ejemplo, en fotogrametría digital se utilizan, desde hace muchos años, métodos de búsqueda de puntos homólogos entre imágenes.
- Facilidad para cambiar de zona de trabajo. En las tabletas digitalizadoras y en los restituidores, cada vez que se quiere cambiar de zona, hay que cambiar físicamente los documentos y marcar los puntos necesarios para georreferenciar el mapa papel o para conseguir la orientación de los fotogramas.
- Facilidad para distribuir los trabajos de vectorización entre distintos operadores (y que pueden encontrarse en localizaciones alejadas).

Rasterización (Vectorial -> Ráster)

Consiste en transformar datos que se encuentran en modelo vectorial a un modelo ráster. Esta operación es, generalmente, más sencilla que la vectorización.

Una de sus principales aplicaciones es la representación de los datos vectoriales. Véase, por ejemplo, que esta operación se realiza implícitamente cada vez que se renderiza en pantalla un conjunto de datos vectoriales. También, la preparación de salidas para impresión suele pasar por un formato ráster.

Se pueden diferenciar dos fases en esta operación. La primera sería la definición de las propiedades de la retícula: paso de malla o tamaño de píxel, orientación, número filas y columnas, etc. La segunda fase consistiría en verter la información vectorial sobre esa retícula y, para esto, se deben definir qué valores imprimen sobre ella los diferentes elementos vectoriales. Por ejemplo, en caso de que el objetivo del ráster de salida sea una representación visual de los datos, hay que especificar los colores, grosores, símbolos, patrones de relleno y demás propiedades con que se van a mostrar los distintos elementos.

Otra utilización frecuente de la rasterización es la plasmación en modelos de malla regular de una determinada variable, por ejemplo, temperaturas, precipitaciones o elevaciones, a partir de un conjunto de datos vectoriales y utilizando algún método de interpolación para determinar el valor que debe tomar cada punto de la malla. En cartografía, son de especial interés los Modelos Digitales de Elevaciones (MDE) que se generan normalmente a partir de nubes de puntos, con

mayor o menor densidad, y de líneas que definen quiebros o estructuras del terreno o de la superficie de interés. La ventaja de la utilización de estos modelos en malla regular estriba principalmente en la capacidad que proporcionan para realizar análisis sobre ellos. Por ejemplo, un análisis de pendientes, de cuencas visuales o de cuencas vertientes es relativamente sencillo de realizar sobre un modelo ráster.

2.3.2.2. Transformación de coordenadas y georreferenciación

En este apartado se aborda la necesidad que hay, en algunas ocasiones, de transformar un conjunto de datos que se encuentra en un espacio de coordenadas que no tiene referencia de cuál es su localización sobre la Tierra, a otro espacio que sí esté vinculado a la Tierra en un Sistema de Referencia (SR), tanto geodésico como de coordenadas (proyección), determinado. Por ejemplo, en el caso de imágenes, se debe pasar del espacio imagen (columna, fila), al espacio del SR.

Para ello se utilizan funciones de transformación. Estas funciones tienen una serie de parámetros cuyos valores deben determinarse previamente. La forma habitual de determinar estos parámetros es a través de un conjunto de puntos con coordenadas conocidas en ambos espacios de coordenadas. Es conveniente que estos puntos se distribuyan y abarquen toda el área de interés. A partir de ellos, se realiza un ajuste, normalmente de mínimos cuadrados, para determinar la mejor estimación de los parámetros de la transformación.

Existen diferentes tipos de transformaciones. Habitualmente se utilizan las polinómicas, con un número mayor o menor de parámetros, en función de la necesidad o no de modificar la escala y las formas del espacio de coordenadas. Algunas de las más utilizadas y que suelen venir implementadas en software de SIG son las siguientes (se refieren siempre a transformaciones bidimensionales):

- Transformación Helmert:
  - De cuatro parámetros.

$$x' = ax - by + c$$

$$y' = bx + ay + d$$

  - Conserva los ángulos. Por tanto, no permite un escalado diferente entre ejes ni un cambio en el ángulo que forman.
- Transformación afín:
  - De seis parámetros.

$$x' = a_0 + a_1 x + a_2 y$$

$$y' = b_0 + b_1 x + b_2 y$$

  - Permite un escalado diferente entre ejes y un cambio en el ángulo que forman. Por tanto, no conserva ángulos, pero la escala a lo largo de cada eje se mantiene constante.
- Transformación polinómica de doce parámetros:

- Existen transformaciones polinómicas de diferente número de parámetros, pero una de las habituales es las de 12.

$$x' = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy$$

$$y' = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy$$

- Permite variar el escalado y las formas a lo largo de los ejes.

- Transformación proyectiva:
  - De ocho parámetros.

$$x' = \frac{a_0 + a_1 x + a_2 y}{1 + c_1 x + c_2 y}$$

$$y' = \frac{b_0 + b_1 x + b_2 y}{1 + c_1 x + c_2 y}$$

  - Permite variar la escala a lo largo de los ejes. Resulta útil, por ejemplo, en el caso de una fotografía que se ha tomado inclinada con respecto al plano de referencia y, por tanto, tiene una escala distinta según la zona de la imagen.

Como puede deducirse, a más parámetros, se necesitan más puntos comunes para poder determinar los parámetros de la transformación (el número de incógnitas es mayor) y ésta tiene más capacidad para transformar las características del espacio de coordenadas. Esta capacidad puede resultar muy útil en algunas ocasiones, como en el caso que se ha señalado en la transformación proyectiva, pero es importante conocer este hecho. Es decir, no se debe caer en el error de incrementar el número de parámetros para disminuir los residuos en los puntos utilizados para hacer la transformación, ya que puede provocar alteraciones no deseadas y errores notables en zonas alejadas a los puntos comunes.

En este sentido, es muy importante que no se empleen transformaciones de más parámetros para absorber las diferencias entre sistemas de proyección. Por ejemplo, en el caso de georreferenciar un mapa en formato imagen, se debe realizar en el mismo SR del mapa (especialmente en la misma proyección), y lo recomendable es utilizar transformaciones de pocos parámetros, como la Helmert o la afín, aunque en ocasiones puede ser necesario una proyectiva. Una vez georreferenciado, se puede transformar al SR de trabajo, pero utilizando ya las fórmulas propias de los sistemas de proyección (y de referencia geodésicos, si fuese el caso) implicados.

Otro aspecto que se debe reseñar es que, en la transformación de imágenes siempre va a ser necesario realizar, en el paso final, un remuestreo de la misma. Es decir, la malla regular de salida va a ser diferente a la original y debe establecerse un criterio para determinar el valor de los píxeles de destino en función de los originales. Aunque existen diversos métodos de remuestreo, los tres que se utilizan tradicionalmente son los siguientes:

- Vecino más próximo:
  - El valor que se asigna a cada píxel es el que tiene el píxel más cercano en el ráster original, una vez transformado.
  - Es el método más rápido en cuanto a cálculo.
  - Puede producir cierto efecto de escalonamiento en el ráster resultante.

  - Conserva los valores originales de los píxeles, hecho que es importante en el caso de que estos valores tengan un significado (por ejemplo, una imagen clasificada).

- Bilineal:
  - El valor que se asigna a cada píxel es fruto de una interpolación bilineal con los cuatro píxeles más cercanos.
  - Requiere un poco más de tiempo de cálculo que el vecino más próximo.
  - La imagen resultante queda suavizada (como si se hubiese pasado un filtro de paso bajo).
  - No se conservan los valores originales de los píxeles, por lo que no se debe emplear si estos valores tienen un significado.
- Convolución cúbica:
  - En este caso se tienen en cuenta los 16 píxeles más cercanos al píxel de salida, realizando una media ponderada en función de la distancia.
  - Requiere más tiempo de cálculo que la bilineal.
  - Consigue buenos resultados en el caso, por ejemplo, de remuestrear fotografías.
  - Al igual que en la bilineal, no conserva los valores originales de los píxeles, por lo que no se debe emplear si estos valores tienen un significado.

2.3.2.3. Transformación de modelo de datos

En este apartado se va a tratar la necesidad que surge en muchas ocasiones de transformación de datos para adaptarlos desde un modelo lógico a otro. Como se ha comentado previamente, un modelo parte de una abstracción de la realidad para poder trabajar con ella de forma simplificada. Pero la realidad es compleja y, dependiendo de la finalidad que se persigue, se pueden hacer multitud de abstracciones y definir, por tanto, muchos modelos diferentes.

Partiendo de esta base, para reutilizar conjuntos de datos existentes, en la mayoría de ocasiones se hará necesario transformar esos datos desde el modelo original en el que se encuentren al modelo de trabajo que se haya definido.

Estas transformaciones afectarán de manera frecuente a determinadas características del dato como pueden ser:

- **Geométricas**: un mismo elemento de la realidad puede estar modelado con tipo de geometría o con otro. Por ejemplo, un aerogenerador puede representarse mediante una geometría poligonal en un determinado modelo y como una geometría puntual en otro modelo.
- **Atributos**: un mismo elemento de la realidad puede modelarse con una serie de atributos que definan sus características importantes para una finalidad y, a su vez, en otro modelo, ese mismo objeto geográfico será modelado con unas características distintas. En este sentido, se pueden distinguir tres tipos de diferencias:

- Diferentes características (atributos): según la finalidad, interesan unas características u otras. Siguiendo con el ejemplo del aerogenerador, si es modelado por un equipo que se dedique a la conservación de aves, le interesará tener la característica de la altura de máxima de las aspas de ese aerogenerador con el fin de saber posibles interacciones con la migración de las aves. En cambio, si es modelado por una compañía eléctrica, seguramente le interese más la potencia que es capaz de generar.
  En estos casos no se puede establecer una equivalencia en el cambio de modelo.
- Tipo de dato del atributo: atributos que se refieren a la misma característica pueden estar modelados con tipos de datos distintos. Por ejemplo:
  - La característica altura de la presa puede estar modelada con un atributo numérico de tipo *float* en un modelo de datos y, en otro, estar como tipo *integer*. En este caso se puede establecer una conversión de tipos, aunque en el paso de *float* a *integer* se debe asumir la pérdida de decimales.
  - La característica nombre del objeto puede estar modelada como texto de una longitud determinada y, en otro modelo, tener una longitud distinta. En caso de trasladar información de un modelo a otro, se puede producir un truncamiento en la cadena de texto.
- Lista controlada: en ocasiones, los valores de un determinado atributo están restringidos a una lista de valores cerrados, mientras que, en otro modelo, esa misma realidad geográfica puede estar restringida a otra lista de valores, más general o más específica, o diseñada para un ámbito de aplicación distinto.
  Por ejemplo, si en un modelo determinado, en el objeto geográfico edificios existe un campo tipo con la lista de valores *unifamiliar*, *multifamiliar*, *comercial*, *industrial*, mientras que en otro modelo se definen los valores *residencial*, *comercial*, *industrial*, para pasar del primero al segundo se debe hacer una agregación, de forma que se agrupen los valores *unifamiliar* y *multifamiliar* en el valor *residencial*. En cambio, en la transformación inversa no habría forma de distinguir entre los dos valores posibles y se debe tomar una decisión para establecer cómo actuar en esos casos.

- **Cardinalidad:** una misma realidad, a pesar de estar representada con el mismo tipo de geometría (poligonal, lineal o puntual) puede recogerse como una única geometría, en un solo registro, o como un conjunto de ellas, por algún criterio que las divida. Por ejemplo, un rio puede representarse como una línea y en un único registro en un modelo dado, mientras que en otro estar como una sucesión de tramos que empiezan y finalizan cuando a ese río le llegan afluentes. En este segundo caso, el río tendrá *n* tramos, todos ellos con un mismo nombre.
- **Relación**: Un objeto de la realidad puede modelarse en una única tabla, con múltiples atributos, o en varias, relacionándolas a través de un identificador. Por ejemplo, si una organización es la encargada de capturar y mantener la geometría de los núcleos de población y otra organización es la encargada de mantener sus características (habitantes, número de hospitales, número de escuelas, etc.), una forma de modelar es definiendo una relación entre la geometría almacenada en una tabla y los atributos almacenados en otra tabla distinta e independiente, relacionándose a través de un identificador común. En cambio, una tercera otra organización puede estar interesada

y requerir estos datos de forma conjunta, en una única tabla, por lo que debe realizar la unión de los dos conjuntos de datos.

2.4. Información sin coordenadas explícitas: geocodificación de direcciones

La información que se precisa para realizar una correcta clasificación temática de los objetos geográficos capturados o, incluso para dotarles de una denominación oficial, también puede ser obtenida a partir del tratamiento de los datos de direcciones postales o procesos de geocodificación.

El objetivo de la geocodificación es asignar coordenadas geográficas a puntos del territorio, descritos por listado de direcciones. Estas coordenadas nos permitirán integrar estas entidades en un sistema de información geográfica y resolver diferentes problemas que se nos planteen. Para poder llevar a cabo esta asignación se ha de establecer una correspondencia entre un registro del listado de direcciones y un callejero de referencia que contenga la ubicación de cada portal mediante el emparejado de direcciones registro-callejero, lo que se conoce como *address matching*.

Existen dos tipos de geocodificación, la directa, por la cual asignamos coordenadas a una dirección dada o a un lugar y la inversa, en la que se asigna una dirección postal o el topónimo de un lugar a una localización dada por unas coordenadas.

Es importante que exista coherencia entre el callejero que utilice el servicio de geocodificación y los datos de direcciones que se pretenden geocodificar, para que se pueda establecer correctamente el emparejamiento entre ambos conjuntos de datos. Si este requisito no se cumple puede darse un alto índice de incertidumbres, de faltas de coincidencias o una precisión insuficiente en el posicionamiento obtenido.

Los problemas que nos podemos encontrar a la hora de realizar la geocodificación pueden ser diversos, siendo los más frecuentes los siguientes:

- datos sobre los que no se ha aplicado ningún tipo de normalización. Suelen presentar ausencia de una estructuración adecuada y conocida o hacen un uso muy libre de artículos, preposiciones y abreviaturas y, por norma general, no mantienen ninguna vinculación con un nomenclátor de referencia oficial.
- falta de actualización de las direcciones y de comprobación de su veracidad. En este sentido es importante señalar la gran relevancia que adquiere el mantener una política de actualización y aseguramiento de calidad de las direcciones por parte de las administraciones públicas responsables, que evite la degradación de los registros.

Ilustración 2-74: Ejemplo de diferentes nomenclaturas de calles

2.4.1. Normalización de las direcciones postales

Una dirección postal se puede definir como una jerarquía de componentes, siendo los más frecuentes:

- Unidades administrativas (región, provincia, etc.)
- Nombres geográficos, generalmente vinculados a agrupaciones de asentamientos de población, con un creciente nivel de detalle (municipio, entidad singular, núcleo de población, etc.).
- Tipo de vía (Calle, Avenida, Plaza, etc.)
- Nombre de la vía
- Localizador, o identificador de mayor detalle dentro de los componentes de la dirección, que generalmente se corresponde con el número o el nombre del inmueble.
- Identificador postal

Con la evolución de las tecnologías, en los últimos años se ha incrementado notablemente el número de dispositivos que, de forma generalizada y masiva, utilizan aplicaciones que precisan de datos de localización, entre ellos las direcciones. Ello, sumado a los casos tradicionales como el geomarketing, los servicios de distribución postal o la localización de los servicios públicos y emergencias, pone de manifiesto la necesidad extendida de disponer de datos completos, fiables y precisos. Por tanto, disponer de una base de datos de direcciones para su uso como callejero de referencia, actualizada y con información fiable, es estratégico para optimizar los procesos y la gestión de todas estas actividades, y garantizar la consecución de los objetivos perseguidos.

El proceso de la **normalización de direcciones postales** persigue la armonización e interoperabilidad de los datos, e implica corregir, transformar y estructurar los datos para que puedan ser utilizados por aplicaciones de valor añadido como las que requieren procesos de geocodificación; el éxito de estos procesos vendrá determinado por la buena calidad y estructura de la información que consigamos con la normalización.

Causas de la problemática. El tratamiento de las direcciones postales es un proceso complejo motivado fundamentalmente por la existencia de múltiples proveedores de datos que emplean diferentes modos de recopilación, siendo habitual encontrar conjuntos de datos heterogéneos, poco o nada estructurados y con diferentes grados de precisión.

farmacias	direccion
Farmacia Mercedes Romano Garcia	Calle La Roche Sur Yon 7, 10005, Cáceres
Farmacia Francisco Javier Mateos Vicente Y Elvira Mateos Gonzalez	Calle Doctor Fleming 3, 10001, Cáceres
Farmacia Jose Javier Mogollon Carrasco	Avda De Alemania 8, 10001, Cáceres
Farmacia Maria Antonia Sanchez Gundin Y Maria Cristina Gonzalez Sanchez	Calle Rio Danubio 14, 10005, Cáceres
Farmacia Santiago Rodriguez Briega	Avda Isabel De Moctezuma 1, 10005, Cáceres
Farmacia Jose Gallego Caboblanco Y Jose Gallego Rodriguez	Avda Juan Solano Pedrero 13, 10005, Cáceres
Farmacia Juan Francisco Garcia De Casasola Garcia	Avda Antonio Hurtado 22, 10002, Cáceres
Farmacia Luis Conde Garcia	Calle Islas Azores 15, 10001, Cáceres
Farmacia Maria Isabel Torres Perez	Avda Virgen De La Montaña 23, 10004, Cáceres
Farmacia Francisco Javier Sanguino Sanchez	Calle Padre Bayle 9, 10004, Cáceres
Farmacia Maria San Pedro Ledesma	Calle San Anton 7, 10004, Cáceres
Farmacia Josefina Becerra Muñoz	Calle Amberes 6, 10003, Cáceres

Ilustración 2-75: Tabla ejemplo de direcciones sin normalizar de farmacias en Cáceres

En relación a la estructuración de los datos, es frecuente encontrar que la información viene desagregada en una serie de campos o componentes de los que habrá que establecer su correspondencia con los recomendados en los estándares.

	A	B	C	D	E	F
1	n_mun	dir_tipo_via	Dir_Via	dir_cp	n_mun	N_PROV
2	Ababuj	C/	Iglesia, 8	44155	Ababuj	TERUEL
3	Abades	C/	Santo Cristo, 1	40141	Abades	SEGOVIA
4	Abadía	Plaza	de España, 3	10748	Abadía	CÁCERES
5	Abadín	C/	General Franco, 29	27730	Abadín	LUGO
6	Abadiño		San Trokaz Plaza, s/n	48220	Abadiño	BIZKAIA
7	Abáigar	C/	San Vicente, s/n	31280	Abáigar	NAVARRA
8	Abajas	Ctra.	Villarcayo, s/n	09141	Abajas	BURGOS
9	Ábalos	Plaza	Martín Fernández de Navarrete, 1	26339	Ábalos	LA RIOJA
10	Abaltzisketa	Plaza	de San Juan, s/n	20269	Abaltzisketa	GIPUZKOA
11	Abánades	Calle	Real,1	19432	Abánades	GUADALAJARA

Ilustración 2-76: Tabla ejemplo de direcciones normalizadas por campos

Respecto a la heterogeneidad y precisión de los datos, es habitual encontrar los problemas que a continuación se enumeran (recopilados en base a la experiencia en el tratamiento de las direcciones en España, que se estima son también frecuentes en otros países):

- Empleo de distintas nomenclaturas en el tipo de vía (Ej. “C.”, “C/”, “CL”: todas hacen referencia a la misma entidad: “CALLE”).
- Uso de abreviaturas no estandarizadas en el nombre de las calles (Ej.: Calle José María Castro, Calle J. Mª. Castro, Calle J.M. Castro, Calle J. María Castro, Calle José M. Castro).
- Empleo de caracteres especiales (Ej.: acentos y diéresis).
- Utilización del nombre del inmueble o del lugar en lugar de su dirección (Ej.: Iglesia de Santa María).
- Aplicación de mismas denominaciones (tipo vía + nombre de vía) pero que están situadas en poblaciones distinta, lo que provoca que las direcciones no sean unívocas.

Proceso. El tratamiento de mejora de la calidad de las direcciones postales puede resumirse como un proceso que: 1) identifica direcciones o registros muy difíciles, casi imposibles de localizar, 2) corrige direcciones erróneas (ej.: datos que provocan devoluciones de correo o que servicios públicos no lleguen al lugar concreto del suceso) y 3) verifica y actualiza los registros.

Para realizar la normalización se pueden emplear herramientas que ayuden a estructurar, optimizar y normalizar la información de la dirección postal; si no se dispone de ellas habrá que analizar el conjunto de datos y extraer las diferentes casuísticas que se presenten y crear una serie de patrones o criterios de estandarización y unas tablas de equivalencia, que aplicaremos sobre el conjunto de datos para tratar de homogeneizar los registros.

El proceso de normalización se compone fundamentalmente de una primera fase de estandarización y una segunda de segmentación o desagregación.

En la <u>fase de estandarización</u> se depuran los registros y se asignan valores estandarizados. Esta fase es primordial para garantizar homogeneidad en el contenido de los datos, y en ella se:

- Unifica el contenido de los registros en letras mayúsculas
- Asignan palabras completas e inequívocas a las distintas nomenclaturas de tipos de vía que aparecen (ej.: “C.”, “C/”, “CL”, “c”,” c.”, “c/”, “cl” =” CALLE”), y a las distintas abreviaturas que se presenten (ej.: “Dña.” = “Doña”, “Pdte.” o “Pte.” = “Presidente”, “Dr.” o “Doc.” = ”Doctor”).

- Eliminan tildes y partículas (ej.: Avenida de las Naciones, Avenida Naciones (de las), Avenida Naciones) que puedan interferir en la búsqueda posterior de la dirección o en cruces con otras bases de datos disponibles de direcciones.

En la segunda fase de segmentación o de desagregación, se estructura la información en cada uno de los componentes que conforman la dirección para que estos puedan ser posteriormente usados en servicios de geocodificación. Esta fase es fundamental, ya que cuanto más estructurada reciba la información el geocodificador más sencillo le resultará al algoritmo identificar el resultado correcto. A continuación, con un ejemplo de una dirección clásica en España, se ilustra cómo una misma dirección (el primer registro es el correcto) puede tener muchas representaciones según las variantes en su estructura, el orden de especificar sus componentes o la manera de escribir el contenido de cada uno de ellos.

Tipo de Vía	Nombre de Vía	Número	Provincia	Municipio	Cód. Postal
Calle	Juan Montilla	12	Jaén	Úbeda	23400
Calle Juan Montilla, 12, Úbeda, Jaén					
Calle Juan Montilla, 12			Úbeda, Jaén		
Calle	Juan Montilla, 12, Jaén			Úbeda	
Juan Montilla 12, Úbeda, Jaén					
12, Juan Montilla, Úbeda, Jaén					

C/ Juan Montilla Nº 12, Úbeda (Jaén)
Call. Juan Montilla S/N, Úbeda, Jaén
C/ J. Montilla 12, Úbeda Jaén
C. J. Montilla (de) 12, 23400, Jaén

Ilustración 2-77: Ejemplo de variantes de representación de una dirección: arriba, según la estructura y el orden de los componentes, y abajo, según el contenido de estos.

A la hora de normalizar las direcciones puede utilizarse como referencia el modelo de datos definido en las Especificaciones de Direcciones para el desarrollo de la Directiva 2007/2/CE, por la que se establece una Infraestructura de Información Espacial en Europa (INSPIRE), de acuerdo a la cual, una dirección postal se compone de: nombre de las unidades administrativas (*AdminUnitName*), nombre de entidad inferior a la unidad administrativa de mayor detalle (*Address Area Name*), nombre de la vía (*ThoroughfareName*), localizador (*Locator*), elemento postal (*Portal Descriptor*).

El orden de estos componentes en una dirección postal debería seguir un determinado esquema, por ejemplo: tipo de vía, nombre de la vía, número, nombre de la unidad administrativa (provincia y municipio) y finalmente el elemento postal. Este esquema variará dependiendo de la naturaleza de los datos que encontremos en cada país, siendo flexible en su estructura y en los elementos que lo componen.

2.4.2. Geocodificación: callejeros y geocodificadores

Cada servicio de geocodificación utiliza como referencia un callejero determinado que, en resumen, es una base de datos geográfica formada por un conjunto direcciones, es decir, registros cuyos atributos son los componentes de éstas, a los que se les ha dotado de una geometría puntual asociada generalmente al localizador o componente de mayor detalle de la dirección. Es frecuente que esta ubicación espacial se corresponda con la entrada acceso al edificio o con el centroide de la parcela catastral.

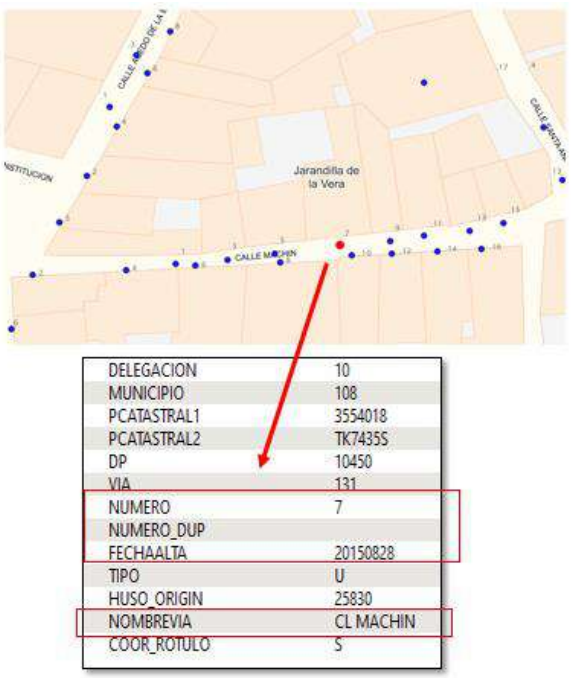


Ilustración 2-78: Ejemplo de datos completos de un registro de dirección. Fuente Catastro (España)

La calidad y eficacia del callejero de referencia vendrá determinada por la completitud y la rigurosidad de los atributos de la dirección, no debiendo existir registros con nombres de vía vacío (NULL), del tipo SIN NOMBRE o NO IDENTIFICADO, o con errores ortográficos, ni con ausencias de tipos de vía, numeración de vías y códigos postales.

Por ello, cómo ya se comentó anteriormente, será de vital importancia que exista la máxima coherencia, armonización e interoperabilidad entre este callejero de referencia y los datos de direcciones para, en cualquier caso, tratar de minimizar las disparidades potenciales entre los datos a procesar.

Por tanto, la característica más importante de un servicio de geocodificación es su flexibilidad, de manera que permita obtener buenos resultados incluso a partir de datos de entrada poco estructurados o inciertos. No debe ser exigente en los requisitos del formato de entrada de manera que una simple coma, guion o tilde haga fracasar el emparejamiento, y debe ser capaz de entender las diferentes tipologías de vías, independientemente de cómo se expresen, y de analizar y estructurar los principales componentes de una dirección (nombre de la vía, numeración y población), lo que técnicamente se denomina *parseado.*

Ejemplo de funcionamiento de un servicio de geocodificación de direcciones

Se detallan a continuación, y brevemente, las fases principales que explican el funcionamiento del proceso de un servicio de geocodificación estándar.

- **Análisis de la dirección solicitada**

Divide la dirección en sus componentes basándose en un “estilo de direcciones” que debería ser conforme con un estándar. Para aumentar las posibilidades de encontrar la dirección, el localizador de direcciones podrá realizar diferentes divisiones de elementos.

Ejemplo: C/ de Gral. Ibáñez de Ibero, nº3, Madrid, 28003

Tipo de Vía	Nombre de Vía	Número	Provincia	Municipio	Cód. Postal
Calle	Gral. Ibáñez de Íbero	Nº 3	Madrid	Madrid	28003
C/de Gral. Ibáñez de Ibero		3	Madrid	Madrid	28003

Ilustración 2-79: Estructuración de componentes de la dirección

- **Creación de varias representaciones de la dirección**

Muchos elementos de una dirección a menudo se escriben con una abreviatura que pueden presentarse en varios formatos. En ocasiones los artículos, las preposiciones y los guiones pueden sobrar. Por tanto, en esta fase se debe establecer una lista de valores para cada componente de la dirección.

	Posibles Valores			
	Valor 1	Valor 2	Valor 3	Valor 4
Tipo de Vía	Calle	c/	Carrer	C.
Nombre de vía	Gral. Ibáñez de Íbero	de Gral Ibañez de Ibero	General Ibáñez Íbero	General Ibañez Ibero
Número	3	Nº3		
Provincia	Madrid			
Municipio	Madrid			
Código Postal	28003			

Ilustración 2-80: Definición de la lista de valores para cada componente de la dirección

- **Búsqueda en el callejero de referencia**

Una vez analizada la dirección y generada una lista de valores para cada componente se obtiene una lista de representaciones de la dirección solicitada (ej.: solo se muestran algunas):

LISTADO DE REPRESENTACIONES DE LA DIRECCION SOLICITADA
Calle Gral. Ibáñez de Íbero, 3, Madrid, Madrid, 28003
Calle de Gral Ibañez de Ibero, 3, Madrid, Madrid, 28003
Calle General Ibañez Ibero, nº 3, Madrid, Madrid, 28003
C/ Gral. Ibáñez de Íbero, 3, Madrid, Madrid, 28003
C/ de Gral Ibañez de Ibero, 3, Madrid, Madrid, 28003
C/ General Ibañez Ibero, 3, Madrid, Madrid, 28003
C/ General Ibáñez Íbero, nº3, Madrid, Madrid, 28003
C/ General Ibañez Ibero, nº 3, Madrid, Madrid, 28003
c. Gral. Ibáñez de Íbero, 3, Madrid, Madrid, 28003
c. de Gral Ibañez de Ibero,nº 3, Madrid, Madrid, 28003
c. General Ibáñez Íbero, nº3, Madrid, Madrid, 28003

Ilustración 2-81: Ejemplos de variantes de representación de una dirección

Posteriormente, comienza la búsqueda de registros del callejero de referencia que concuerdan con los elementos de la dirección de entrada, según unos criterios de búsqueda. La concordancia entre ambos elementos se analiza como el grado de similitud entre las cadenas de caracteres de la dirección de entrada y las cadenas de caracteres de los registros del callejero de referencia. Para cuantificar esa similitud y detectar posibles candidatos se usan funciones de distancia como

puede ser la *Distancia Levenshtein* o la *Distancia de Hamming*, que indican el grado de concordancia entre ambas cadenas y, por tanto, facilitan la identificación de los resultados que mejor se ajusten a los parámetros de entrada.

- **Puntuación establecida de cada concordancia potencial**

Se realizan búsquedas sucesivas con criterios cada vez menos restringidos, comenzando por la búsqueda de similitud de todos los componentes al 100%, de modo que si hay éxito se ofrece el resultado del registro unívoco; si esto no sucede, se definen unos umbrales de similitud menos restrictivos (p. ej.: 95-85%) para que el sistema ofrezca una serie de potenciales candidatos. Posteriormente, se ordenan los candidatos atendiendo al grado de mayor similitud con la dirección de entrada.

- **Geocodificación del mejor candidato**

Por último, el geocodificador ofrece como resultado las coordenadas de aquel registro que presenta grado mayor de similitud respecto de los datos de entrada, seguido del resto potenciales candidatos para que el usuario finalmente decida cuál es el que mejor se ajusta a sus requisitos.

Servicios de Geocodificación existentes

A continuación, se citan algunos de los geocodificadores más utilizados cuya elección dependerá en gran medida del ámbito geográfico y de los requisitos del caso de uso que se trate (tipo de precisión, coste, formatos de salida, condiciones de software, etc.).

Nominatim es un servicio de geocodificación de uso libre que utiliza como callejero de referencia las direcciones y determinados objetos geográficos de la base de datos colaborativa OpenStreetMap. Dada la falta de homogeneidad de datos de direcciones de OpenStreetMap (sobre todo en ámbitos rurales hay una gran escasez de datos) el éxito y la precisión en la localización puede no ofrecer muchas garantías dependiendo de la zona de búsqueda.

Google Maps dispone de un servicio de geocodificación que permite realizar un número determinado de localizaciones al día. Para obtener un número mayor se debe utilizar la API Geocoding de Google Maps cuyo uso no es gratuito.

ArcGIS dispone de un servicio de geocodificación al que se puede acceder fácilmente disponiendo de una conexión a Internet y de las credenciales adecuadas, a través de ARCGIS Online.

OpenCage usan también las fuentes de OpenStreetMap, pero las mejoran con otros recursos de fuentes de datos abiertas.

A continuación, se citan dos herramientas de carácter local (España), a título de ejemplo y como referencia de buenas prácticas que pueden ser desarrolladas a nivel nacional en otros países.

Servicio de callejero y datos catastrales no protegidos y servicio de conversión de coordenadas de la dirección General de Catastro. Usando de forma combinada estos servicios WEB se puede acceder al callejero catastral (provincias, municipios, calles, números) y a los datos no protegidos de los inmuebles (todos excepto titularidad y valor), de manera que dando como parámetro de entrada una dirección postal se obtiene la referencia catastral de la finca correspondiente.

Posteriormente el servicio de conversor de coordenadas proporciona, a partir de esta Referencia Catastral, las coordenadas geográficas X e Y del centroide de la parcela, indicando el sistema de referencia en el que está almacenado el dato en la Dirección General de Catastro.

Instituto Geográfico Nacional, Proyecto CartoCiudad. CartoCiudad es un conjunto de datos de red viaria urbana e interurbana continua de ámbito nacional, generado a partir de datos oficiales. Su fuente de datos principal es la red viaria procedente de la Información Geográfica de Referencia de Redes de Transporte (IGR-RT) del IGN que a su vez recoge códigos de vías urbanas (callejero) del INE (Instituto Nacional de Estadística) y de la Dirección General de Catastro para dotar a todas ellas de oficialidad en cuestiones de nombres y tipologías de vías. Además, los datos de direcciones se complementan con la denominación de las poblaciones procedentes de la Información Geográfica de Referencia de Poblaciones del IGN (IGR-Poblaciones), los códigos postales de Correos y la denominación oficial de las unidades administrativas del Registro Central de Cartografía. Dentro del Proyecto Cartociudad se han desarrollado una serie de geoprocesos que siguen la especificación WPS del OGC y que operan sobre la información espacial georreferenciada de la que dispone la base de datos de Cartociudad (*https://www.cartociudad.es/portal/web/guest/calculos*). Entre estos procesos existe el de geocodificación, tanto directa como inversa, y que permiten realizar los procesos de manera individual o masiva. La geocodificación directa ofrece la posibilidad de geocodificar mediante un texto libre, en el que se introduce una dirección y se obtienen unos resultados más próximos a la búsqueda, o mediante un texto rígido en el que se debe insertar, obligatoriamente, todos los campos que componen la dirección.

Existen muchos otros servicios como pueden ser: la API de Yahoo Maps, la API de MapQuest, la API de Microsoft Virtual Earth, Pelias, HERE Geocoder REST API, Mapbox Geocoding API, Bing Maps Locations API, GeocodeEarth, GeocodeFar o MapBox.

Cada uno de ellos con sus propias características, condiciones de acceso, y datos de referencia que pueden estar delimitados a zonas determinadas. La mayoría de estos servicios son de pago, aunque muchos permiten usar cuentas gratuitas, conocidas como freemium, que tienen limitaciones con relación al número de peticiones por día o mes, tiempo (petición por segundo), restricciones de IP y CORS o soporte.

CAPÍTULO 3: INTEGRACIÓN DE INFORMACIÓN SEMÁNTICA

Ejemplos de fuentes de datos y técnicas de integración

Descripción breve

Técnicas para la actualización de datos respecto de fuentes de referencia por métodos de integración alfanumérica y conflación geométrica. Principales fuentes y servicios webs

Índice de contenidos

3.1. Integración basada en atributos alfanuméricos 1

3.2. Integración por criterios espaciales 4

3.2.1. Integración de datos superficiales 4

3.2.1.1. Selección de candidatos 4

3.2.1.2. Medida de la similitud 4

3.2.1.3. Operaciones para la mejora de la identificación 7

3.2.1.4. Combinación de entidades 8

3.2.2. Integración de datos lineales 8

3.2.2.1. Metodología a aplicar en la conflación de elementos lineales 9

3.2.2.2. Caso de uso: conflación en una red de caminos 10

Selección de la información de referencia 10

Diseño de proceso de transformación para la integración de la información semántica 11

3.2.2.3. Caso de uso: detección de cambios en redes de transporte 12

Datos de partida 12

Proceso de conflación 13

Resultados obtenidos 17

3.3. Fuentes de datos en la red utilizables en la actualización 18

3.3.1. Bots: extracción de datos por técnicas de Web Scraping 18

3.3.1.1. Fundamento de Web Scraping. Tipologías 18

Web Scraping Estático 19

Web Scraping dinámico 19

Web Scraping de una sola página o Crawling Vertical 19

Web Scraping de varias páginas o Crawling Horizontal y Vertical 19

Web Scraping con llamadas a API 20

3.3.1.2. Web Scraping en las Agencias Cartográficas 20

Desarrollo de la metodología de extracción de datos. 20

Consideraciones en Web Scraping 21

3.3.2. OSM: extracción de cambios respecto de esta fuente de datos 21

Ventajas e inconvenientes de OSM 23

Inconvenientes 23

Ventajas 23

Descarga de datos OSM 24

Consultas con lenguaje QL.....25
Consultas XML.....30
Limitaciones en las descargas.....34
Tratamiento de datos descargados.....35
3.3.3. WFS: extracción de datos desde servicios de descarga.....42
Ejemplo buenas prácticas de un sistema de extracción de datos.....42

Índice de ilustraciones

Ilustración 3- 1: Intersección de perímetros con un margen de tolerancia.....6
Ilustración 3- 2: Traslación de los dos candidatos a un origen común.....7
Ilustración 3- 3: Fuentes con diferente nivel de fragmentación.....8
Ilustración 3- 4: Ejemplo simplificado de conflación con discrepancias topológicas y geométricas.....9
Ilustración 3- 5: Ejemplo simplificado de conflación con semánticas.....10
Ilustración 3- 6: Comprobación de coincidencias semánticas.....11
Ilustración 3- 7: tramos del producto C con predominancia dentro del producto A.....12
Ilustración 3- 8: tramos cartografía que requieren de revisión visual.....12
Ilustración 3- 9: Elemento de OSM a procesar (izq) y elemento de OSM tras buffer y envelope (der).....13
Ilustración 3- 10: Elementos extraídos de IGR-RT en base al envelope (izq) y elementos de IGR-RT filtrados al buffer (der).....14
Ilustración 3- 11: Tramos de IGR-RT agrupados por id_vial, clase y tipo_vial (izq) y tramos IGR-RT tras disolución (der).....14
Ilustración 3- 12: Elemento OSM interpolado (izq) y relación entre interpolación y tramos disueltos de IGR-RT (der).....15
Ilustración 3- 13: Tramos de IGR-RT interpolados (izq) y relación entre interpolaciones y elemento OSM (der).....15
Ilustración 3- 14: Evaluación de resultados del proceso.....15
Ilustración 3- 15: Procesado de elementos con más de un tramo IGR-RT candidato.....16
Ilustración 3- 16: Generación de zonas de revisión.....17
Ilustración 3- 17: Arquitectura cliente –servidor.....18
Ilustración 3- 18: Ejemplo de Web para la extracción o Crawling vertical.....19
Ilustración 3- 19: Sintaxis básica de llamadas QL a Overpass.....25
Ilustración 3- 20: Rectángulo de consulta.....26
Ilustración 3- 21: Ejecución secuencial de sentencias QL.....27
Ilustración 3- 22: Unión de sentencias QL.....27
Ilustración 3- 23: Resta de sentencias QL (1).....28
Ilustración 3- 24: Resta de sentencias QL (2).....28
Ilustración 3- 25: Archivos SHP generados con datos descargados de OSM.....41
Ilustración 3-26: Esquema del procedimiento de actualización de edificios de BTN, desde la consulta WFS a la conflación.....43

Índice de tablas

Tabla 3- 1: Ejemplo de tablas a relacionar.....3
Tabla 3- 2: Ejemplo de decisión basada en la ratio de solape 5
Tabla 3- 3: Ejemplo de tabla de decisión con dos criterios.....7
Tabla 3- 4: Estadísticas de los resultados 17
Tabla 3- 5: Principales servidores públicos de Overpass 25

3.1. Integración basada en atributos alfanuméricos

En este apartado se van a analizar las características de la atribución alfanumérica o componente semántica (o temática) de la información geográfica, es decir, la naturaleza o propiedades del objeto geográfico que nos interesa modelar. La componente semántica va íntimamente ligada a la componente espacial, la cual le sirve de soporte y establece un marco geográfico materializado por la posición en un sistema de referencia.

Nos enfrentamos ahora a la tarea de incorporar a nuestro modelo de datos información semántica de una fuente de datos exógena, es decir, procedente de otro organismo u otro productor de información. Debido a la relación íntima de las componentes semántica y espacial, es habitual, en cualquier proceso de integración, tener que manejar y tratar ambas componentes para agregar la información a los objetos geográficos de nuestro modelo de datos. No obstante, las explicaciones que se exponen en este apartado se centran exclusivamente en la componente semántica, es decir, en la comparación de los atributos no geométricos de los conjuntos de datos que queremos relacionar.

Al integrar información semántica procedente de una fuente de datos exógena es importante definir una estrategia adecuada. La mecánica habitual para comparar dos conjuntos de datos es enlazarlos mediante uno o varios atributos que aseguren la relación inequívoca entre los elementos de ambos conjuntos. Si excluimos el análisis geométrico -que se explicará más adelante- el primer paso consiste en identificar aquellos atributos comunes o similares que permiten emparejar los elementos. Esta parte del proceso quizá es la más importante ya que suele exigir el estudio del modelo datos o especificaciones técnicas de la fuente exógena en cuestión. Una vez relacionados los conjuntos de datos, el siguiente paso es incorporar, con las manipulaciones necesarias, la información requerida al modelo de datos destino.

Una vez identificados los atributos que sirven de enlace, el siguiente paso es realizar las consultas de relación que une los elementos de los dos conjuntos de datos a través dichos atributos. Los principales factores a tener en cuenta para relacionar atributos temáticos son:

- **Tipo de dato** con el que están definidos los atributos. Se trata de la propiedad más importante porque determina qué información puede almacenar el atributo. Los atributos temáticos pueden ser, a grandes rasgos, de los siguientes tipos: **numéricos**: *enteros* (pueden ser positivos, negativos y el cero) y *reales*; a**lfanuméricos**: contienen letras, símbolos y cifras numéricas; **booleanos**: solo admiten los valores *verdadero* o *falso* y **memo**: información alfanumérica de longitud ilimitada.

  Vemos un ejemplo: el valor '*98234*' puede almacenarse en un atributo numérico y también en un atributo alfanumérico. A primera vista, el contenido de los dos atributos es el mismo, pero sus propiedades son bien diferentes. Dependiendo de la aplicación que se esté usando, este problema puede ser importante y requerir de una conversión previa de alguno de los dos atributos antes de poder compararlos.

- **Formato** del contenido de los atributos. Es la forma o modo de presentación de lo que está almacenado en los atributos. Vamos a destacar algunos casos habituales:

  - **Orden** del contenido: en el siguiente ejemplo si el *atributo1* contiene el valor '*Los Urquiza*' y el *atributo2* el valor '*Urquiza, Los*', será difícil compararlos usando un operador '=', siendo necesario realizar alguna operación de manipulación de los textos previa a la comparación.

- **Mayúsculas, minúsculas, tildes y otros caracteres especiales**: '*Picón de la Cigüeña*', '*Picon de la Ciguena', 'picón de la ciGüeÑa*' son valores de atributo muy similares, pero no iguales. Aunque cualquier manipulación puede alterar los resultados, es conveniente modificar los campos de texto que se quieren relacionar eliminando tildes, poniendo todos los caracteres en mayúsculas o minúsculas, o teniendo precaución con caracteres como '*Ñ*', '-', '.', etc.
- **Espacios en blanco**: es otro problema habitual por el que falla la coincidencia de atributos de texto. Generalmente por error, muchas veces las cadenas de texto tienen espacios en blanco delante o detrás, o más de un espacio entre palabras.
- **Abreviaturas**: hay que tener en cuenta el uso de convenciones ortográficas para acortar la escritura de ciertos términos o expresiones. Por ejemplo, las parejas de valores '*Ntra. Sra. de Sta. María*', '*Nuestra Señora de Santa María*' y '*Avenida de la Paz*', '*Avda. De la Paz*' son equivalentes pero diferentes. Puede resultar interesante disponer de una lista de las abreviaturas más comunes para facilitar las tareas de comparación.

Es importante destacar que, según la aplicación que se use, existen utilidades ya diseñadas que permiten hacer comparaciones entre cadenas de textos y obtener un índice o porcentaje de similitud, lo cual ayuda a solucionar todas las diferencias de formato mencionadas anteriormente. Estas utilidades están basadas en *algoritmos de coincidencia aproximada* (*fuzzy matching algorithm*) y su aplicación está en la comparación de tablas o fuentes dispares que no contienen identificadores únicos o claves primarias y externas apropiadas. Ejemplos de estas utilidades son el transformador '*FuzzyStringComparer*' del software *FME* o las funciones '*hamming_distance_levensheim*' y '*longest_common_soundex*' del software *QGIS*.

- **Tamaño** de los atributos. Aunque es un problema más ligado al almacenamiento del contenido de un atributo en otro que a la comparación de atributos, conviene tenerlo en cuenta. Por ejemplo, el valor de un atributo de texto definido como *longitud 255* no podremos volcarlo en otro de *longitud 50*, independientemente de la longitud real de los textos almacenados; o si la aplicación utilizada sí permite volcar el contenido, corremos el riesgo de truncar el texto original si es mayor de 50 caracteres. El mismo razonamiento se puede aplicar a los atributos numéricos, por ejemplo, el valor de un atributo *Double* no cabe en un atributo definido como *Integer*.
- **Combinaciones** de atributos. Para su explicación, veamos el siguiente ejemplo en el que *atributo1=Nombre*, *atributo2=Apellido* y *atributo3=Apellido, Nombre*. Para relacionarlos es necesario concatenar los dos primeros atributos con un orden específico e insertar los caracteres '*,* ' entre ellos, de tal manera que generemos un nuevo atributo combinando otros de origen. De la misma manera que hemos concatenado o sumado atributos, la solución al ejemplo anterior la podemos obtener dividiendo *atributo3* en la posición que ocupa el carácter ','. Es decir, a partir de uno o varios atributos creamos otro u otros atributos, en la forma que los necesitamos para realizar la comparación. A todas estas situaciones le son de aplicación las operaciones con cadenas de textos tales como concatenar, cortar, dividir, buscar, extraer, recortar, etc.

- **Multiplicidad.** Al comparar los elementos de dos fuentes de datos mediante un atributo común, no siempre hay una relación 1:1 entre los valores de dicho atributo. Vamos a verlo con el siguiente ejemplo:

Aeropuertos tabla1		
Nombre	**Terminal**	**Vuelos**
A	1XD3	500
A	J728	1000
A	B6T4	100
B	UR4T	85
C	12TY	200
D	H7K2	100
D	U723	260

Aeropuertos tabla2		
Tipo	**Uso**	**Nombre**
aeródromo	mixto	A
deportivo	civil	B
aeródromo	civil	C
base aérea	militar	D
aeródromo	civil	E

Tabla 3- 1: Ejemplo de tablas a relacionar

Si relacionamos las tablas por el atributo común *Nombre*, el aeropuerto '*A*' se relaciona con tres elementos de la *tabla 3-1*, el aeropuerto '*D*' con 2, los aeropuertos '*B*' y '*C*' con 1, y el aeropuerto '*E*' con ninguno.

- **Identificadores exógenos.** La función principal de un identificador es distinguir de manera inequívoca al elemento que representa. Es evidente que cualquier tarea de comparación resulta mucho más sencilla si disponemos de algún atributo común que tenga esta característica. Es frecuente que los organismos oficiales generen identificadores de los objetos u elementos sobre los que ejercen su competencia. Si estos elementos tienen cabida en nuestro catálogo de objetos geográficos, es absolutamente recomendable incluir estos identificadores exógenos en nuestro modelo de datos.

  Por ejemplo, los valores '*Castillo de Villanueva del Cañedo*' y '*Castillo de Villanueva del Cañedo o Castillo del Buen Amor*' representan los nombres de un mismo castillo en España, pero son diferentes ya que el segundo contiene dos nombres distintos del mismo elemento. La relación entre ambos valores será mucho más sólida a través del identificador que le haya asociado el titular. En este caso sería el '*RI-51-0000850-00000*' de *Bien de Interés* Cultural proporcionado por el *Ministerio de Cultura de España*.

  Los identificadores exógenes, además de facilitar el control de las variaciones o actualizaciones de nuestros elementos, también ayudan a la detección de omisiones (el identificador no se encuentra en nuestros datos) y comisiones (el identificador no se encuentra en la fuente exógena).

En resumen, descartando el análisis geométrico, la mejor opción para comparar dos conjuntos de datos es el uso de identificadores únicos ya que eliminan ambigüedades. Si no se dispone de ellos, debemos recurrir a otros atributos y prestar atención a la posible falta de concordancia, siendo habituales los conflictos por diferencias en el tipo de dato, el formato, etc. Las utilidades de comparación de cadenas de textos basadas en *algoritmos de coincidencia aproximada* (*fuzzy matching algorithm*) suponen una importante ayuda cuando no disponemos de identificadores únicos.

3.2. Integración por criterios espaciales

En el mundo de la producción cartográfica es muy habitual encontrar productos cartográficos relativos al mismo ámbito geográfico y con temáticas comunes, confeccionados por diferentes proveedores. Cada una de las cartografías está producida con unos objetivos que pueden tener similitudes, pero frecuentemente presentan discrepancias, generalmente vinculadas a los procesos de captura de los datos empleados, las escalas, los procesos de edición y formación cartográfica, los requisitos de usuario o el hecho de no emplear los mismos estándares.

El proceso de integración de datos espaciales para un proyecto cartográfico se enfrenta a problemas ocasionados por la falta de consistencia entre la información procedente de las distintas fuentes que se combinan. Las tareas de conflación tratan de resolver esos problemas para producir un conjunto de datos unificado que incorpore la información más actualizada y de mayor calidad que pueda extraerse de cada una de las fuentes.

Como fase preparatoria al proceso, conviene analizar con detenimiento los conjuntos de datos que van a intervenir en la conflación con el objeto de homogenizar sus sistemas de referencia, interpretar correctamente los modelos de datos e identificar nexos de unión, y detectar cualquier otro aspecto que facilite la ejecución del proceso.

Este apartado va a estar dedicado a la identificación, en distintas fuentes, de las entidades que son la representación del mismo objeto del mundo real, en base a criterios geométricos: primero, para el caso de la conflación entre geometrías de tipo superficial y después para las de tipo lineal que será además ilustrado con dos casos de uso.

3.2.1. Integración de datos superficiales

3.2.1.1. Selección de candidatos

En general, los algoritmos de identificación deben realizar primero una selección de candidatos mediante el análisis de la proximidad espacial entre los miembros de los dos conjuntos de datos que participan en la conflación. Según cuál sea el grado general de semejanza entre estos dos conjuntos de datos, el técnico debe elegir el criterio de proximidad que resulte más apropiado. Un criterio muy frecuente es seleccionar todos los candidatos del conjunto de datos B que se solapan con cada entidad del conjunto A. En otros escenarios más desfavorables, el criterio debe ampliarse para incluir también los candidatos de B que no solapan con el elemento de A pero que se encuentran a menos de una cierta distancia.

Otras ocasiones, por el contrario, son propicias para realizar un análisis espacial más liviano, que no tenga en cuenta la geometría completa del polígono, sino únicamente las coordenadas de un centroide representativo. En este supuesto, cada entidad de A se relacionaría normalmente con un solo candidato de B, aquel en cuyo interior se encontrase el centroide de A.

3.2.1.2. Medida de la similitud

A veces, la selección de candidatos es el principio y el fin del proceso de identificación. Pensemos, por ejemplo, en dos mapas con las divisiones administrativas de un mismo territorio que únicamente presentan tenues variaciones en la geometría de las fronteras de unas regiones con otras. El proceso de selección de candidatos (por el método del centroide) habría proporcionado una relación 1:1 entre los polígonos de uno y otro mapa, que puede ser aceptada sin más consideraciones.

En cambio, en un caso de conflación más general sí que será necesario evaluar cuantitativamente el grado de similitud que presenta la geometría de cada par de candidatos y, en función de esta medida, decidir si ambos representan (o no) el mismo objeto del mundo real. La ventaja de tratar con geometrías de tipo superficial es que admiten la medida de su superficie de solape, que es un criterio muy revelador de la similitud entre dos polígonos. No existe un concepto equivalente que sea aplicable a otros tipos de conflación, como la de puntos con puntos o líneas con líneas.

Siendo A un elemento del primer conjunto de datos y siendo B un elemento del segundo conjunto entre los que existe un cierto solape A∩B, se pueden plantear diversas ratios que relacionan la superficie de solape con la superficie de cada elemento individual o con la de su polígono de unión. Quizás esta última sea la más relevante de todas. Denominando D al valor de similitud obtenido y S a la superficie, resulta:

$$D(A,B) = \frac{S(A \cap B)}{S(A \cup B)}$$

Esta expresión da como resultado un valor comprendido en el intervalo [0,1]. Cuanto más alto sea el valor, más similitud se supone que existe entre los polígonos A y B. La citada expresión, sola o combinada con otras ratios de superficie como $\frac{S(A \cap B)}{S(A)}$ o $\frac{S(A \cap B)}{S(B)}$ u otros indicadores, permite definir una tabla de decisión tan compleja como sea necesario. Por ejemplo, en un caso sencillo podrían establecerse los siguientes criterios:

CRITERIO	**DECISIÓN**
D(A,B) ≥ 0.8	Se asume la correspondencia entre A y B
0.4 ≤ D(A,B) < 0.8	Se toman en consideración otros criterios adicionales
D(A,B) < 0.4	Se descarta la correspondencia entre A y B

Tabla 3- 2: Ejemplo de decisión basada en la ratio de solape

Los valores introducidos en la tabla son arbitrarios. No hay regla alguna que marque el valor de cada umbral; sólo es posible determinar cuáles son los intervalos idóneos tras un estudio atento de los datos.

El criterio basado en la medida de la superficie de solape tiene una lectura más clara en los extremos superior e inferior de la escala que en los valores medios. Si la ratio es igual a 1, la identificación puede darse por segura, ya que la equivalencia del área de intersección y de unión implica la coincidencia de forma entre los dos elementos. Pero en la gama media de valores pueden aparecer parejas de candidatos que un ojo humano clasifica claramente como homólogos junto a otras parejas que descarta, también claramente. Y ambas parejas con una ratio de solape muy parecida.

Por todo ello, es conveniente complementar las medidas de superficie con otras que también tengan en cuenta la forma de los polígonos, por lo menos en aquellos casos en los que la ratio de solape no resulte concluyente. Teniendo en cuenta que la forma de un polígono viene dada por su contorno y que éste está constituido por una polilínea sin más particularidad que la coincidencia del primer vértice con el último, debe admitirse que las mismas técnicas de conflación mencionadas en el apartado dedicado a la integración de datos lineales serán también válidas para evaluar la coincidencia de forma entre geometrías de tipo área.

En particular, puede ser interesante calcular algún cociente entre perímetros para utilizarlo como una medida de la similitud de la forma, por extensión del método basado en el área de solape. Siguiendo esta analogía, llamemos F_A a la frontera del polígono A y F_B a la frontera del polígono B. La longitud de la intersección entre ambas L ($F_A \cap F_B$) es la cantidad que debería tomarse como numerador en las nuevas expresiones, haciendo las veces del término S (A ∩ B) definido anteriormente.

El problema es que la intersección de dos polilíneas, salvo que sean total o parcialmente coincidentes, es un conjunto discreto de puntos que no tiene medida de longitud. Para evitar este fenómeno de discontinuidad de la intersección lineal, es preferible pasar por alto los desajustes entre las fronteras de A y B cuando estos no superen unas pequeñas cantidades. Tal efecto puede lograrse mediante la constitución de un búfer en torno a la frontera de A con una anchura dada por un cierto parámetro de tolerancia. De esta manera, la intersección entre la frontera de B y el interior del búfer se puede asimilar a la intersección con la propia frontera de A.

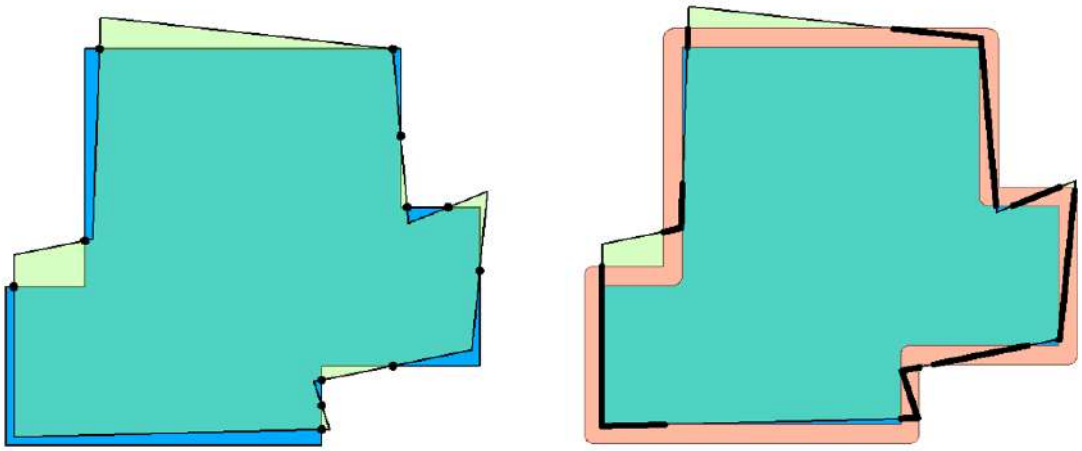


Ilustración 3- 1: Intersección de perímetros con un margen de tolerancia

Cuanto mayor sea la longitud de la intersección así obtenida, mayor será también el grado de similitud entre las figuras. Para darle un carácter adimensional a la expresión y para que su valor quede comprendido entre 0 y 1, la longitud obtenida se divide entre el perímetro total, ya sea el del polígono B o el de la unión de A y B, según se considere más conveniente. La fórmula final presenta un paralelismo evidente con la ratio de solape:

$$F(A,B) = \frac{L(F_A \cap F_B)}{L(F_{A \cup B})}$$

Combinando ambos criterios, la tabla de decisión podría quedar configurada del siguiente modo:

		CRITERIO DE PERÍMETRO		
		F(A,B) ≥ 0.9	0.7 ≤ F(A,B) < 0.9	F(A,B) < 0.7
CRITERIO DE SUPERFICIE	D(A,B) ≥ 0.8	A=B	A=B	A≠B
	0.4 ≤ D(A,B) < 0.8	A=B	A≠B	A≠B
	D(A,B) < 0.4	A≠B	A≠B	A≠B

Tabla 3- 3: Ejemplo de tabla de decisión con dos criterios

3.2.1.3. Operaciones para la mejora de la identificación

En algunos trabajos de conflación puede resultar aconsejable realizar un proceso de simplificación de vértices como paso previo a la evaluación de las medidas de similitud. Con esta operación se consigue dar énfasis a los rasgos esenciales que definen la forma de las figuras, además de reducir el volumen de cómputo necesario.

Otra transformación que puede resultar útil en algunas situaciones consiste en referir las coordenadas de todos los vértices de cada polígono a su centroide o, dicho de otra forma, aplicar a cada polígono una traslación que lleve su centroide al origen de coordenadas. Naturalmente, esta traslación sólo puede realizarse después de haber completado la selección de candidatos, pues en caso contrario, todas las entidades existentes solaparían unas con otras.

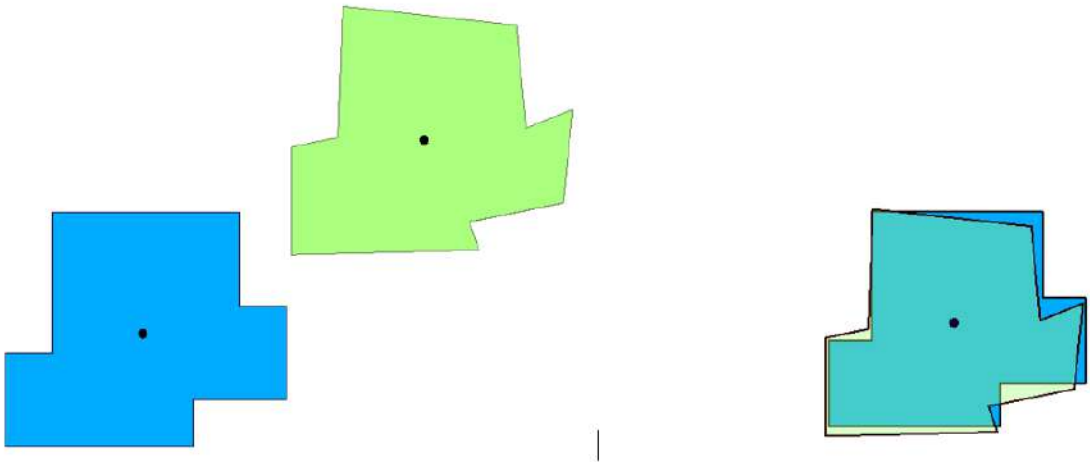


Ilustración 3- 2: Traslación de los dos candidatos a un origen común

Para comprender la utilidad que tiene referir las coordenadas al centroide, pensemos que las medidas de similitud presentadas (tanto la que se basa en la intersección de áreas como en la de perímetros) dan resultado nulo si las dos entidades analizadas se encuentran separadas entre sí. Esta propiedad convierte a tales medidas en indicadores poco adecuados para resolver ciertos casos reales, como el representado en la parte izquierda de la ilustración anterior. Son casos en los que las diferencias entre entidades candidatas radican más en su posición que en su forma.

Referir las coordenadas de los polígonos al centroide es una manera simplificada de abordar este problema, pero puede dar resultados más que aceptables en muchos casos, sin requerir apenas cálculo. Una solución más elaborada podría ir enfocada a la identificación de vértices homólogos dentro de las figuras.

Una vez compensado el desplazamiento relativo de los polígonos, las medidas de similitud ya son aplicables como en un caso normal. En general, en un mismo trabajo de conflación coexisten parejas de candidatos cuya identificación se ve favorecida por el cambio del origen de coordenadas, junto a otras a las que no las afecta o les causa perjuicio. Puede ser conveniente realizar un análisis previo de cada pareja basándose en algún criterio espacial para decidir en qué casos conviene aplicar la traslación o, directamente, pueden obtenerse por

separado los indicadores de similitud con y sin traslación, y escoger el resultado que sea más favorable en cada caso.

Otro importante factor que perturba el proceso de identificación se produce cuando las entidades presentan un nivel de fragmentación diferente en cada fuente de datos. Este problema es análogo al que se produce en la conflación lineal cuando las dos redes no tienen la misma división en tramos. En un caso y otro (superficial y lineal), la consecuencia es la misma: el grado de similitud entre entidades individuales es débil, aunque se aprecie un buen ajuste conjunto.

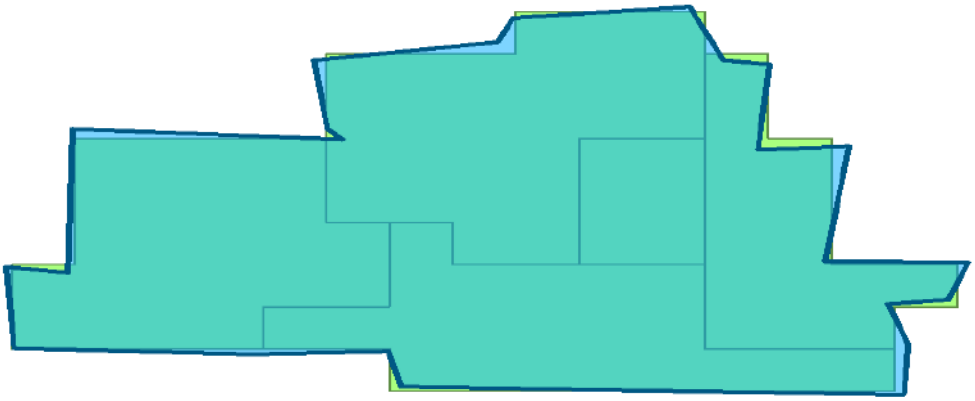


Ilustración 3- 3: Fuentes con diferente nivel de fragmentación

Este problema puede atenuarse si previamente se ha realizado la unión geométrica de todas las entidades del conjunto de datos B vinculadas a cada entidad de A (o viceversa). Cuando son contiguas, esa unión equivale a disolverlas en un único polígono; cuando no lo son, se tratan como una colección de geometrías. En este último caso, las medidas de la intersección de área y perímetro se obtienen como sumatorio de las entidades individuales.

Si las medidas de similitud calculadas a partir de la geometría conjunta no superan un valor predeterminado, la solución colectiva queda descartada, se deshace el agrupamiento y se repite el cálculo para cada una de las entidades de forma independiente.

3.2.1.4. Combinación de entidades

Una vez obtenido el conjunto de correspondencias entre entidades, la información de las mismas se combina siguiendo reglas predefinidas. Esto afecta tanto a la geometría como a los atributos. Si la relación es 1:1, se puede dar preferencia absoluta a los atributos de una de las dos entidades o realizar una mezcla entre ambas para generar el conjunto de valores atribuido a la entidad combinada. El criterio que se adopte en cada caso puede estar en relación con el grado de similitud observado entre las dos entidades. Si la correspondencia entre entidades es de tipo 1:N o M:N, será necesario establecer fórmulas de combinación que dependerán de la naturaleza de cada atributo.

También existen formas diferentes de combinar las geometrías. Puede ocurrir que una de las fuentes de datos tenga preferencia por poseer una mejor precisión espacial. Si ambas fuentes tienen precisión similar, se puede calcular una geometría promedio entre las dos.

3.2.2. Integración de datos lineales

Al igual que ocurre con los datos de carácter superficial, la integración de objetos lineales por procesos de conflación es una práctica muy frecuente en el ámbito de la integración de datos

existentes, con especial aplicación en los conjuntos de datos de redes como son, por ejemplo, las redes de transporte, las hidrográficas o las energéticas (eléctricas, gasoductos, etc.).

El proceso de conflación permite la detección y subsanación de discrepancias semánticas, topológicas y geométricas del conjunto de datos a actualizar respecto del considerado como referencia.

3.2.2.1. Metodología a aplicar en la conflación de elementos lineales

Una vez identificados y analizados los dos conjuntos de datos que van a participar en el proceso de conflación, el siguiente paso es la identificación de los objetos homólogos. Se trata de identificar las discrepancias que existen entre el conjunto de datos a mejorar y el considerado de referencia, las cuales pueden ser de naturaleza: topológica, geométrica o semántica (alfanumérica).

En la siguiente ilustración se muestra un ejemplo gráfico simplificado de conflación de elementos lineales donde se presentan discrepancias geométricas y topológicas entre los objetos geográficos homólogos de los dos conjuntos de datos, correspondientes con entidades de carreteras en la realidad, y su posible solución tras aplicar los procesos que los modifican según los datos de referencia.

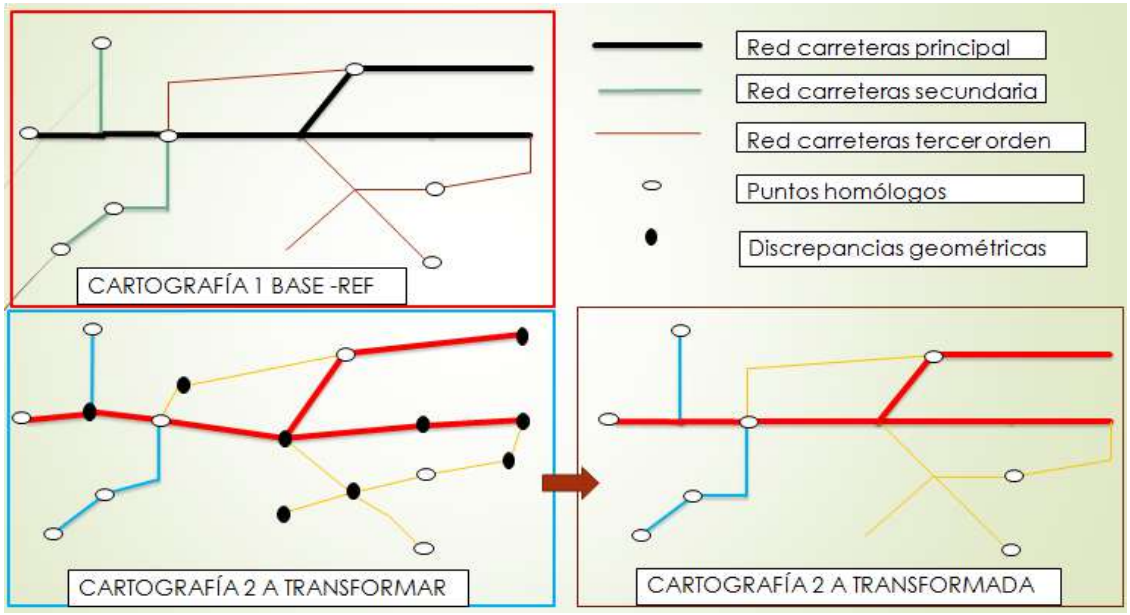


Ilustración 3- 4: Ejemplo simplificado de conflación con discrepancias topológicas y geométricas

Como ya se ha indicado anteriormente, el proceso de conflación también puede ser aplicado para la depuración y actualización de la información semántica que esté asociada a los objetos geográficos homólogos de los dos conjuntos de datos. Igualmente, será necesario establecer los criterios que permiten identificar el grado de similitud de los objetos homólogos y, cuándo éste sea aceptable, proceder al análisis de las discrepancias semánticas.

En la siguiente ilustración se muestra un ejemplo simplificado donde se presentan las discrepancias entre dos cartografías con respecto al atributo *nombre* aplicado al objeto geográfico caminos:

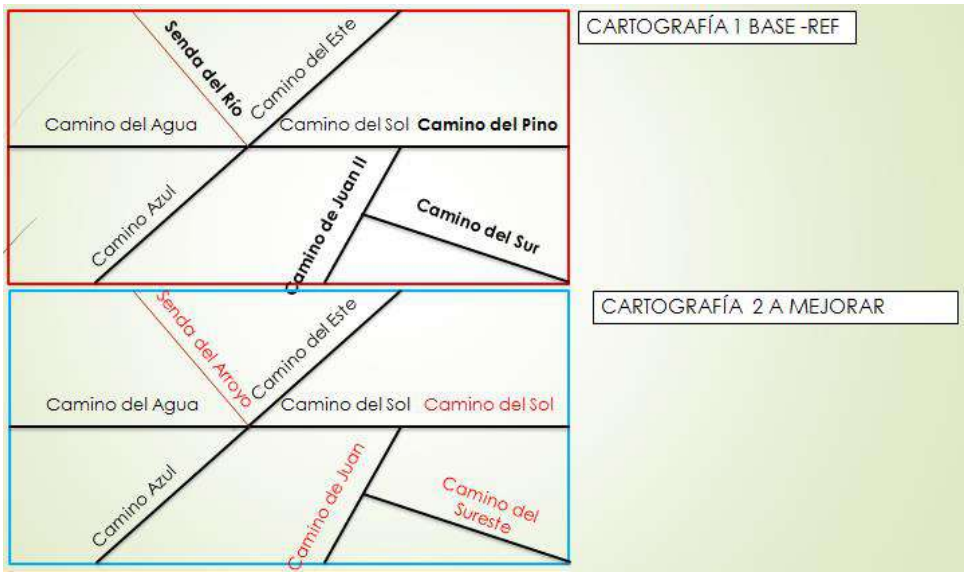


Ilustración 3- 5: Ejemplo simplificado de conflación con semánticas

La fase final consiste en la aplicación del proceso en el que se implementan el conjunto de reglas, algoritmos y transformaciones a aplicar al producto cartográfico susceptible de ser corregido respecto al producto de referencia para reducir así las discrepancias observadas.

3.2.2.2. Caso de uso: conflación en una red de caminos

Para ilustrar el contenido del apartado anterior, a continuación, se muestra la metodología empleada de un caso real de conflación en la red de caminos en el ámbito de la región de La Rioja (España) ejecutado en el seno del IGN de España:

Selección de la información de referencia

Se dispone de 3 conjuntos de datos de información geográfica de caminos para el mismo ámbito, que van a intervenir en la conflación:

- Conjunto de datos A de red de Caminos de Agencia Cartográfica Regional La Rioja (posible producto a ser la referencia semántica).
- Conjunto de datos B de Base Topográfica Nacional: red clásica de Caminos (producto auxiliar).
- Conjunto de datos C de la IGR –RT (Redes de Transporte): red de Caminos (producto a mejorar).

Este proceso de conflación se ejecuta sobre el objeto geográfico **Camino** incluido en la temática de Redes de Transporte. Este objeto tiene, entre otros, el atributo alfanumérico **'Nombre',** para registrar la denominación de los mismos. La problemática se inicia cuando se detecta que en los datos del producto C (IGR-RT) el nombre de los caminos es muy discrepante respecto al producto A (datos de La Rioja), donde solo hay una coincidencia del 35 - 40 % de los registros.

Para asegurar la calidad semántica de los datos A, se ejecuta un análisis previo entre **el producto A** y un **segundo conjunto de datos auxiliar B** que representa la antigua red de caminos de la Base Topográfica Nacional BTN, donde la coincidencia para los nombres de dicho objeto geográfico es del 80 – 85%. Por lo tanto, se concluye que el **producto A** se puede tomar

como referencia ya que es un producto más actualizado y es el de mayor escala de los 3 empleados en el análisis.

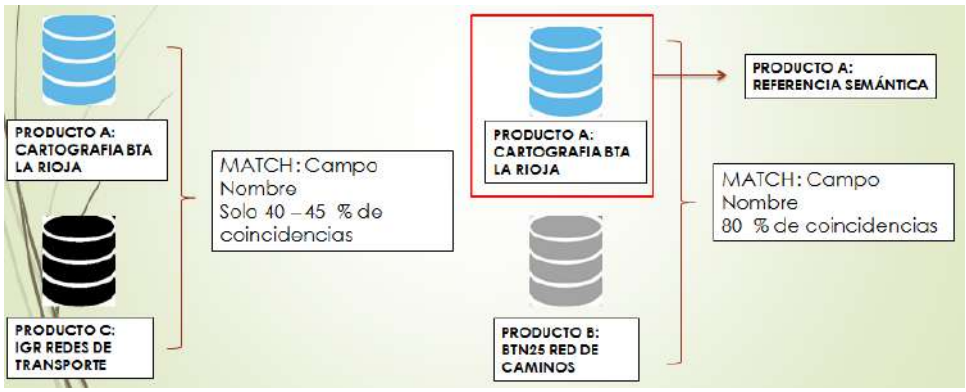


Ilustración 3- 6: Comprobación de coincidencias semánticas

Diseño de proceso de transformación para la integración de la información semántica

A continuación, se muestra cada una de las fases adoptadas para el caso real de la red de caminos de la Rioja.

FASE 1. Reproyección del producto A

En esta fase previa se reproyecta el producto A en el mismo Sistema de Referencia destino que el producto C para poder ejecutar la conflación con éxito.

FASE 2. Obtención de las áreas de influencia del producto A

En esta fase se implementa un *buffer* con un radio de influencia de 10 m respecto a la red de caminos del producto A, para poder estudiar la posterior inclusión de las geometrías de la cartografía C y, por tanto, poder proceder a la subsanación de los nombres de los caminos.

FASE 3. Análisis de relaciones espaciales entre las áreas de influencia y la red de caminos del producto C. Reglas de asignación semántica. Se analizan dos tipos de relaciones espaciales:

- Relación espacial en la que el producto C está incluido completamente dentro de las áreas de influencia: aquí se produce la asignación directa del nombre de los caminos del producto A sobre el producto C.
- Relación espacial en la que el producto C interseca con las áreas de influencia del producto A, donde hay 4 casuísticas:
    - Tramos cuyo porcentaje de inclusión dentro del *buffer* > 65 % y su longitud es mayor a 100 m: asignación directa.
    - Tramos cuyo porcentaje de inclusión dentro del *buffer* > 90 % y su longitud es menor a 100 m: asignación directa.

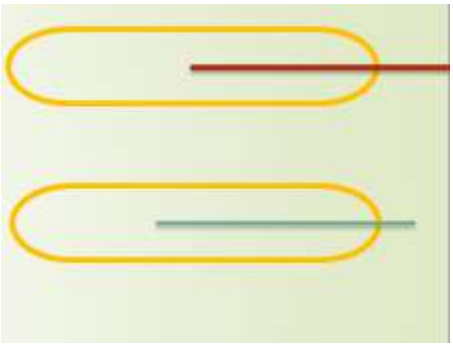


Ilustración 3- 7: tramos del producto C con predominancia dentro del producto A

- Tramos cuyo porcentaje de inclusión dentro del *buffer* < 65 % y su longitud es mayor a 100 m: supone una revisión visual
- Tramos cuyo porcentaje de inclusión dentro del *buffer* < 90 % y su longitud es menor a 100 m: supone una revisión visual

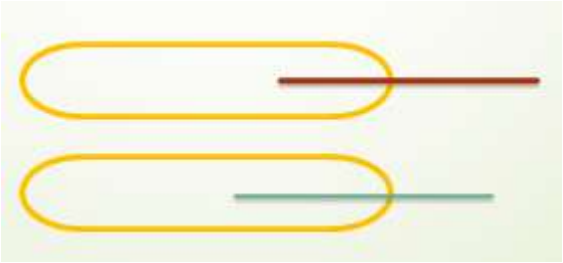


Ilustración 3- 8: tramos cartografía que requieren de revisión visual

FASE 4. Conclusiones

La capa del producto C susceptible de mejora dispone de 36.451 entidades geométricas de las que, una vez ejecutada la conflación semántica, se consiguen los siguientes resultados:

- 35.469 entidades cambian el nombre de los caminos de manera automática y por asignación directa usando como referencia el producto A. El 97.3 % de los datos.
- 982 entidades requieren de una revisión visual para asignar de forma tradicional el nombre de los caminos usando como referencia el producto A. El 2.7 % de los datos.

3.2.2.3. Caso de uso: detección de cambios en redes de transporte

En el siguiente ejemplo se muestra la metodología de conflación para detectar cambios geométricos entre bases de datos. Se trata de acotar las zonas de actualización en los productos cartográficos a partir de archivos que contienen cambios en la realidad que provienen de terceros.

Datos de partida

En este caso de uso se pretende desarrollar un planeamiento que facilite la actualización continua de la red viaria contenida en el producto IGR de Redes de Transporte (IGR -RT) del IGN de España. Como fuente continua de cambios de utilizarán archivos de la base de datos del proyecto *Open Street Maps,* descargados periódicamente.

En el modelo IGR-RT, los datos de viales se encuentran fundamentalmente en dos tablas, la tabla *rt_tramo_l* que se enfoca en la parte física de la red (contiene la geometría lineal del

trazado y otros atributos) y que por su estructura de red planar genera tramos entre todos los nodos de la red y la tabla *rt_vial_a* que recoge la información semántica de los viales. Ambas tablas se unen por una tabla relación *rrt_tramo_vial* que materializa una relación M:N entre ambas tablas. Debido a la estructura del producto y sus estrictos controles de calidad, todos los elementos de ambas tablas tienen una serie de atributos que han de ser rellenados siguiendo unos parámetros dados.

Por otro lado, en el modelo OSM, no existe una estructura tan rígida para los datos, ni unos controles de calidad tan exhaustivos, de modo que la cantidad y calidad de los atributos no es utilizable para comparaciones en la gran mayoría de los casos. Además, el modelo de almacenamiento utilizado genera viales completos, teniendo una sola geometría por vial desde su inicio hasta su final, al contrario que IGR-RT que los divide en cada intersección de tramos, cuando cambia el valor de algunos de los atributos y cuando interseca con una división territorial de carácter administrativa (ej. línea límite de municipio).

Cabe destacar que los criterios de digitalización en IGR-RT están fuertemente reglados y no generan grandes discrepancias entre zonas digitalizadas por distintos operadores, mientras que los criterios de digitalización en OSM son muy laxos, lo que hace que cada operador pueda digitalizar basándose en su propio criterio. Esto puede provocar incoherencias entre los datos derivadas de estas diferencias, por ejemplo, carreteras digitalizadas con más de un eje en el caso de carreteras multicarril, aceras digitalizadas como ejes, rotondas con múltiples criterios de selección de eje, etc.

Proceso de conflación

El proceso se ha realizado íntegramente mediante scripts de Python que procesan cada elemento de OSM de manera individual, de modo que un elemento no se ve afectado por los resultados de elementos contiguos, lo que evita errores de descarte en zonas con conflictos de digitalización.

En primer lugar, se lee la línea a analizar, se realiza un *buffer* a una distancia de 10 metros sobre ella y un *envelope* sobre el *buffer* previo.

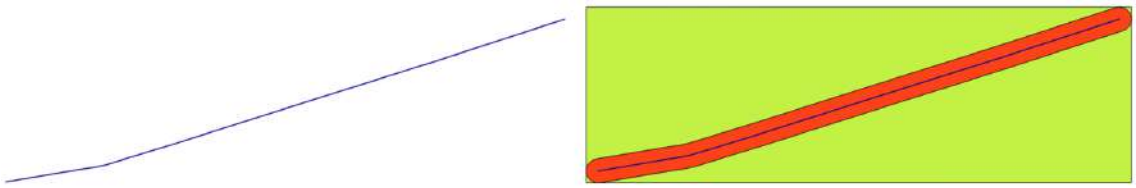


Ilustración 3- 9: Elemento de OSM a procesar (izq) y elemento de OSM tras buffer y envelope (der)

El motivo para realizar el *envelope* es para facilitar la consulta a base de datos de IGR-RT. La base de datos contiene millones de registros y los elementos que se descargan de OSM pueden tener longitudes de varios kilómetros con geometrías altamente complejas, por lo que la consulta sería sustancialmente más lenta. Es recomendable obtener más elementos en menos tiempo que menos elementos en mucho más tiempo.

Con la geometría del *envelope* se hará una consulta SQL a la base de datos IGR-RT para obtener todos los elementos que intersecten con ese cuadro, de los cuales se pedirán los siguientes atributos:

- Geometría.
- Identificadores (Id_tramo, del tramo; Id_vial, del vial)

- Tipo_vial (carretera, urbano, etc.)
- Clase (autopista, carretera convencional, etc.)
- Tipo_tramo (troncal, enlace, rotonda, etc.)

Posteriormente, de vuelta en Python se filtrarán los elementos que intersecten con el *buffer* original.

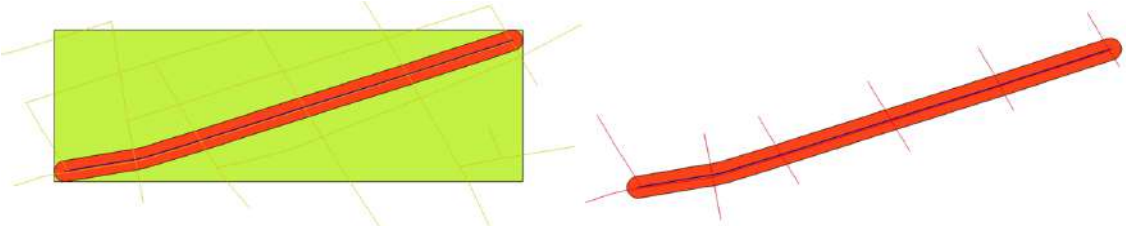


Ilustración 3- 10: Elementos extraídos de IGR-RT en base al envelope (izq) y elementos de IGR-RT filtrados al buffer (der)

Si en este punto no se hubiese obtenido ningún elemento procedente de IGR-RT que intersecase con el *buffer*, se marcaría el objeto de OSM como posible ALTA a integrar en la base de datos de IGR-RT. Realmente no se integra la geometría de OSM en IGR-RT si no que este ALTA tiene el significado de alta en la detección de objetos que no existen en IGR-RT y que por tanto habrá de capturar por la metodología habitual con la que se genera este producto (digitalización sobre orto, etc.).

A continuación, se toman los elementos de IGR-RT que han llegado a este punto y se agrupan según los campos id_vial, tipo y clase y se disuelven en base a ellos. Es importante tener en cuenta al disolver que por la naturaleza de algunos viales puede haber más de un resultado en la disolución, por ejemplo, en viales urbanos que en ocasiones tienen pequeñas ramificaciones de acceso a edificios o a zonas peatonales, por lo que se diseñó un algoritmo de disolución que fuese capaz de devolver todas las opciones de disolución posibles para un grupo sin generar geometrías múltiples.

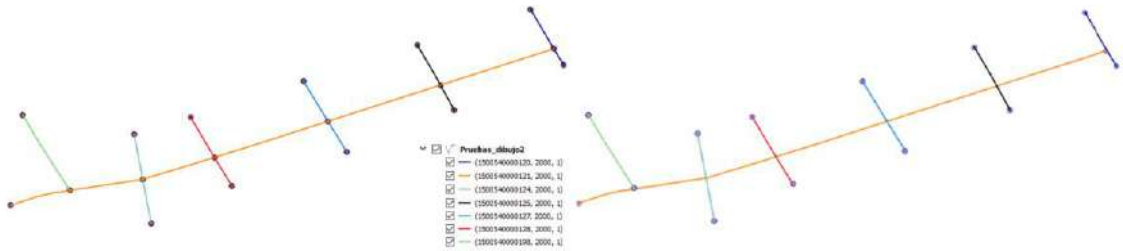


Ilustración 3- 11: Tramos de IGR-RT agrupados por id_vial, clase y tipo_vial (izq) y tramos IGR-RT tras disolución (der)

En el siguiente paso se vuelve a tomar la línea descargada de OSM y se realiza una interpolación que genera puntos sobre la línea espaciados un metro de distancia. Tomando estos puntos y las entidades disueltas, para cada una de las entidades se calcula qué porcentaje de puntos de la entidad de OSM está situado a menos de cinco metros.

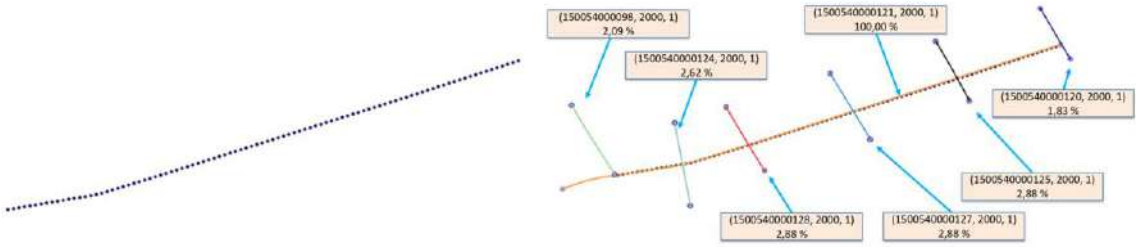


Ilustración 3- 12: Elemento OSM interpolado (izq.) y relación entre interpolación y tramos disueltos de IGR-RT (der.)

Con los porcentajes calculados, se toma el de mayor porcentaje y se descarta el resto. Hecho esto, se retoman las entidades de IGR-RT que forman la entidad disuelta y al igual que previamente se hizo con la entidad de OSM, se interpolan generando puntos cada metro. Para cada tramo se calculará el porcentaje de vértices que están a menos de 4 metros de distancia de la entidad de OSM y se eliminarán todos aquellos que tengan un porcentaje inferior al 50%.

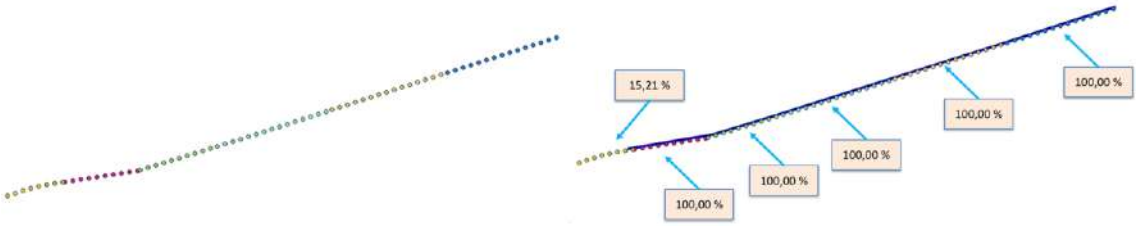


Ilustración 3- 13: Tramos de IGR-RT interpolados (izq.) y relación entre interpolaciones y elemento OSM (der.)

Tras estas comparaciones se pueden presentar tres alternativas. La primera de ellas es que no haya ningún elemento de IGR-RT candidato a ser semejante al elemento de OSM, por lo que el elemento OSM se marcaría como posible elemento a incorporar a IGR-RT. La segunda alternativa es que haya un solo elemento de IGR-RT candidato a ser semejante al elemento de OSM, en este caso se compararán las longitudes de ambos elementos aplicando una serie de tolerancias que disminuyen a medida que aumenta la longitud de los tramos. En caso de que esté dentro de las tolerancias, se considerará que los elementos son iguales y por lo tanto no ha habido cambios, en caso de que no entren en tolerancia, se marcará el elemento como elemento a revisar para una posible edición en IGR-RT. La tercera alternativa es que para un elemento de OSM haya más de un elemento IGR-RT candidato a ser semejante, en este caso se sigue un nuevo proceso de revisión exhaustiva.

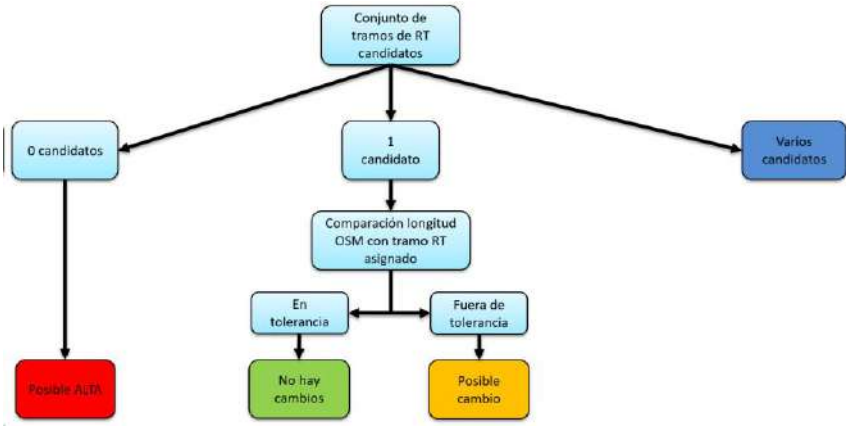


Ilustración 3- 14: Evaluación de resultados del proceso

Lo primero que se comprueba es si el punto inicial y el punto final del tramo de OSM son iguales, en este caso es altamente probable que se trate de una rotonda; si de los elementos obtenidos de IGR-RT alguno tiene su atributo tipo_tramo con el valor rotonda, se considera que no hay edición.

En caso de que no se considere rotonda, se compara la longitud del elemento OSM con la suma de longitudes de IGR-RT, de modo que si está dentro de las tolerancias se determina que no hay cambios y si está fuera de tolerancia sigue procesándose. En el siguiente paso se comprueba qué elemento tiene una longitud mayor, en caso de que sea el elemento de OSM, se considera que no hay elementos suficientes de IGR-RT para cubrir la longitud del elemento de OSM por lo que es probable que falte algún tramo y por tanto se marca para revisión. En caso de que la suma de longitudes de IGR-RT sea mayor, se recupera el resultado de la interpolación de los tramos de IGR-RT respecto al elemento OSM, se resta el tramo que tenga menor relación y se vuelven a comparar las longitudes. Si siguen sin entrar en tolerancia, se avanza en el procesado del elemento.

En el siguiente paso, se toma de nuevo la interpolación del elemento OSM y se ejecuta un algoritmo de paralelismo, el cual calcula el porcentaje de puntos que están a menos de 4 metros de las entidades IGR-RT y la distancia media de los puntos, utilizando un cálculo combinado de ambos parámetros que establece que, a mayor distancia media de los puntos, mayor ha de ser el porcentaje de puntos que estén a menos de 4 metros, y se decide si se encuentran en tolerancia o no.

Los elementos que siguen sin entrar en las tolerancias pasan a un último algoritmo que utiliza "fuerza bruta de computación" para intentar resolver el problema. Aunque el algoritmo es prácticamente idéntico para elementos urbanos e interurbanos, se aplican pequeñas diferencias en la evaluación de resultados. Este algoritmo retoma todos los tramos de IGR-RT que pasaron el primer filtro (intersecar con el *buffer* a 10 metros) y los disuelve con un algoritmo especial que genera todas las combinaciones posibles que existen, pudiendo llegar a generar miles de caminos alternativos en tramos de varios kilómetros. Este algoritmo se considera un último recurso, ya que su ejecución para un solo tramo puede durar horas. Tras aplicar este algoritmo, todos aquellos elementos de OSM que no hayan pasado los filtros serán considerados elementos a revisar para posible edición en IGR-RT.

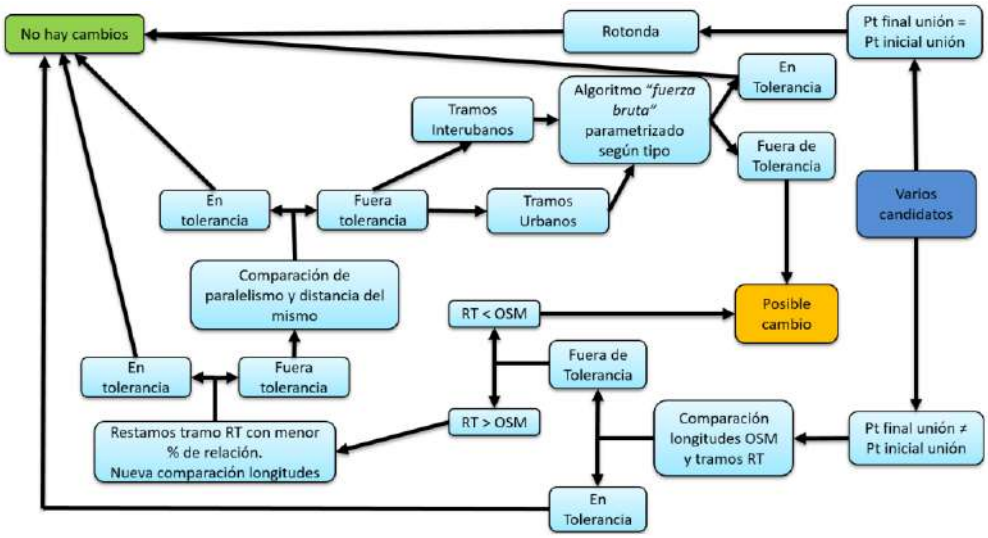


Ilustración 3- 15: Procesado de elementos con más de un tramo IGR-RT candidato

Una vez determinados los elementos individuales que han de ser revisados, se procede a agruparlos en zonas de revisión. Para ello se realiza un *buffer* a 100 metros a todos los elementos de OSM que se han marcado para revisión y posteriormente se disuelven para formar un archivo de zonas de revisión con el que el equipo de editores pueda trabajar.

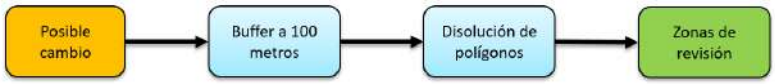


Ilustración 3- 16: Generación de zonas de revisión

Resultados obtenidos

Los archivos de cambios extraídos de OSM se obtienen a intervalos semanales y su contenido se divide en red principal y red urbana. En la red principal se producen una media de 5,193 cambios en OSM, de los cuales 678 son marcados para revisión en IGR-RT (13.06%) generando 291 zonas de revisión semanales. En la red urbana se producen una media de 6,960 cambios semanales en OSM, de los cuales 2,623 son marcados para revisión en IGR-RT (38.5%) generando 775 zonas de revisión.

Se evalúan los resultados de la conflación seleccionando aleatoriamente 500 elementos del archivo de una semana. A partir de esa selección se genera un archivo *shape* sin información de los resultados del proceso para revisión visual y evaluación de los elementos. Adicionalmente se generan dos archivos Excel, uno sin diagnóstico, para que el revisor apunte sus evaluaciones y otro con diagnóstico para posterior cruce de resultados. Se revisan los 500 elementos visualmente, comparando los datos de OSM con los de IGR-RT y apoyándose en ortofotos en los casos necesarios.

De los 393 elementos de OSM para los que el software consiguió establecer relación con elementos de IGR-RT marcándolos como elementos sin cambio, 386 eran correctos y 7 incorrectos. De los 107 elementos que el software marcó como elementos a revisar para posible edición, 102 eran cambios reales que en mayor o menor medida afectaban al producto IGR-RT y 5 no lo eran.

MATRIZ DE CONFUSIÓN		Predicho	
		No cambio	Cambio
Real	No cambio	386	5
	Cambio	7	102

Precisión	(VP+VN)/(VP+FP+FN+VN)	97,60%
Exactitud	VP/(VP+FP)	95,33 %
Sensibilidad	VP/(VP+FN)	93,58 %
Especifidad	VN/(VN+FP)	98,72 %

Tabla 3- 4: Estadísticas de los resultados

3.3. Fuentes de datos en la red utilizables en la actualización

En esta sección se explican tres tipos de fuentes de datos existentes en la red que son de especial relevancia para las agencias cartográficas a la hora de realizar los procesos de actualización de sus conjuntos de datos.

Así, se describe cómo la información de las webs que se consideren de referencia puede ser rastreada por bots diseñados en la búsqueda de palabras claves evocadoras de cambios en los objetos geográficos (ej. en el ámbito de las carreteras podrían ser: cesión, inauguración, trazado, etc.). También se dedica un amplio apartado a explicar las diversas formas de consulta y descarga que existen sobre el conjunto de cartografía colaborativa OSM. El último punto se centra en la utilidad de los servicios estándar de descarga WFS, ilustrándolo con un ejemplo de buenas prácticas sobre un sistema creado para la consulta y el consumo de los datos que se facilitan a través de este sistema.

3.3.1. Bots: extracción de datos por técnicas de Web Scraping

Una de las técnicas que permite facilitar la búsqueda de cambios en el territorio para actualizar la información geográfica en el seno de una agencia cartográfica son las metodologías de Web Scraping o extracción de datos de páginas web. Live Stats (web de estadísticas sobre el contenido que se está produciendo en la red en tiempo real) contabiliza la existencia de 1,718 millones de portales a fecha de octubre de 2019, de los cuales se calcula que sólo 200 millones están activos.

En efecto, el ecosistema de páginas web es amplio, complejo y diverso, alojando muchos millones de terabytes de información. Esta gran cantidad de datos es explotada hoy en día por organismos y agencias que, mediante la extracción regular, programada y siguiendo un conjunto de reglas, pueden obtener un subconjunto de datos aplicables en la consecución de los objetivos de dichas organizaciones.

3.3.1.1. Fundamento de Web Scraping. Tipologías

Estas técnicas se basan en la arquitectura cliente - servidor que es el soporte de la comunicación para la red de Internet, donde n clientes realizan peticiones o requerimientos a uno o varios servidores. Al hacer clic en la web, formalmente se hace una petición de información al servidor que devuelve literalmente el código HTML de la página web con toda la información alfanumérica susceptible de ser extraída.

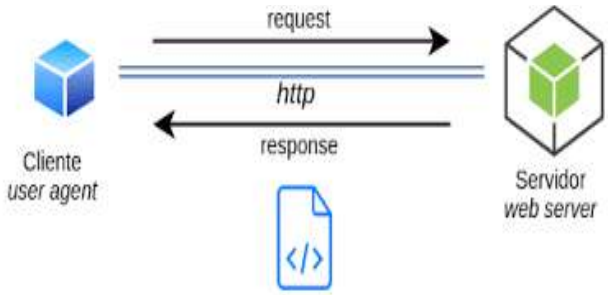


Ilustración 3- 17: Arquitectura cliente –servidor

A continuación, se muestra la clasificación de tipologías de Web Scraping más habituales:

Web Scraping Estático

Esta extracción se ejecuta sobre webs estáticas, que son aquellas en las que la totalidad de su contenido (todos sus ítems) se cargan de una sola vez después de recibir la respuesta por parte del servidor. Esta tipología es bastante sencilla de implementar mediante algoritmos de extracción ya que las webs estáticas tienen un funcionamiento sencillo.

Web Scraping dinámico

En esta segunda tipología se ejecuta sobre web dinámicas, que son aquellas en las que el servidor carga solo parcialmente parte del contenido de dicha web en la primera respuesta del mismo. Se conoce como carga dinámica, por tanto, dichas cargas posteriores se ejecutarán según acciones de interacción del usuario desde la máquina cliente con su navegador. Esta tipología es más compleja que la anterior ya que se debe conocer el funcionamiento del servidor.

Web Scraping de una sola página o Crawling Vertical

Aquí la extracción se ejecuta sobre webs en las que todos los posibles ítems de extracción se encuentran alojados en una página que está en el mismo dominio. En este caso el algoritmo de extracción emplea solo una regla vertical de búsqueda y extracción que es rápida y sencilla.

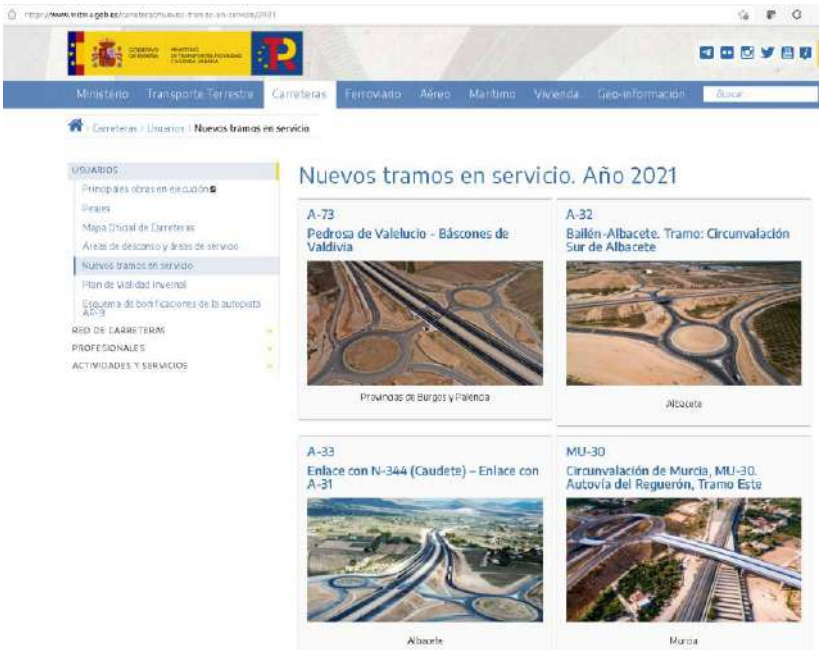


Ilustración 3- 18: Ejemplo de Web para la extracción o Crawling vertical

Web Scraping de varias páginas o Crawling Horizontal y Vertical

En este caso se ejecuta sobre webs en las que todos los posibles ítems de extracción se encuentran alojados en diferentes subpáginas ubicadas en el mismo dominio de la página principal. Esta tipología es también bastante sencilla ya que requiere:

- Una regla de navegación horizontal, para acceder a cada una de las subpáginas.
- Una regla de extracción vertical, para la extracción de los ítems de cada subpágina.

Web Scraping con llamadas a API

Existe un conjunto de páginas web que disponen de API públicas y documentadas, que permiten disponer de un mecanismo formal de extracción de datos alfanuméricos normalmente mediante formato JSON.

3.3.1.2. Web Scraping en las Agencias Cartográficas

La aplicación de estas técnicas en los procesos de actualización de la información geográfica puede resultar ser de gran utilidad. Son técnicas que se enfocan a realizar el seguimiento de aquellas webs que se consideran de mayor relevancia en la comunicación de variaciones surgidas en cada temática, bien por la trascendencia de la novedad, bien por la frecuencia de su publicación.

La búsqueda y extracción de información de las URLS seleccionadas permite a la organización diseñar y ejecutar disparadores que, de manera selectiva y automática, notifican los objetos geográficos que potencialmente deben ser actualizados. Siguiendo la terminología dentro de la disciplina de Web Scraping, la agencia cartográfica productora necesita reconocer aquellos **ítems de extracción** que permiten finalmente actualizar los objetos geográficos.

Desarrollo de la metodología de extracción de datos.

La aplicación de este proceso implica 4 fases.

FASE 1. Búsqueda y selección de las URLS.

Antes de diseñar los algoritmos de extracción, el productor cartográfico debe concretar los objetos geográficos que son susceptibles de ser actualizados mediante estas técnicas, así como seleccionar las URLS de las webs a rastrear (aquéllas donde se publican numerosos ítems de interés para la temática a actualizar): ej. los objetos geográficos “carreteras” y web del Ministerio de Transportes.

FASE 2. Reconocimiento de tipologías de extracción para cada URL

Una vez preseleccionadas las webs en la fase 1, se procede a su inspección visual y a la inspección de su árbol HTML para determinar de qué tipo web se trata y qué tipología de extracción se va a implementar para cada URL.

FASE 3. Diseño e implementación de script de extracción para cada URL

En tercer lugar, se implementa el script que tendrá un conjunto de reglas, procedimientos y características que van a condicionar la extracción de datos, y que son las siguientes:

- URL de entrada para la extracción de dicha Página Web.
- Número máximo de ítems a extraer.
- Tiempo de espera, entre extracción y extracción.
- Número de solicitudes concurrentes.
- Regla horizontal de navegación (solo para extracciones horizontales y verticales).
- Regla vertical de extracción (obligatoria siempre).
- Expresiones regulares para depurar y sistematizar la extracción de los ítems.

FASE 4. Georreferenciación de los ítems extraídos

En ocasiones las webs pueden proporcionar las coordenadas de los ítems en el propio árbol HTML y su extracción es trivial, al igual que el resto de campos o características que permiten la actualización de los objetos geográficos. Si no fuera el caso, la metodología puede resolverse mediante peticiones a un Servidor de Mapas como GeoServer partiendo de las características alfanuméricas extraídas en la fase 3 (siempre y cuando de ellas se pueda interpretar y deducir datos que permitan localizar el evento en la base de datos geográfica en la que se apoya el servidor de mapas) y que permitirán a GeoServer devolver un BBOX o una geometría puntual para el ítem objeto de estudio.

FASE 5. Registro de los ítems en BBDD Postgres - PostGIS

Finalmente, se procede a guardar en una Base de datos (por ejemplo, en Postgres con extensión espacial PostGIS) los ítems extraídos en las fases precedentes y que representan elementos potenciales de cambios en el territorio por ser de nueva construcción o por que han sufrido una importante modificación.

Estos datos se podrán "paquetizar" como órdenes de trabajo, para que, posteriormente, los editores cartográficos analicen cada caso y, selectivamente, editen el objeto. Esta metodología es muy útil y optimiza los tiempos respecto a las tradicionales revisiones masivas de datos, pero no hay que olvidar que, al estar focalizadas en objetos concretos, deben ser complementadas con otras para garantizar la completitud de la actualización del conjunto de datos.

Consideraciones en Web Scraping

A la hora de diseñar, implementar y ejecutar extracciones hay que tener en cuenta un conjunto de consideraciones éticas que implican una gran responsabilidad para quien quiere extraer datos masivamente de diferentes webs. Se aconseja cumplir los siguientes principios generales de una extracción ética:

- No enviar requerimientos o solicitudes de manera desmesurada a una página web, ya que podemos colapsar la máquina servidora.
- Emplear timeouts o tiempos de espera que no sean cortos entre cada requerimiento.
- Cerciorarse de que la información sea legalmente susceptible de ser extraída y compartida por terceros.
- Limitar el número de ítems de extracción a los necesarios y programar las extracciones espaciadas en el tiempo.

Desde el punto de vista operativo, a la hora de diseñar los algoritmos de extracción de páginas web, hay que tener en cuenta que el propio ecosistema Web es cambiante y por tanto la estructura del árbol HTML o su propia URL puede ser modificada por necesidades del propietario, haciendo que los scripts de extracción ya no funcionen, lo que supone un coste de mantenimiento del código que siempre hay que tener en cuenta.

3.3.2. OSM: extracción de cambios respecto de esta fuente de datos

En general, cuando se habla de cartografía online, las primeras referencias que el gran público suele conocer son las aplicaciones desarrolladas por las grandes compañías de software (*Google*, *Bing*, *Apple*…) o bien las generadas por las instituciones oficiales de los distintos países para servir su cartografía oficial. Sin embargo, existen otras alternativas, basadas en proyectos colaborativos, cuyo objetivo es generar mapas editables y libres.

Uno de los proyectos de cartografía colaborativa más conocido es *Open Street Maps (OSM)*, el cual se enfoca principalmente en la cantidad de los datos generados, dando más importancia a los propios datos y a su actualización que a los mapas que se derivan de dichos datos.

El proyecto nació en Gran Bretaña en el año 2004, siendo una de sus mayores motivaciones el alto precio que los mapas tenían en ese momento. Además, existían numerosas restricciones de uso sobre la cartografía a la que se podía acceder, por ejemplo, la imposibilidad de corregir errores encontrados, añadir nuevos datos o emplear estos mapas en otros proyectos sin pagos adicionales por uso.

Aunque en esos años también estaban surgiendo algunos de los grandes proyectos cartográficos de internet, como por ejemplo *Google Maps*, estos proyectos se basaban en la iniciativa privada, eran deliberadamente opacos tanto en su código como en sus fuentes y tenían fuerte carácter comercial, lo que provocaba una desigual cobertura del territorio en función de los beneficios esperados, ya que siempre predominaría el trabajo de actualización sobre las grandes ciudades en detrimento de las poblaciones pequeñas.

La iniciativa reunió a una gran cantidad de usuarios que desinteresadamente han ido cartografiando el mundo y aportando al proyecto sus conocimientos y sus datos para obtener el que actualmente es uno de los proyectos de cartografía con cobertura mundial más importante. Se publica bajo una licencia "Open Database License" que ofrece a los usuarios una serie de derechos de uso, siempre y cuando se cumplan unas determinadas condiciones. Si se traslada a números, en enero de 2022 la comunidad de OSM tiene más de 8 millones de usuarios y recibe diariamente más de 3 millones de ediciones.

Si bien la mayoría de los usuarios son personas individuales que colaboran voluntariamente, en los últimos años ha habido un gran incremento del interés en la actualización y revisión de OSM por parte de grandes empresas tecnológicas e incluso de organismos oficiales. Mantener OSM con una gran calidad de datos y alta actualización, proporciona ventajas a sus respectivos negocios, ya que permiten mejorar los cálculos de ruta para empresas de transporte de viajeros o de reparto de mercancías, actualizar los datos de sus propias aplicaciones cartográficas basándose en OSM o simplemente el beneficio económico de dejar de utilizar otros proveedores de mapas que son de pago, como es el caso de *Google Maps*.

La actualización de los datos se hace a través de diferentes aplicaciones multiplataforma que permiten la edición tanto on-line como off-line y que incluyen herramientas para la validación de los datos editados.

OSM se basa en una estructura de datos con topología que se almacenan en el datum WGS84 lat/lon (EPSG:4326) y se divide en 26 temas y aproximadamente tipos de 1000 entidades. Los elementos básicos que componen el modelo son:

- Nodos (*nodes*): Son puntos que recogen posiciones geográficas.
- Vías (*ways*): Son listas ordenadas de nodos que representan líneas o polígonos.
- Relaciones (*relationships*): Son agrupaciones de nodos, vías u otras relaciones con alguna propiedad común.
- Etiquetas (*tags*): Son los elementos que contienen la información semántica asociada a los nodos, vías y relaciones. Su estructura es un par clave:valor, donde la clave da nombre a la propiedad y el valor representa el valor de la misma.

Ventajas e inconvenientes de OSM

A la hora de utilizar datos provenientes de OSM en actualización cartográfica, se deben de tener en cuenta una serie de ventajas e inconvenientes que influirán tanto en la calidad de los datos como en las decisiones que se tomarán respecto a los mismos.

Inconvenientes

Calidad irregular de los datos

El carácter voluntario de las aportaciones provoca que haya zonas con mejor o peor calidad de datos en función del número de usuarios que la conozcan o que quieran participar en su actualización. En general, zonas más pobladas o con más interés para la comunidad de usuarios, tendrán mejores datos y zonas menos pobladas o de menor interés, podrán tener unos datos de calidad inferior, aunque esto no es necesariamente así.

Además, al no existir un modelo de datos semántico en el que se determine un mínimo de atributos obligatorio para cada uno de los objetos geográficos genera grandes disparidades de completitud entre los datos. En ocasiones, se dan ejemplos de entidades contiguas o relativamente cercanas que, en función de quien las haya digitalizado, la riqueza de sus atributos es muy distinta.

Ausencia de metadatos

No existen metadatos completos sobre los elementos registrados en OSM, por lo que no es posible obtener una trazabilidad real de los datos que componen OSM. Si bien se puede acceder al historial de versiones de un determinado dato, no se puede saber qué motivó el cambio, en base a que se hizo o que fuente se utilizó.

Fuentes de datos no oficiales

Las fuentes de los datos no están respaldadas por ninguna fuente oficial, es el usuario quien decide qué fuente utiliza para la actualización y no siempre tiene que ser una fuente fiable u oficial. Este problema se acrecienta por la propia falta de metadatos.

Ventajas

Gran cantidad de datos

Es uno de los puntos fuertes de OSM, su gran cantidad de datos. En enero de 2022 existían aproximadamente 7500 millones de nodos, 830 millones de vías y 9.5 millones de relaciones repartidos por todo el mundo.

Alto ritmo de actualización

La base de datos recibe más de 3 millones de actualizaciones diarias realizadas por más de 5000 usuarios únicos al día y el ritmo de crecimiento anual de la base de datos (elementos nuevos) es superior al 10%.

Además, diferentes colectivos de usuarios organizan reuniones periódicas para cartografiar nuevas zonas y se realizan eventos para cartografiar masivamente zonas en las que ha habido desastres humanitarios durante los periodos de emergencia.

Licencia Open Source

Se basa en una licencia *Open DataBase License* que nos garantiza los siguientes derechos:

- Compartir: Copiar, distribuir y usar la base de datos.
- Crear: Producir trabajos a partir de la base de datos.
- Adaptar: Modificar, crear y construir sobre la base de datos.

Con las condiciones de:

- Atribuir: Requiere que en los trabajos derivados se cite la fuente original.
- Compartir como: Se ha de mantener al menos la misma licencia o utilizar una menos restrictiva en los trabajos derivados.
- Mantener abierto: No se podrá restringir el acceso a los datos.

Descarga de datos OSM

A medida que la cantidad de datos recogidos en OSM ha ido creciendo, el volumen de la base de datos se ha ido haciendo más difícil de manejar, lo cual sumado a las constantes operaciones que generan en ella los usuarios, hace que la descarga en bloque de la base de datos o la interacción directa no sea práctica.

Con el objetivo de mitigar este problema surgen las APIs de OSM. En la actualidad existen varias y cada una tiene un objetivo específico, hay APIs de edición que permitirán obtener y guardar datos geográficos no procesados en la base de datos, APIs para mostrar mapas embebidos en páginas web o APIs de lectura que permitirán descargar fragmentos de la base de datos para utilizar en proyectos propios.

Dentro de las APIs de descarga, una de las más utilizadas es *Overpass API*, que es una API de solo lectura que ofrece partes personalizadas seleccionadas de los datos contenidos en OSM. Actúa como una base de datos web, de modo que el cliente le envía una consulta a través de la API y recibe el conjunto de datos que corresponde a dicha consulta. A diferencia de otras API, *Overpass* está optimizado para consumir pequeños conjuntos de datos en lapsos cortos de tiempo, utilizando como criterios de búsqueda, entre otros, la ubicación de los objetos, el tipo, las propiedades de sus etiquetas o la proximidad entre objetos. Además, permitirá utilizar criterios únicos o combinarlos.

El funcionamiento de esta API es muy sencillo; de forma periódica realiza una copia completa de la zona geográfica que quiere publicar en un servidor y posteriormente, en periodos de tiempo mucho más cortos que los utilizados para la copia completa, va añadiendo archivos de cambios que contengan las modificaciones realizadas. Con estos datos ya es posible responder de manera muy ágil a las consultas que se vayan solicitando. Existen varios servidores públicos cuyas zonas pueden ser globales, es decir, contienen una copia completa de todo el territorio o pueden ser regionales, es decir, pueden contener una zona limitada, como por ejemplo un país. Otra ventaja de esta API es que es de código abierto, tiene una instalación sencilla y unos requisitos de hardware razonables, por lo que cualquiera podrá generar su propia instancia para servir una zona de interés. En la siguiente tabla se pueden ver algunos de los servidores públicos más importantes de *Overpass*.

Nombre del Servidor	**Cobertura de datos**	**URL de acceso**
Instancia principal de la API de Overpass	Global	https://overpass-api.de/api/interpreter
		https://lz4.overpass-api.de/api/interpreter
		https://z.overpass-api.de/api/interpreter

Instancia de la API Overpass Rusia	Global	https://overpass.openstreetmap.ru/api/interpreter
Instancia de la API Overpass Francia	Global	https://overpass.openstreetmap.fr/api/interpreter
Instancia de la API Overpass Suiza	Regional: Suiza	https://overpass.osm.ch/api/interpreter
Instancia de la API Overpass de Kumi Systems	Global	https://overpass.kumi.systems/api/interpreter

Tabla 3- 5: Principales servidores públicos de Overpass

Overpass está preparado para interpretar dos lenguajes distintos, por un lado, el lenguaje XML que nos permitirá enviar peticiones HTTP POST y por otro lado el lenguaje *Overpass Query Lenguage* (QL) nos permitirá la obtención de datos utilizando peticiones HTTP GET a través de URL.

Adicionalmente existe un tercer intérprete, que utiliza la capa de compatibilidad XAPI (*Extended API)* y que permite utilizar una sintaxis modificada de la API principal de OSM (la utilizada para edición) pero con comandos más poderosos y a la vez más complejos que se ejecutaban a través de HTTP GET. Aunque sigue siendo interpretada, actualmente se desaconseja su uso en nuevos proyectos ya que su bajo rendimiento ha hecho que deje de haber soporte técnico para su uso.

Consultas con lenguaje QL

La forma más sencilla de realizar una descarga de datos procedentes de OSM con *Overpass* es utilizar el lenguaje QL a través de cualquier navegador que tengamos instalado. En la siguiente imagen se muestra un ejemplo de consulta.

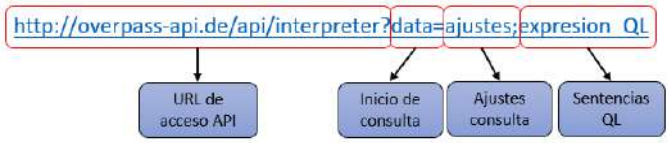


Ilustración 3- 19: Sintaxis básica de llamadas QL a Overpass

En primer lugar, está la URL de la API a la que se va a dirigir la consulta, en este caso se ha utilizado la URL de la API principal, seguidamente irá la palabra clave *data* seguida de un igual que dará comienzo a la consulta y seguidamente irían los ajustes de la consulta, que son opcionales y sirven para definir parámetros como el formato de archivos de descarga, el tiempo máximo de ejecución, el tamaño máximo de los elementos, etc. Finalmente, se pondrán las sentencias en lenguaje QL que serán las que seleccionen los datos a descargar.

Ajustes

Los ajustes van a permitir personalizar de manera avanzada los resultados de las consultas. Entre estos ajustes se pueden encontrar:

- **Timeout**: valor entero positivo que define el tiempo máximo de ejecución de una consulta, cuyo valor por defecto es 180. Si se supera ese tiempo el servidor devolverá un error de *timeout* y se parará la consulta. Si las capacidades del servidor al que se

realiza la consulta son limitadas, modificar este tiempo para ponerlo más alto puede hacer que el servidor rechace la consulta antes de realizarla. Ejemplo: [timeout:180]

- **Límite de elementos**: valor entero positivo que define el máximo de memoria que podrá ocupar la consulta en el servidor, por defecto es 536870912 (512 MB). Si se supera este valor se cancelará la ejecución y si se establece un valor más alto, el servidor puede rechazar la consulta antes de realizarla. Ejemplo: [maxsize:536870912]
- **Formato de salida**: define el formato de salida de la consulta, por defecto *xml*. Los formatos disponibles son *xml*, *json*, *csv*, *custom* y *popup*. Ejemplo: [output:json]
- **Bounding box global**: define el área del mapa sobre el que se realizarán la consulta, si no se especifica, por defecto, se realiza sobre todo el planeta. Este *bounding box* se aplica a todas las consultas que se hagan en la llamada, de tal modo que, si en alguna de las consultas se pone otro *bounding box*, se obtendrá como resultado la intersección entre ambos. El orden de las coordenadas es sur, oeste, norte, este. Ejemplo: [bbox:-23,-43.3,-22.8,-43.1].
- **Fecha**: establece una fecha para las consultas, por defecto la actual, de modo que los resultados que se obtienen de la consulta son el estado de la base de datos en esa determinada fecha. Ejemplo: [date:"2015-10-28T19:20:00Z"]
- **Diferencia entre dos fechas**: Permite descargar solo las diferencias de la base de datos entre dos fechas concretas. La fecha de partida es obligatoria y la fecha de fin es opcional, ya que en su ausencia se tomará la actual. El valor por defecto de este parámetro es desde el inicio de los tiempos hasta el momento actual de la consulta. Ejemplo: [diff:"2012-09-14T15:00:00Z","2012-09-21T15:00:00Z"]
- **Diferencia "aumentada" entre dos fechas**: Funciona igual que el parámetro anterior, pero añade información sobre qué ha pasado con los elementos que no están en la última fecha, añadiendo al resultado la fecha de eliminación y la indicación "*visible=false*". Ejemplo: [adiff:"2012-09-14T15:00:00Z","2012-09-21T15:00:00Z"]

Expresiones QL

Las expresiones en lenguaje QL son un conjunto de sentencias separadas por el carácter punto y coma (;) que comenzarán con la palabra clave *node*, *way*, *relation* o *área* para definir el tipo de elemento de OSM que queremos obtener y terminará con la sentencia *out;* para indicar el final de la consulta. Todas las sentencias que aparezcan en una expresión QL se interpretarán de forma secuencial. Un ejemplo sencillo sería la siguiente expresión:

```
https://overpass-api.de/api/interpreter?data=node(50.745,7.17,50.75,7.18);out;
```

Esta expresión consta de dos sentencias, la sentencia *node(50.745,7.17,50.75,7.18);* que obtiene el conjunto de nodos que está contenido en el rectángulo que forman las coordenadas que se pasan como parámetro.

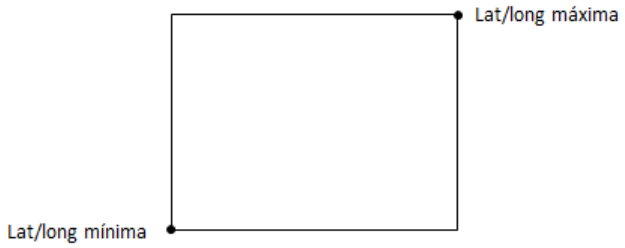


Ilustración 3- 20: Rectángulo de consulta

A continuación, la sentencia *out;* da por finalizada la expresión y prepara los datos para su descarga en el equipo cliente. Se puede utilizar la sentencia *out meta*; en lugar de *out;* y se añadirán los datos de autor, fecha de edición y versión a los datos descargados.

Como ya se comentó anteriormente, las sentencias QL que componen una expresión se ejecutan de forma secuencial. Véase el siguiente ejemplo (omitimos la URL del intérprete para ver la expresión en una sola línea):

```
data=node(50.745,7.175,50.746,7.176);node(50.749,7.179,50.75,7.18);out;
```

En este caso, el intérprete haría la consulta de los nodos en el primer rectángulo, luego realizaría la consulta de los nodos en el segundo rectángulo y finalmente ejecutaría la descarga. Al ejecutarse las órdenes de modo secuencial, lo que se va a suceder es que cuando se ejecute la segunda orden, se eliminan de memoria los resultados de la anterior y la sentencia *out*; solo devuelve los datos de la segunda consulta.

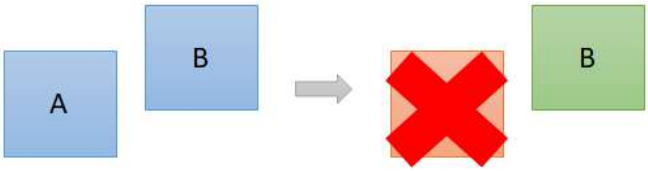


Ilustración 3- 21: Ejecución secuencial de sentencias QL

El lenguaje QL permite agrupar los resultados de varias sentencias. Para ello tan solo hay que agrupar las sentencias dentro de un paréntesis, de este modo se ejecutan de manera independiente y se suman los resultados. Se puede ver en la siguiente expresión.

```
data=(node(50.745,7.175,50.746,7.176);node(50.749,7.179,50.75,7.18););out;
```

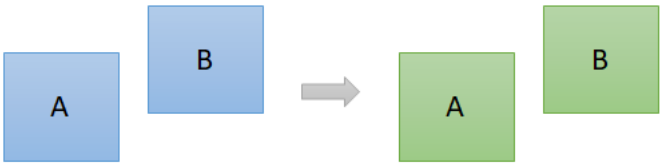


Ilustración 3- 22: Unión de sentencias QL

Del mismo modo que se pueden sumar sentencias, también se pueden ejecutar restas. Este método permitirá ejecutar las dos sentencias y restar a los resultados de la sentencia A, los resultados de la sentencia B, de modo que solamente se obtendrán los elementos de A que a su vez no estén en B.

```
data=(node(50.745,7.175,50.746,7.176);-node(50.7455,7.1755,50.7465,7.1765););out;
```

Ilustración 3- 23: Resta de sentencias QL (1)

```
data=(node(50.745,7.175,50.746,7.176);-node(50.7452,7.1752,50.7458,7.1758););out;
```

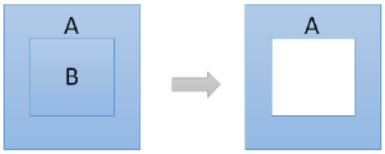


Ilustración 3- 24: Resta de sentencias QL (2)

Estas son las opciones de trabajo con conjuntos de sentencias más sencillas, sin embargo, el lenguaje QL nos permite hacer expresiones más complejas, en las que se pueden hacer condicionales, bucles o comparaciones entre otras cosas. Se pueden encontrar más detalles en wiki de OSM https://wiki.openstreetmap.org/wiki/Overpass_API/Overpass_QL

A la hora de hacer las consultas, además de filtrar espacialmente, en muchas ocasiones se necesitará establecer filtros más concretos, para obtener solamente los objetos espaciales que realmente interesen a los objetivos del proyecto. Por ejemplo, si solo se desease descargar las carreteras de una zona, en ese caso se tendría que especificar en primer lugar el tipo de geometría que se quiere descargar (*way*) y además de añadir las coordenadas de la zona que se desea descargar, se deberá añadir la etiqueta de los elementos a obtener, como se ve en la siguiente sentencia.

```
data=way[highway](50.745,7.175,50.746,7.176);out;
```

Sin embargo, al ejecutar esa consulta, se obtendrá una carretera, el listado de nodos que la componen y ninguna coordenada. El motivo es que, por la definición del modelo, las líneas son conjuntos de nodos y los nodos son los objetos que tienen las coordenadas, al pedir solamente las líneas, las devuelve, pero como en base de datos las vías no almacenan coordenadas, no las devolverá. Hay una manera muy simple de solucionar esto y es utilizar el operador “>” que devolverá todos los elementos que componen los elementos devueltos por la primera consulta. Se utiliza como se ve en la siguiente consulta.

```
data=(way[highway](50.745,7.175,50.746,7.176);>;);out;
```

Se puede ver que en la consulta se añaden los paréntesis para agrupar, ya que internamente la expresión se interpreta como dos sentencias, una primera sentencia que devuelve los elementos de tipo *highway* y una segunda consulta que devuelve los elementos relacionados con los elementos de la primera. Si no se utilizan los paréntesis para realizar la unión, las sentencias se ejecutarían secuencialmente y solo se obtendrían los resultados de la segunda sentencia. Por tanto, si realmente el objetivo fuese conseguir los nodos que forman las carreteras, pero no las carreteras como tal, la sentencia sería la siguiente.

```
data=way[highway](50.745,7.175,50.746,7.176);>;out;
```

Es importante no confundir el operador “>” con el operador “<” ya que tienen distintos significados. El primero, que es el visto en el ejemplo anterior, devolverá todos los elementos que componen las entidades de un conjunto de datos, mientras que el segundo devolverá todos los elementos que están compuestos por los elementos de ese conjunto de datos. Véase el siguiente ejemplo.

```
data=(node(50.745,7.175,50.746,7.176);<;);out;
```

Con esta expresión, se obtendrían todos los nodos y además todas las vías y relaciones de las que alguno de esos nodos formase parte.

Aún es posible filtrar más los datos, de modo que se podrían obtener tan solo un tipo de carreteras, por ejemplo, las residenciales. Sería muy sencillo, ya que sólo habría que añadir a la clave *highway* un valor.

```
data=(way[highway=residential](50.745,7.175,50.746,7.176);>;);out;
```

En caso de que el objetivo fuese el contrario, es decir, obtener todas las carreteras menos las de tipo residencial, se utilizaría una resta de sentencias.

```
((way[highway](50.745,7.175,50.746,7.176);-
way[highway=residential](50.745,7.175,50.746,7.176););>;);out;
```

En estos casos hay que tener mucha precaución con los paréntesis, ya que en un primer paréntesis se hace la resta de elementos y en el segundo se hace la suma del resultado de la resta más los elementos relacionados.

Como último punto, hay que recordar que con los ajustes se podrán establecer filtros temporales, es decir, se podrá conocer el estado de la base de datos en determinados puntos temporales o conocer los cambios producidos entre un rango de fechas determinado. Por ejemplo, en la siguiente sentencia, se obtendrá como resultado los elementos que estaban en la base de datos a 1 de enero de 2017:

```
data=[date:"2017-01-01T00:00:00Z"];node(50.745,7.175,50.746,7.176);out meta;
```

Se ha sustituido la sentencia *out;* por *out meta;* para poder ver el *timestamp* de los elementos en la salida y así comprobar que no hay ningún elemento con fecha posterior a la establecida en la consulta. Del mismo modo, se podrían obtener los cambios en la base de datos entre el 1 de enero de 2017 y el 1 de enero de 2018.

```
data=[diff:"2017-01-01T00:00:00Z","2018-01-
01T00:00:00Z"];node(50.745,7.175,50.746,7.176);out;
```

El lenguaje QL es muy rico y permite muchas expresiones, pero no es objetivo fundamental de este temario, por lo que se remite nuevamente a la documentación de la wiki de OSM para más información.

Consultas XML

Las consultas XML se realizan por el método HTTP POST y se necesita enviar un archivo al servidor para que lea su contenido y devuelva un archivo con la respuesta. En este caso no se podrá hacer directamente desde el navegador si no que se necesitará software específico para enviar los archivos o utilizar un lenguaje de programación. Los archivos XML que se utilizarán son muy sencillos. A continuación, se puede ver un ejemplo de código.

```
<osm-script>
        <query type="node">
                <bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
        </query>
        <print/>
</osm-script>
```

Para enviar los archivos al servidor, se utilizará un script de Python con el código que se puede ver a continuación.

```
import requests

fichero_xml = r'D:\PruebasOSM.xml'
fichero_salida = r'D:\SalidaConsulta.xml'
ruta_interprete = 'http://overpass-api.de/api/interpreter'

with open(fichero_xml, 'rb') as f:
  r = requests.post(ruta_interprete, data=f)

with open(fichero_salida, 'wb') as sal:
        sal.writelines(r)
```

En caso de no querer utilizar Python para realizar las descargas, se puede acceder a la herramienta https://overpass-turbo.eu/ que permite ejecutar expresiones para probar los resultados que se obtendrían.

Ajustes

Los ajustes que se utilizan son semejantes a los que se usan en el lenguaje QL y se utilizan añadiéndolos como atributos a la etiqueta *<osm-script>*. Se exponen algunos ejemplos de uso de los elementos que coinciden con el lenguaje QL y se añade explicación de algunos elementos adicionales:

- **Timeout**: Ejemplo: <osm-script timeout="180">.
- **Formato de salida**: Ejemplo: <osm-script output="json">
- **Límite de elemento**: Ejemplo: <osm-script element-limit="10824">
- **Fecha**: Ejemplo: <osm-script date="2014-09-21T15:00:00Z">
- **Fecha de referencia**: Establece una fecha a partir de la cual se tienen en cuenta los elementos para la consulta, es decir, si se establece esta fecha se obtendrán todos los objetos incluidos o modificados desde esta fecha. Combinado con el atributo fecha (date), permite obtener cambios en la base de datos de OSM para rangos de tiempo determinados. Ejemplo: <osm-script from="2014-09-14T15:00:00Z">

Expresiones XML

La funcionalidad que se puede obtener con el uso de archivos XML y peticiones post frente a las peticiones get en QL es muy similar. Al contrario que en QL, donde toda la expresión iba en una sola línea, en XML se ha de utilizar el característico anidamiento de etiquetas del lenguaje, de modo que cada consulta irá dentro de una etiqueta *<query>*.

En esta parte, se va a explicar cómo hacer en XML los mismos ejemplos que se han mostrado en QL. En el primer ejemplo de QL se explicaba como descargar todos los nodos incluidos en una determinada área. En XML el código sería el siguiente.

```
<osm-script>
        <query type="node">
                <bbox-query e="7.18" n="50.75" s="50.745" w="7.17"/>
        </query>
        <print/>
</osm-script>
```

Interpretando el código, se puede ver una etiqueta genérica *<osm-script>* que llevarán todos los archivos XML que se envíen al servidor, posteriormente una etiqueta *<query>* a la que se añade el atributo *type* que será el que especifique el tipo de geometría que se desea descargar. Anidada a la etiqueta *<query>* está la etiqueta *<bbox-query>* cuyos atributos son las coordenadas del área a descargar. Una vez cerrada la etiqueta *<query>* está la etiqueta *<print>* que hace la misma función que la sentencia *out* en QL, es decir, escribir los resultados en el archivo que se devuelve al cliente. Al igual que en QL, se pueden pedir los datos de autor, fecha de edición y versión añadiendo el atributo *mode* con el valor *meta* a la etiqueta, es decir: <print mode="meta"/>

Al igual que en QL, las sentencias se ejecutan secuencialmente y en caso de que no se hagan operaciones con los resultados, se mostrarán los resultados de la última sentencia. Véase la siguiente expresión como ejemplo.

```
<osm-script>
        <query type="node">
                <bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
        </query>
        <query type="node">
                <bbox-query e="7.18" n="50.75" s="50.749" w="7.179"/>
        </query>
        <print/>
</osm-script>
```

Si se desea la suma de las dos consultas, habría que anidar las consultas dentro de una etiqueta *<union>*.

```
<osm-script>
	<union>
		<query type="node">
			<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
		</query>
		<query type="node">
			<bbox-query e="7.18" n="50.75" s="50.749" w="7.179"/>
		</query>
	</union>
	<print/>
</osm-script>
```

Al igual que en QL, se podrán realizar restas de sentencias, en ese caso, se sustituirá la etiqueta *<union>* por la etiqueta *<difference>*.

```
<osm-script>
	<difference>
		<query type="node">
			<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
		</query>
		<query type="node">
			<bbox-query e="7.18" n="50.75" s="50.749" w="7.179"/>
		</query>
	</difference>
	<print/>
</osm-script>
```

En ocasiones se pueden querer filtrar los elementos por sus etiquetas, en ese caso habrá que anidar dentro de *<query>* la etiqueta *<has-kv>* con el atributo *k*.

```
<osm-script>
	<query type="way">
		<has-kv k="highway"/>
		<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
	</query>
	<print/>
</osm-script>
```

Tal como sucedía en QL, esta consulta devolverá todos los objetos de clase *highway* pero sin las coordenadas de sus nodos, para obtenerlas habrá hacer unión de dos consultas, pero la segunda será simplemente una etiqueta *<recurse>* con el atributo *type="down"* que devolverá todos los elementos que forman los elementos de la primera consulta, es decir, todos los nodos que forman las vías de la primera consulta.

```
<osm-script>
	<union>
		<query type="way">
			<has-kv k="highway"/>
			<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
		</query>
		<recurse type="down"/>
	</union>
	<print/>
</osm-script>
```

Esta etiqueta funciona como los operadores “<” y “>” de QL. De este modo, *type=”down”* hará la misma función que “>”, es decir, devolver todos los elementos que componen los elementos de la consulta anterior y *type=”up”* hará la misma función que “<”, es decir, devolver todos los elementos que están compuestos por los elementos de la consulta anterior.

Para filtrar aún más los datos y obtener sólo un determinado grupo dentro de un conjunto, por ejemplo, las vías residenciales dentro de los elementos de la clase *highway* habrá que añadir el atributo *v* a la etiqueta *<has-kv>*.

```
<osm-script>
	<query type="way">
		<has-kv k="highway" v="residential"/>
		<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
	</query>
	<print/>
</osm-script>
```

En caso de querer todas menos las de tipo residencial, se usaría nuevamente la resta.

```
<osm-script>
	<difference>
		<query type="way">
			<has-kv k="highway"/>
			<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
		</query>
		<query type="way">
			<has-kv k="highway" v="residential"/>
			<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
		</query>
	</difference>
	<print/>
</osm-script>
```

También se podrán hacer consultas a una fecha concreta de la base de datos.

```
<osm-script date="2018-01-01T00:00:00Z">
	<query type="node">
		<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
	</query>
	<print mode="meta"/>
</osm-script>
```

Y finalmente, consultar a un intervalo concreto de tiempo.

```
<osm-script date="2018-01-01T00:00:00Z" from="2017-01-01T00:00:00Z">
	<query type="node">
		<bbox-query e="7.176" n="50.746" s="50.745" w="7.175"/>
	</query>
	<print mode="meta"/>
</osm-script>
```

Limitaciones en las descargas

Un punto importante que hay que considerar a la hora de utilizar la API de OSM son sus limitaciones, por la naturaleza libre del servicio se han establecido límites tanto en los tamaños de archivo que se pueden descargar, como en el número de peticiones que se pueden hacer.

Si se va a utilizar OSM como fuente de datos en un proyecto, descargando datos de un servidor público y se requiere usar grandes extensiones de terreno, como por ejemplo un país, es necesario tenerlo en cuenta. Es recomendable dividir el terreno en zonas, ya sea por divisiones administrativas o por cuadriculas sobre el territorio, siendo esta última la manera más sencilla de hacerlo. Además, por el límite de peticiones, no se recomiendan hacer peticiones en paralelo, se recomienda serializarlas y para mayor seguridad espaciarlas unos segundos en el tiempo.

En caso de usar un script de descarga automática, se recomienda establecer mecanismos que comprueben que los archivos descargados son válidos antes de continuar con los siguientes. Un ejemplo de descarga con comprobación, a partir del script que se propone para la descarga de archivos con consultas XML sería el siguiente.

```
import requests
import time

fichero_xml = r'D:\PruebasOSM.xml'
fichero_salida = r'D:\SalidaConsulta.xml'
ruta_interprete = 'http://overpass-api.de/api/interpreter'

descarga_correcta = False
while not descarga_correcta:
   with open(fichero_xml, 'rb') as f:
      try:
         r = requests.post(ruta_interprete, data=f)
      except:
         print("Intento de descarga fallido")

      if r.content.decode('utf-8').__contains__('Error'):
         print("Intento de descarga fallido")
      else:
         with open(fichero_salida, 'wb') as sal:
            sal.writelines(r)
         descarga_correcta = True
         print("Fichero descargado correctamente")
      time.sleep(5)
```

A los comandos previos se ha añadido una comprobación que hace que el proceso de descarga se lance recursivamente en caso de error, por un lado, con las cláusulas *try/except* se monitorizan los errores de conexión que se pudieran dar y con un *if* se busca la palabra “Error” dentro del contenido del archivo descargado.

Aunque este código hará que se ejecuten los intentos necesarios para descargar los archivos, si el error que se obtiene no es por la sobrecarga del servidor y es por algún error en la generación de los archivos XML, entrará en un bucle infinito, por lo que se recomienda añadir algún contador de intentos para para la ejecución del bucle en caso de error recurrente.

Tratamiento de datos descargados

Una vez descargados los archivos de datos de OSM a través de la API, el siguiente paso será convertirlos en archivos con geometría que sean más sencillos de trabajar en programas SIG y que se puedan compartir con otros usuarios de la organización. La mejor manera de hacer estas conversiones es automatizarlas mediante sencillos scripts de Python utilizando las librerías *osmread, shapely* y *gdal*.

Para este ejemplo, se usará un archivo con todas las vías de una zona y todos los nodos que forman dichas vías. Se descargará con el archivo XML que se encuentra a continuación.

```
<osm-script>
        <union>
                <query type="way">
                        <bbox-query e="-69.225" n="18.742" s="18.730" w="-69.240"/>
                </query>
                <recurse type="down"/>
        </union>
        <print mode="meta"/>
</osm-script>
```

El siguiente fragmento de código Python muestra un ejemplo de código para abrir un archivo XML, interpretarlo y extraer los nodos.

```
import osmread

fichero_descargado = r'D:\SalidaConsulta.xml'
osm = osmread.parse_file(fichero_descargado)

for el in osm:
   if isinstance(el, osmread.Node):
      print(el)
```

Su funcionamiento es sencillo, llama a la función *parse_file* de la librería *osmread* y convierte los datos en un conjunto de objetos de OSM. Posteriormente, se recorre el conjunto de objetos y se pide que se muestren por pantalla los elementos que sean de tipo nodos. El resultado sería semejante al que se ve a continuación (sólo se muestra un nodo).

```
Node(id=359156755, version=4, changeset=82204998, timestamp=1584223684,
uid=72235, tags={'power': 'tower'}, lon=-69.342427, lat=18.4928011)
```

El objeto nodo es un objeto cuyas propiedades son las siguientes:

- ID: Identificador único en la base de datos.
- *Version*: Versión del elemento en la base de datos, 1 si es una nueva inserción, va incrementando con cada edición del elemento.
- C*hangeset*: Identificador de la sesión en la que se dio de alta o modifico el elemento.
- *Timestamp*: Fecha del cambio en segundos.
- *UID*: Identificador del usuario que realizo el cambio.
- *Tags*: Diccionario de atributos del elemento, serán pares clave:valor y habrá tantos como atributos tenga el elemento.
- *Lon*: Coordenada longitud del elemento.
- *Lat*: Coordenada latitud del elemento.

En el caso de los nodos, la generación de geometrías es sencilla, solamente habrá que utilizar el módulo *geometry* de la librería *shapely*. Este módulo permite crear puntos en WKT a partir de pares de coordenadas.

```
import osmread
from shapely.geometry import Point

fichero_descargado = r'D:\SalidaConsulta.xml'
osm = osmread.parse_file(fichero_descargado)

for el in osm:
   if isinstance(el, osmread.Node):
      pt_wkt = Point(el.lon, el.lat)
      print(pt_wkt)
```

Una vez obtenida la geometría en WKT, ya se podría escribir a un archivo. Sin embargo, en este momento surge un problema con los atributos que tendrá la tabla de datos del archivo, ya que, aunque hay unos atributos fijos, el atributo *tags* contiene todos los atributos específicos de cada uno de los elementos y dado que estos atributos específicos son voluntarios en OMS, la variabilidad es muy grande. Una posible solución, si no se conocen profundamente los datos es crear tantos atributos en el archivo de salida como atributos distintos existan el archivo descargado. Para ello habría que leer todos los campos que hay en los distintos elementos e ir almacenándolos en una lista. En el siguiente fragmento de código se modifica el bucle de lectura para almacenar los campos a crear.

```
campos_crear = []
for el in osm:
   if isinstance(el, osmread.Node):
      pt_wkt = Point(el.lon, el.lat)
      for t in el.tags:
         if t not in campos_crear:
            campos_crear.append(t)
print(campos_crear)
```

En este caso solo existe el campo *power*. Cuando se genere el archivo de líneas, se verá que la variabilidad es más elevada.

A continuación, se generaría la capa con el módulo *ogr* de la librería *gdal* y posteriormente se rellena con los datos descargados. Para no tener que leer dos veces los nodos desde el archivo, durante la lectura de los campos, se añaden los nodos a escribir a un diccionario. Además de los campos definidos por los *tags* de cada nodo, se añade el identificador de cada nodo al archivo. Durante la ejecución se puede mostrar algún *warning* de la librería *gdal*, es debida al tamaño de los enteros que forman los identificadores pero no supone ningún problema.

```
import osmread
from shapely.geometry import Point, LineString, Polygon
from osgeo import ogr, osr

fichero_descargado = r'D:\SalidaConsulta.xml'
osm = osmread.parse_file(fichero_descargado)

campos_crear = []
nodos = {}
for el in osm:
   if isinstance(el, osmread.Node):
      nodos[el.id] = el
      for t in el.tags:
         if t not in campos_crear:
            campos_crear.append(t)

fichero_shp = r'D:\Ejemplo_Nodos.shp'
driver = ogr.GetDriverByName('Esri Shapefile')
ds = driver.CreateDataSource(fichero_shp)
srs = osr.SpatialReference()
srs.ImportFromEPSG(4326)
layer = ds.CreateLayer('', srs, ogr.wkbPoint)
layer.CreateField(ogr.FieldDefn('id', ogr.OFTInteger))
for campo in campos_crear:
   layer.CreateField(ogr.FieldDefn(campo, ogr.OFTString))
defn = layer.GetLayerDefn()

for n in nodos:
   nodo = nodos[n]
   pt_wkt = Point(nodo.lon, nodo.lat)
   campos = nodo.tags
   feature = ogr.Feature(defn)
   feature.SetField('id', nodo.id)
   for campo in campos:
      feature.SetField(campo, campos[campo])
   geom = ogr.CreateGeometryFromWkb(pt_wkt.wkb)
   feature.SetGeometry(geom)
   layer.CreateFeature(feature)
```

Para escribir las vías en un archivo, el proceso sería semejante, sin embargo, si se accede a cualquiera de los objetos de tipo *way* que se generan en la lectura del archivo XML se puede observar que esos elementos no tienen geometría por sí mismos.

```
osm = osmread.parse_file(fichero_descargado)
for el in osm:
   if isinstance(el, osmread.Way):
      print(el)
```

```
Way(id=370901897, version=2, changeset=70824005, timestamp=1559340036,
uid=3782655, tags={'landuse': 'residential'}, nodes=(3745697505, 3745697503,
6514239200, 3745697502, 3745697499, 3745697506, 3745697497, 6514239201,
3745697501, 3745697500, 6514239202, 3745697504, 3745697505))
```

En lugar de geometría, la vía se compone de un listado de nodos, que son los que realmente contienen las geometrías. Por lo tanto, siempre que se vaya a escribir un archivo de vías, es necesario leer previamente los nodos.

Otro problema que aparecerá es que en el conjunto de vías se encuentran tanto los elementos lineales como los elementos superficiales, por lo que habrá que filtrarlos antes de escribirlos. Para saber si un elemento es lineal o superficial habrá que comprobar si su punto inicial y final es el mismo. Los siguientes fragmentos de código irán a continuación del código que genera el archivo de puntos y generarán los archivos de líneas y polígonos.

```
osm = osmread.parse_file(fichero_descargado)
lineas_escribir = {}
areas_escribir = {}
campos_areas = []
campos_lineas = []

for el in osm:
    if isinstance(el, osmread.Way):
        nodos_way = []
        if el.nodes[0] != el.nodes[-1]:
            lineas_escribir[el.id] = el
            for t in el.tags:
                if t not in campos_lineas:
                    campos_lineas.append(t)
        else:
            areas_escribir[el.id] = el
            for t in el.tags:
                if t not in campos_areas:
                    campos_areas.append(t)
```

En este primer fragmento se vuelve a cargar el archivo XML (la función *parse_file* genera un objeto de *Generator* de Python y no es reutilizable) y posteriormente se buscan todos los elementos de tipo *way* separándolos en dos grupos, uno de líneas y otro de áreas. Además de los elementos, se almacenan los distintos atributos de cada elemento.

```
fichero_shp = r'D:\Ejemplo_Lineas.shp'
driver = ogr.GetDriverByName('Esri Shapefile')
ds = driver.CreateDataSource(fichero_shp)
srs = osr.SpatialReference()
srs.ImportFromEPSG(4326)
layer = ds.CreateLayer('', srs, ogr.wkbLineString)
layer.CreateField(ogr.FieldDefn('id', ogr.OFTInteger))
for campo in campos_lineas:
   layer.CreateField(ogr.FieldDefn(campo, ogr.OFTString))
defn = layer.GetLayerDefn()
```

Este fragmento generará el archivo *shape* vacío en el que posteriormente se almacenarán las líneas.

```
for l in lineas_escribir:
   linea = lineas_escribir[l]
   nodos_linea = linea.nodes
   lista_nodos = []
   for nodo in nodos_linea:
      lista_nodos.append((nodos[nodo].lon, nodos[nodo].lat))
   lin_wkt = LineString(lista_nodos)
   campos = linea.tags
   feature = ogr.Feature(defn)
   feature.SetField('id', linea.id)
   for campo in campos:
      feature.SetField(campo, campos[campo])
   geom = ogr.CreateGeometryFromWkb(lin_wkt.wkb)
   feature.SetGeometry(geom)
   layer.CreateFeature(feature)
```

Al igual que los puntos, las líneas se escribirán en el archivo, con la diferencia de que, en el momento de generar los nodos, se toman los identificadores de los mismos y se consultan sus coordenadas en el diccionario de nodos que se creó al inicio.

En un último fragmento de código se genera y rellena el archivo de polígonos.

```
fichero_shp = r'D:\Ejemplo_Areas.shp'
driver = ogr.GetDriverByName('Esri Shapefile')
ds = driver.CreateDataSource(fichero_shp)
srs = osr.SpatialReference()
srs.ImportFromEPSG(4326)
layer = ds.CreateLayer('', srs, ogr.wkbPolygon)
layer.CreateField(ogr.FieldDefn('id', ogr.OFTInteger))
for campo in campos_areas:
   layer.CreateField(ogr.FieldDefn(campo, ogr.OFTString))
defn = layer.GetLayerDefn()

for a in areas_escribir:
   area = areas_escribir[a]
   nodos_area = area.nodes
   lista_nodos = []
   for nodo in nodos_area:
      lista_nodos.append((nodos[nodo].lon, nodos[nodo].lat))
   area_wkt = Polygon(lista_nodos)
   campos = area.tags
   feature = ogr.Feature(defn)
   feature.SetField('id', area.id)
   for campo in campos:
      feature.SetField(campo, campos[campo])
   geom = ogr.CreateGeometryFromWkb(area_wkt.wkb)
   feature.SetGeometry(geom)
   layer.CreateFeature(feature)
```

Al final del proceso deberían obtenerse tres archivos *shape* con puntos, líneas y polígonos que se podrá visualizar en cualquier SIG.

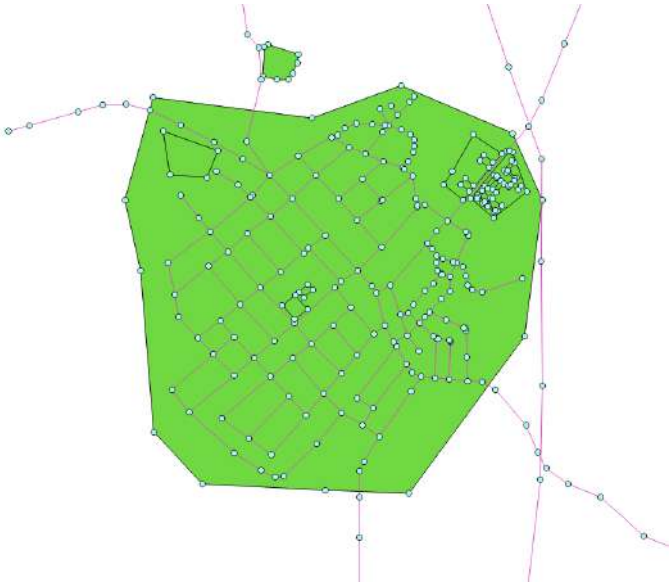


Ilustración 3- 25: Archivos SHP generados con datos descargados de OSM

3.3.3. WFS: extracción de datos desde servicios de descarga

Otra prometedora fuente de información geográfica disponible en la red son los servicios web basados en estándares OGC. Numerosos organismos han habilitado servicios que permiten acceder a sus datos abiertos a través de protocolos normalizados. Entre los diversos tipos de servicio existentes, merece especial atención el llamado Web Feature Service (WFS), ya que:

- Proporciona información geográfica de naturaleza vectorial.
- La información se distribuye en un conjunto de clases de entidad.
- Cada entidad posee geometría y atributos.
- El servicio permite realizar consultas basadas en criterios espaciales y temáticos.
- La comunicación se realiza por medio de HTTP.

Por todas estas características, los servicios WFS se prestan bien a la creación de sistemas automáticos de extracción de información orientados a la detección de cambios en el territorio, con una filosofía parecida a la aplicada en el apartado dedicado al web scraping, pero con las ventajas que proporciona la normalización del lenguaje de intercambio de información.

Un factor que condiciona radicalmente el diseño del sistema de detección de cambios basado en WFS es la existencia de atributos que informen sobre el ciclo de vida de las entidades (fecha de creación, borrado, etc.). Cuando estos atributos existen y el servicio WFS permite realizar consultas sobre los mismos, el propio servidor será el encargado de proporcionar el conjunto de cambios ocurridos durante el periodo de interés. La responsabilidad del cliente en este caso se reduce a formular una consulta con los parámetros adecuados, basándose en la fecha de la consulta anterior y en la actual. En cambio, si el servicio no admite consultas sobre el ciclo de vida de las entidades, la detección de altas, bajas y modificaciones recae en el cliente. Será necesario en tal caso descargar el conjunto completo de datos cada vez que se ejecute la consulta y efectuar un proceso de comparación con la copia local procedente de la anterior extracción.

La necesidad de solicitar grandes conjuntos de datos choca a menudo con la falta de idoneidad del servicio para las descargas masivas. A partir de la versión 2.0.0, el estándar WFS permite las respuestas paginadas. Gracias a ello, consultas que devuelven un volumen de datos importante pueden realizarse sin necesidad de descargar todos los registros de una sola vez. Mediante la paginación de la respuesta, el cliente puede desplazarse de forma secuencial por el conjunto de resultados, accediendo en cada página a un grupo de registros de tamaño fijo. Sin embargo, algunos servicios disponibles no implementan la versión 2.0.0 o, directamente, limitan el tamaño de sus respuestas a un número máximo de registros, lo que obliga al cliente a buscar soluciones alternativas para obtener el conjunto completo de resultados.

Ejemplo buenas prácticas de un sistema de extracción de datos

El IGN de España ha desarrollado un sistema de obtención de información geográfica mediante el envío de consultas periódicas al servicio WFS de la Dirección General de Catastro. El propósito del sistema es establecer un mecanismo de vigilancia sobre la capa de construcciones catastrales, con vistas a la actualización de la tabla de edificaciones de la Base Topográfica Nacional (BTN) a partir de las variaciones detectadas. Dado el enorme volumen de cambios que se identifican diariamente (unos 40.000 movimientos entre altas y bajas, como

promedio diario), no se ha planteado por el momento desencadenar órdenes de trabajo a partir de las variaciones individuales detectadas. En vez de llevar a cabo esa gestión individualizada, se prefiere obtener el conjunto de variaciones acumuladas a lo largo de un periodo de tiempo y utilizarlo como referencia para realizar procesos de actualización de ámbito provincial, aplicando técnicas de conflación.

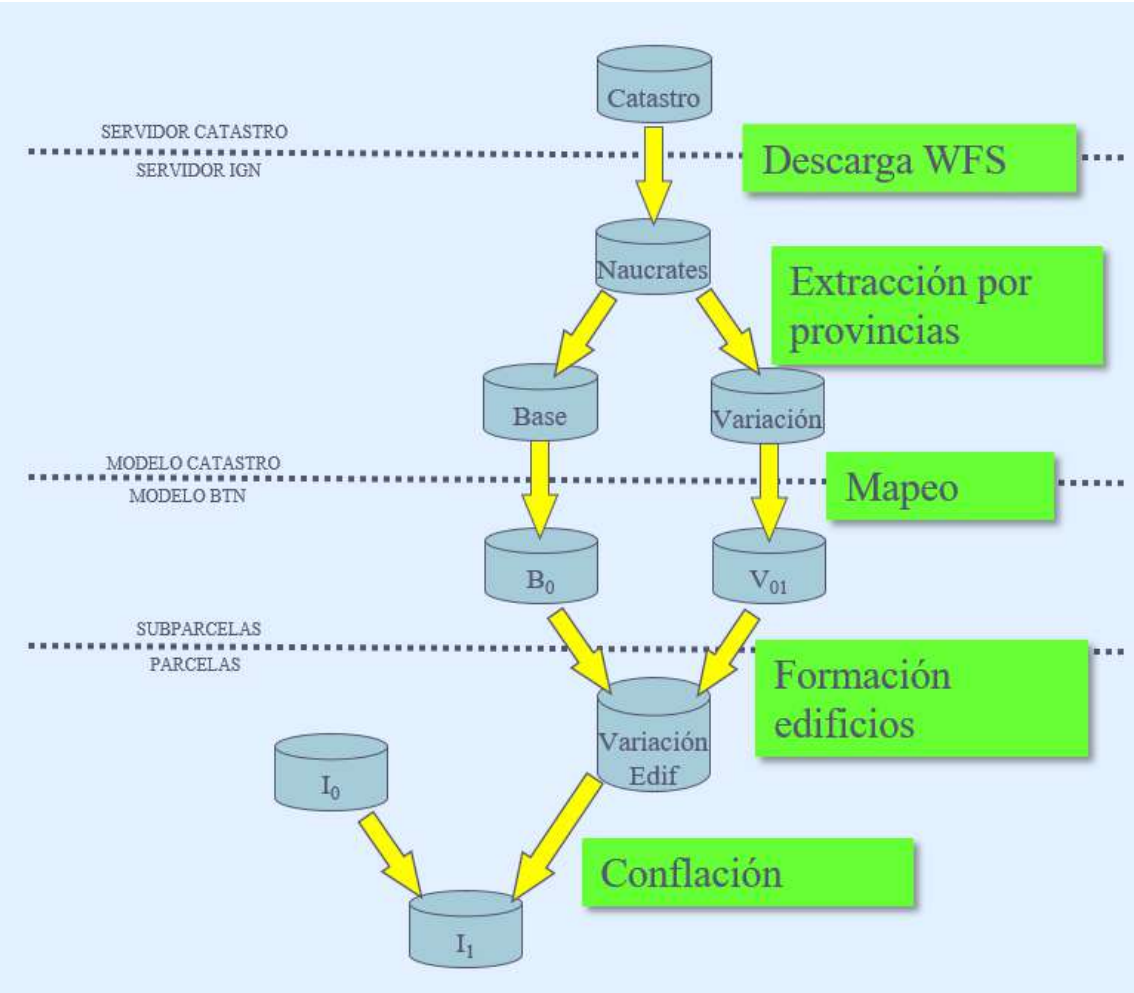


Ilustración 3-26: Esquema del procedimiento de actualización de edificios de BTN, desde la consulta WFS a la conflación

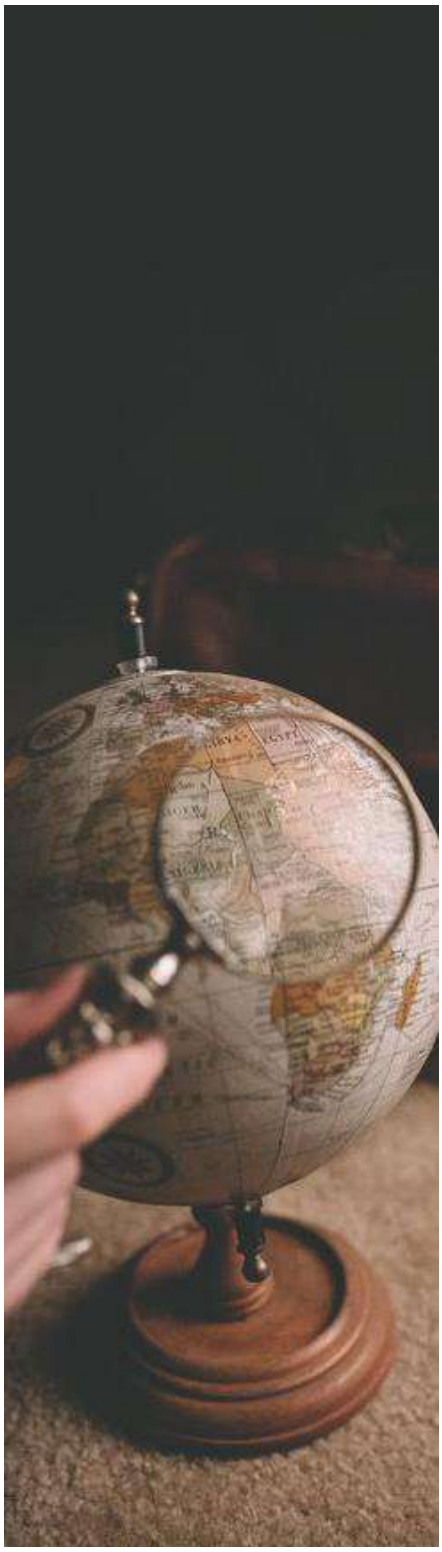
El componente más elemental del sistema de extracción de datos es la petición al servicio WFS. Debido a las restricciones de descarga antes mencionadas, en general es necesario enviar varias peticiones hasta reunir el conjunto de altas y bajas registradas en un determinado municipio durante en el intervalo de tiempo que media entre la última consulta realizada al servicio y la fecha actual. Por tanto, cada petición debe incluir, entre otros argumentos, un criterio de selección por fecha de alta o de baja limitado al periodo considerado.

Sobre la base de estas peticiones elementales y aplicando un esquema de procesamiento por lotes, se obtienen todos los datos comprendidos en el ámbito geográfico de interés, que puede ser una provincia, un conjunto de varias provincias o –más generalmente- la totalidad del territorio nacional. La ejecución completa del conjunto de consultas para todo el país dura un día aproximadamente. En este espacio de tiempo están incluidas las pausas que se han programado entre petición y petición para evitar la saturación del servidor.

Los datos recibidos se escriben en una base de datos PostGIS, dentro de una tabla que reproduce el conjunto de atributos presentes en el servicio WFS. En esta base de datos también queda anotada la fecha de la última petición realizada a cada municipio, que marcará la fecha inicial del siguiente periodo de consulta.

Una segunda modalidad de obtención de datos WFS está orientada a descubrir el conjunto de entidades catastrales que se encontraban vigentes en una fecha determinada. En particular, este tipo de consulta ha servido para marcar el punto de partida de la base de datos local, cuya fecha se ha fijado arbitrariamente en el 1 de septiembre de 2016. Se ha extraído del servicio WFS e incorporado a la base de datos PostGIS una copia de todas las entidades existentes en dicha fecha (50 millones de registros). Sobre esa base y teniendo en cuenta las altas y bajas ocurridas posteriormente (también registradas en la base de datos local), es posible obtener mediante consulta un conjunto de datos representativo de la realidad existente en cualquier momento comprendido entre el punto de partida y la fecha actual.

Una de las prácticas del curso está dedicada a la extracción selectiva de información de un servicio WFS. En ella se detalla el conjunto de reglas que rigen la comunicación con el servicio WFS y se ponen los cimientos de un sistema de extracción de información como el que hemos descrito.



CAPÍTULO 4: CALIDAD DE LA INFORMACIÓN GEOGRÁFICA

Norma ISO 19157 y Estrategias de Gestión de la Calidad en los datos geográficos

Descripción breve

Exposición del contenido de la Norma ISO 19157 desde el punto de vista práctico. Aplicación de la norma en el control de procesos y de producto. Estrategias y recomendaciones para el control de calidad de datos geográficos. Gestión del Calidad: calidad de procesos y producto final. Principales normas relativas a la calidad de la Información Geográfica y estrategias para su aplicación.

Índice de contenidos

4.1. Consideraciones preliminares 1

4.1.1. Conceptos básicos 1

4.1.2. Etapas 1

4.1.3. Control de procesos vs control de producto final 2

4.2. Norma ISO 19157 3

4.2.1. La norma ISO 19157: objetivo 3

4.2.2. Estructura y aspectos principales de la norma 3

4.2.3. Componentes de la calidad de datos 5

4.2.3.1. Unidad de calidad de datos 5

4.2.3.2. Elementos de calidad de datos 6

Compleción: 6

Consistencia lógica 6

Exactitud posicional: 6

Exactitud temática: 6

Calidad temporal: 6

Elemento de usabilidad: 7

4.2.3.3. Descriptores de los elementos de la calidad de los datos 7

a) Medidas 7

b) Métodos de Evaluación 7

c) Resultado 8

4.2.3.4. Elementos de metacalidad 8

4.2.4. Guía práctica para su aplicación 9

4.2.4.1. Fase preliminar: Especificaciones de producto 9

4.2.4.2. Fase 1: Definir la unidad de calidad de los datos. 9

a) Ámbito 9

b) Elementos de calidad de los datos 10

4.2.4.3. Fase 2: Definir la metodología de evaluación 11

4.2.4.4. Fase 3: Definir el nivel de conformidad 13

4.2.4.5. Fase 4: Resultado de la evaluación 13

4.2.4.6. Fase 5: Informe de resultados 14

4.3. Automatización de los controles de calidad 14

4.3.1. Importancia de las especificaciones del producto en el control de calidad 14
4.3.2. Definición de variables sujetas a control 14
4.3.3. Posibilidad de automatizar el control de calidad 15
4.3.3.1. Compleción, exactitud posicional, exactitud temática, calidad temporal 15
4.3.3.2. Consistencia lógica 17
4.3.4. Herramientas para la automatización 17
4.3.4.1. Software específico GIS (Geographic Information Systems) 18
4.3.4.2. Software específico ETL (Extract Transform Load) 18
4.3.4.3. Desarrollo de scripts propios 18
4.3.5. Control de procesos 18
4.3.5.1. Por qué es necesario 18
4.3.5.2. Elementos de calidad durante la etapa de producción 19
a) Compleción 19
b) Consistencia lógica 21
c) Exactitud posicional 26
d) Exactitud temática 28
4.3.6. Control de calidad del producto final 29
4.3.6.1. Metodología para la creación de informes de errores 30
4.3.6.2. Validación de datos 31
a) Compleción 31
b) Consistencia lógica 33
c) Exactitud posicional 37
d) Calidad temporal y exactitud temática 39
4.4. Aspectos generales de calidad posicional 39
4.5. Muestreos 45
4.5.1. Necesidad de muestreo y marco normativo 45
4.5.2. Definiciones y estrategia de muestreo 45
4.5.3. Muestreos probabilísticos 47
4.5.3.1. Estimación de parámetros e intervalo de confianza 47
4.5.3.2. Tamaño de muestra a partir de error de estimación 49
4.5.3.3. Introducción a las normas ISO 2859 y 3951 50

Índice de ilustraciones

Ilustración 4-1: Rueda de Deming o espiral de mejora continua 1
Ilustración 4-2: Estructura del contenido de la ISO 19157 4
Ilustración 4-3: Síntesis de los elementos de calidad de los datos y sus descriptores 8
Ilustración 4-4: Resumen del listado de los elementos de calidad aplicados en los ámbitos de datos. Extracto del índice del anexo de calidad de la IGR-RT 11
Ilustración 4-5: Extracto del modelo de datos de IGR-R en relación a los datos de portales y puntos kilométricos 12
Ilustración 4-6: Datos vectoriales de Santo Domingo procedentes de Openstreetmap 16
Ilustración 4-7: Imagen de satélite de la zona de Santo Domingo 16
Ilustración 4-8: Cuadro de diálogo para creación de cuadrícula de revisión en QGIS 20
Ilustración 4-9: Cuadrícula formada por elementos de 10 km*10km de lado para revisión de la red viaria de alta capacidad en Península Ibérica e Islas Baleares 20
Ilustración 4-10: Cuadro de diálogo de ArcGIS Pro para creación de nueva capa 22
Ilustración 4-11: Tabla de definición de dominios en ArcGIS Pro 23
Ilustración 4-12: Extracto de las especificaciones de la Información Geográfica de Referencia (tramos de red ferroviaria) 24
Ilustración 4-13: Tabla de definición de subtipos en ArcGIS Pro 24
Ilustración 4-14:Ejemplo de uso de clave externa 25
Ilustración 4-15: Herramienta de edición topológica de ArcGIS Pro. Fuente: https://pro.arcgis.com/es/pro-app/latest/help/editing/move-topology-edges-and-nodes.htm 26
Ilustración 4-16: Herramienta de edición topológica de QGIS 26
Ilustración 4-17: Comparación entre dos escalas de visualización (a la izquierda, 5,000; a la derecha, 1,000) 26
Ilustración 4-18: Ejemplos de reglas topológicas que establecen relaciones entre elementos de distintas capas (fuente: http://webhelp.esri.com/arcgisdesktop/9.2/pdf/topology_rules_poster.pdf) 27
Ilustración 4-19: Ejemplo de simbolización de red viaria con distintos estilos en función de si se trata de un camino, una vía urbana, una carretera convencional u una autovía 29
Ilustración 4-20: Extracto de la tabla en base a la cual se identifican los errores de la red viaria de la IGR-RT. 30
Ilustración 4-21: Comparación entre dos conjuntos de datos lineales (IGR-RT y MOCI) 32
Ilustración 4-22: Camino digitalizado con número reducido de vértices. Véase que los vértices se encuentran bien posicionados, pero hay partes de la línea que se alejan notablemente de la verdad terreno. 40
Ilustración 4-23: Comparativa de camino a evaluar (línea roja) con búferes a 3 y 8 metros de la línea considerada como verdad terreno. Se puede determinar qué porcentaje de la línea se encuentra dentro de cada búfer. 41
Ilustración 4-24: Punto de confluencia de los ejes de dos ríos (líneas azules). En amarillo se marca una zona de incertidumbre en la que se podría haber hecho la confluencia de los ejes. 41
Ilustración 4-25: Componentes de la estrategia de muestreo según la ISO 19157:2013. 46

Índice de tablas

Tabla 4-1: Documentación de la medida y el método aplicado para evaluar la consistencia de dominio de los datos de portales y puntos kilométricos 13
Tabla 4-2: Ejemplo de errores de consistencia conceptual que afectan a varios atributos de una misma tabla 36
Tabla 4-3: Ejemplo de errores de consistencia conceptual que afectan a varios atributos de distintas tablas 37
Tabla 4-4: Ejemplo de errores de consistencia topológica 37

4.1. Consideraciones preliminares

4.1.1. Conceptos básicos

Este capítulo aborda la problemática relativa a la gestión de la calidad de la Información Geográfica. Para ello, es necesario definir en primer lugar los siguientes conceptos, a los que se hará alusión durante el resto del presente capítulo:

- Calidad: *grado en que un conjunto de datos cumple una serie de requisitos* que deberemos haber definido previamente, y que deberán ser *adecuados al uso previsto*.
- Inspección de Calidad: conjunto de procesos mediante los cuales se evalúa si el resultado de los trabajos realizados cumple los requisitos definidos inicialmente.
- Gestión de Calidad: conjunto de acciones encaminadas a lograr satisfacer los requisitos establecidos y a lograr una mejora continua de la eficacia y eficiencia de la organización.

De acuerdo con lo anterior, la aplicación de metodologías de Gestión de Calidad deberá estructurarse en una serie de etapas, en muchos casos iterativas, que conducirán no solo a obtener un producto de la calidad requerida, sino a incrementar la eficiencia y eficacia de los procesos de producción. Estas etapas pueden resumirse gráficamente mediante la *Rueda de Deming* o espiral de mejora continua, que se muestra a continuación:

Ilustración 4-1: Rueda de Deming o espiral de mejora continua

4.1.2. Etapas

Las 4 etapas que se identifican en la Rueda de Deming tienen su reflejo en uno o varios de los apartados de la presente Guía. Concretamente, en este capítulo se abordarán:

- Etapa 1: planificación.
    - Subetapa 1.1: definición de los requisitos que debe cumplir el producto. En esta fase, debe atenderse a los usos para los que está previsto necesidades de información que debe cubrir, normativa o regulación legal en base a la cual

debe desarrollarse o utilizarse el producto final, características de los productos derivados o que sería deseable obtener a partir de nuestro producto, etc. En la medida de lo posible, los requisitos deberán definirse haciendo referencia a los elementos y subelementos de calidad que se describen en la Norma ISO 19157 y su aplicabilidad a los distintos objetos geográficos. Este aspecto se aborda en el apartado 4.2 del presente capítulo.

  - Subetapa 1.2: definición de las tareas necesarias para obtener un producto que cubra las necesidades identificadas, y de los recursos humanos y materiales necesarios para llevarlas a cabo (quién las lleva a cabo, quién las revisa y quién las aprueba, y de qué software, hardware y datos disponen o deberían disponer para ello).
  - Subetapa 1.3: definición de las etapas y metodología más adecuada de inspección, así como de los indicadores adecuados, que permitan garantizar que los requisitos se han cumplido. Para ello, deberá tenerse en cuenta si se dispone de datos respecto a los cuales validar nuestro producto, y las características de esos datos (periodicidad de actualización, resolución espacial, formato).

- Etapa 2: producción. Se aborda en el resto de los capítulos de la Guía.
- Etapa 3: comprobación de la calidad del producto. Se aborda en los apartados 4.3, 4.4 y 4.5 del presente capítulo.
- Etapa 4: corrección de los errores detectados y establecimiento de estrategias de mejora.

4.1.3. Control de procesos vs control de producto final

A la hora de planificar las metodologías y actividades necesarias para garantizar la calidad del producto desarrollado, debemos tener en cuenta dos enfoques distintos:

- por un lado, el enfoque de **Control de Procesos** se centra en las actividades que permiten **controlar el proceso de producción mientras este se está llevando a cabo**, con el objetivo de prevenir posibles defectos o errores y maximizar la probabilidad de que el producto final tenga la calidad esperada desde el primer momento.
- por otro lado, el enfoque de **Control de Calidad de Producto Final** incluye las tareas de **comprobación** que se realizan **una vez terminado el proceso de producción**, y cuyo objetivo es evitar que un producto erróneo llegue al usuario final.

Ambos enfoques son complementarios y necesarios, ya que ninguno de ellos sirve, por separado, para asegurar de forma eficiente el cumplimiento de los requisitos especificados. El primero de estos enfoques se tratará en el apartado 4.3.5, y el segundo de ellos se desarrollará en el apartado 4.3.6 del presente capítulo.

4.2. Norma ISO 19157

4.2.1. La norma ISO 19157: objetivo

La Norma ISO 19157:2013 es la norma internacional de la familia ISO 19100 de Información Geográfica que establece los principios para describir la calidad de datos geográficos. Sustituye a las anteriores normas ISO 19113 e ISO 19114 y a la especificación técnica ISO/TS 19138.

Esta norma considera la evaluación de la calidad de los datos tanto desde la perspectiva del productor, o adecuación de los datos a las Especificaciones del producto, como la del usuario, o adecuación de los datos a sus requisitos.

Actualmente vivimos en un mundo global con sobreabundancia de datos y en ocasiones es complejo determinar cuál es el conjunto más idóneo según el caso de aplicación que se trate. En este sentido, cabe destacar que el propósito que persigue esta norma con la descripción de la calidad es el de facilitar dicha comparación y la consecuente selección del conjunto de datos más adecuado según los requisitos del caso de uso.

El objetivo de esta norma es proveer:

- unos principios para la **descripción de la calidad de los datos geográficos**, y conceptos, para el manejo de esta información.
- una manera **consistente y estandarizada de determinar y reportar la información de la calidad** de los conjuntos de datos.
- unas pautas o descripciones de los **procedimientos generales para la evaluación de la calidad** de los datos geográficos. Facilita además un conjunto de **medidas de la calidad** aplicables en dichos procedimientos, y reportables en los informes de calidad.

Por último, es importante resaltar que esta norma internacional **no pretende definir niveles mínimos de aceptación** para la calidad de datos geográficos, pues esta cuestión deberá ser determinada finalmente bien por el productor de los datos, según las Especificaciones de su producto, o bien por el usuario, según sus requisitos.

4.2.2. Estructura y aspectos principales de la norma

Como todas las normas de la familia ISO 19100, el contenido inicial de la ISO 19157 se estructura en las clásicas **secciones introductorias**. Así, comienza con una *introducción*, que contiene la descripción conceptual del objetivo de la norma, y continua con las secciones de *alcance*, donde concreta el marco en el que pretende alcanzar su objetivo y el modo, *conformidad*, indicando las pruebas (anexo A) que debe superar un producto parar ser considerado conforme con la norma, *referencias normativas* o conjunto de normas a las que se hace alusión en ésta, *términos y definiciones* más relevantes que se mencionan en el documento cuyo significado el lector debe conocer sin ambigüedad y por último las *abreviaturas* que emplean en todo el documento.

A partir esta sección el contenido de la norma es de mayor densidad por lo que pasa describirse de forma más detalla según sus apartados.

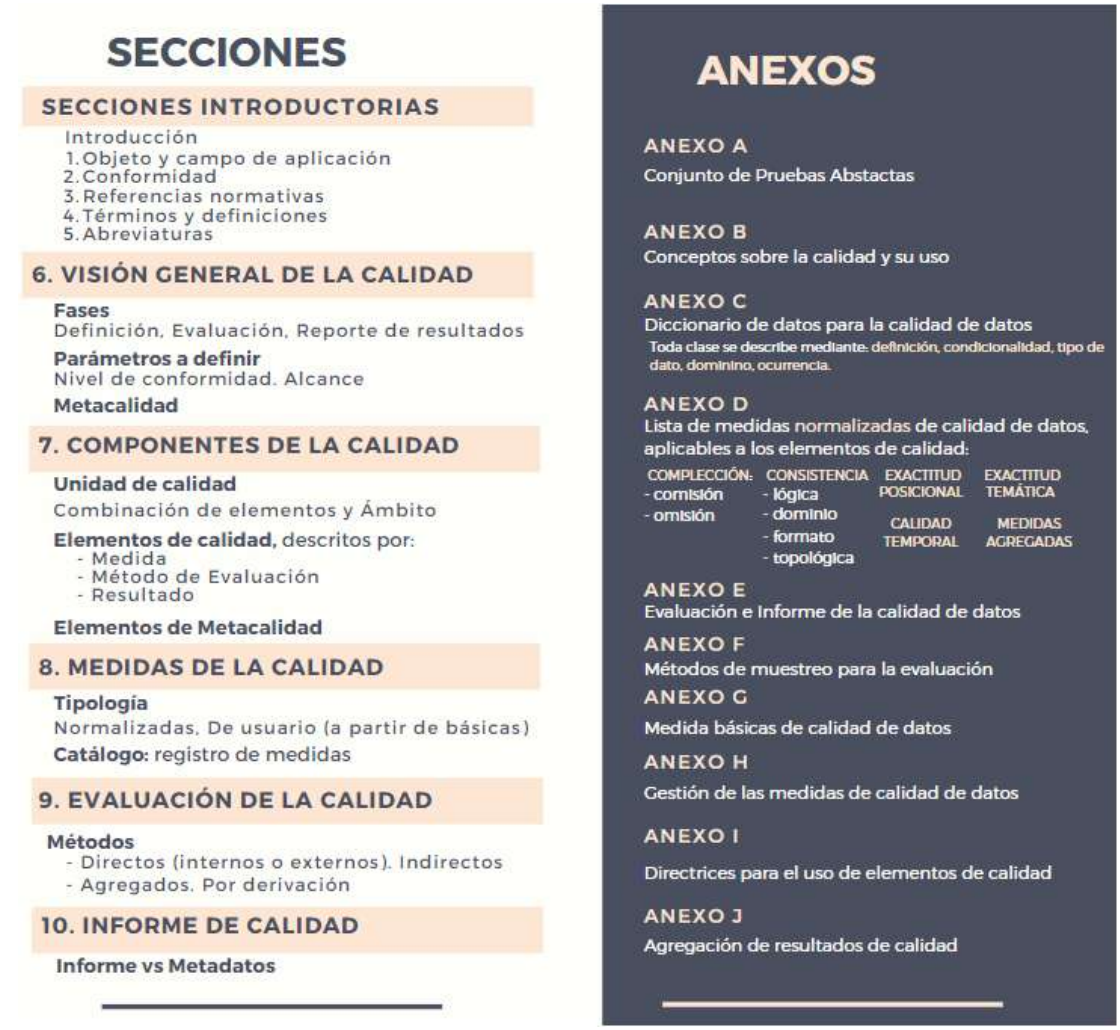


Ilustración 4-2: Estructura del contenido de la ISO 19157

La norma, en su **apartado 6** aporta una **visión general** sobre lo que supone trabajar con la calidad de los datos geográficos, identificando las fases de definición, evaluación y reporte de los resultados.

Así, indica que la calidad de los datos debe ser descrita usando los componentes de calidad, los cuales a su vez deben ser definidos por la organización a partir de los conceptos de calidad de datos vinculados a la información geográfica, y que en la norma se recopilan en el Anexo B. Los aspectos de calidad deben ser especificados en esquemas de aplicación.

Además, es necesario definir los niveles de conformidad de la calidad de los datos (en las Especificaciones de producto o en los requisitos de usuario), de modo que éstos se puedan comparar con los resultados cuantitativos con los que se expresan los elementos de calidad y sus descriptores, y determinar así si el conjunto de datos evaluado es conforme.

Una vez aplicados los procesos de evaluación de la calidad, incluyendo tanto las medidas que permiten obtener valores cuantitativos de los componentes de la calidad como el análisis de conformidad, es necesario reportar los resultados, informando así sobre la calidad de los datos.

No obstante, si el conjunto de datos no ha sido creado conforme a unas Especificaciones de producto que contengan valores de referencia cuantitativos, al menos, aunque sea de forma

descriptiva, el productor puede facilitar información de la calidad a través de los elementos de metadatos propósito, uso y linaje.

Por último, señalar que informar de la calidad de un conjunto de datos es reportar igualmente datos de los que se puede evaluar a su vez la calidad. Esto se denomina **metacalidad.**

En la **sección 7, componentes de la calidad de datos**, se describen y clasifican los componentes que han de emplearse para describir la calidad de los datos: unidad de calidad, elementos de calidad y sus descriptores (medida, método de valuación y resultado) y elementos de metacalidad. En el anexo C, donde se explica el contenido que debe de tener un diccionario de datos, se detallan sus características y atributos.

En la **sección 8, medidas de la calidad de datos**, se detalla la naturaleza de las medidas aplicables y el modo en el que deben describirse, recomendando que siempre que sea viable se empleen las medidas normalizadas (descritas en el anexo D), o si se crean nuevas que empleen las medidas básicas (descritas en el anexo G) y que se registren en un catálogo (conforme al anexo H).

Estos dos apartados son los más relevante para poder implementar correctamente la norma, motivo por el que su contenido se detalla en el siguiente apartado de este documento.

En la **sección 9, evaluación de la calidad de datos,** la norma indica las fases que implica el proceso de evaluación, estructura los métodos aplicables y plantea la posibilidad de generar una evaluación final por *agregación de resultados* de evaluaciones basadas en diferentes elementos de calidad o diferentes ámbitos (descrita con mayor detalle en el anexo J así como la forma de reportarla en el anexo E), o por *derivación de resultados* existentes (ej. la valoración de conformidad es un ejemplo de derivación pues se obtiene tras comparar los umbrales de aceptación con los valores obtenidos al evaluar los componentes de calidad.

Por último, en la **sección 10, informe de la calidad de los datos**, la norma aporta cómo se debe reportar la información de la calidad de los datos. Una de las ideas fundamentales que transmite es que este informe, que puede tener una estructura libre, es distinto y complementario a los metadatos, de forma que estos nunca deberán ser sustituidos por el hecho de elaborar dicho informe. En particular, la norma específica cómo debe ser la relación entre el informe y los metadatos en los casos de evaluaciones por agregación y por derivación. En la agregación, los metadatos deben proporcionar al menos una referencia a los resultados de calidad originales, y pueden incluir información sobre la medida y el método de agregación empleado. En la opción de derivación señala que, además de los metadatos, debería generarse un informe donde facilitar los resultados de calidad originales al que hacer referencia desde los propios metadatos.

En cuanto a los anexos, ya se ha ido mencionando las secciones de la norma desde las que se les hace las principales referencias y en la figura anterior se listan todos ellos.

4.2.3. Componentes de la calidad de datos

4.2.3.1. Unidad de calidad de datos

A la hora de evaluar la calidad de los datos hay que concretar qué datos se van a evaluar y cómo o mediante qué elementos de calidad. Este primer parámetro es lo que se denomina la **unidad de calidad de datos** y se define como la combinación de los elementos de calidad que se van a emplear para la descripción de un ámbito de datos, entendiéndose por ámbito la extensión espacial (ej. una provincia), temporal (ej. un periodo limitado entre dos años) o el conjunto de

características comunes que identifican a los datos a evaluar (ej. el subconjunto de datos de carreteras perteneciente a toda la red de transportes).

4.2.3.2. Elementos de calidad de datos

Los elementos de la calidad de datos, muchas veces denominados directamente como las componentes de la calidad, describen un cierto aspecto de la calidad a partir del resultado obtenido tras aplicar el método que permite evaluar la medida considerada para la cuantificación dicho aspecto de la calidad, es decir, de sus descriptores. Los elementos se clasifican según las categorías que se describen a continuación.

Compleción: Es la presencia o ausencia de objetos geográficos, sus atributos y relaciones y se concreta en comisión, en el caso de datos excedentes que existen en el conjunto de datos o en omisión, cuando se trata de datos ausentes en el conjunto de datos.

Consistencia lógica: es el grado de adherencia a las reglas de la estructura de los datos, de los atributos y de las relaciones. Contempla a su vez cuatro tipologías o elementos de calidad:

- Consistencia conceptual: adherencia a las reglas del modelo conceptual, que describe la estructura de datos (nombres de tablas, atributos, relaciones, etc.).
- Consistencia de dominio: adherencia de los valores de dominio.
- Consistencia de formato: grado en el que los valores se almacenan conforme a la estructura física del conjunto de datos.
- Consistencia topológica: corrección de las características topológicas codificadas explícitamente.

Exactitud posicional: exactitud de la posición de los objetos geográficos en un determinado sistema de referencia. Generalmente se mide por comparación de los datos respecto a una verdad terreno o fuente de referencia considerada como tal. Puede ser:

- Exactitud absoluta o externa: proximidad de los valores reportados a los verdaderos (o considerados de referencia).
- Exactitud relativa o interna: proximidad de las posiciones relativas de los objetos en el conjunto de datos a sus respectivas posiciones relativas verdaderas.
- Exactitud posicional de datos en malla: en el caso de datos en estructura de malla regular, se trata de la proximidad de sus valores de posición a los valores verdaderos.

Exactitud temática: es la exactitud o corrección de los valores y clasificaciones de los atributos, los objetos geográficos y sus relaciones. Se compone por tanto de estos tres elementos:

- Corrección de la clasificación: comparación de las clases asignadas a los objetos o sus atributos respecto de los valores que se consideren de referencia.
- Corrección de atributos no cuantitativos: para evaluar si el valor es correcto o no.
- Exactitud de los atributos cuantitativos: proximidad de su valor al de valor de referencia.

Calidad temporal: calidad relativa a los atributos y a las relaciones temporales de los objetos geográficos, en un punto definido en el tiempo o en un periodo. Se compone de tres elementos:

- Exactitud de una medida de tiempo: proximidad de las medidas de tiempo realizadas a los valores de referencia.

- Consistencia temporal: corrección del orden de los eventos.
- Validez temporal o validez de los datos con respecto al tiempo.

Elemento de usabilidad: elemento que se usa para describir la información específica de la calidad sobre la idoneidad de un conjunto de datos para una aplicación particular o su conformidad con los requisitos, que debe ser empleado cuando dicha evaluación no es posible realizarla a partir de los elementos de calidad anteriormente mencionados.

4.2.3.3. Descriptores de los elementos de la calidad de los datos

Tal y como se ha mencionado anteriormente, la forma de describir a los elementos de la calidad de datos es a través de una medida, un método de evaluación y el resultado.

a) *Medidas*

Este descriptor define **qué** se va a cuantificar a la hora de evaluar un elemento de calidad. Las medidas pueden ser:

- Normalizadas, aquellas que permiten que las evaluaciones y los informes de calidad sean comparables. En el anexo D de la norma, se describen por elemento de calidad al que son aplicables, siendo el productor quien deba seleccionar las idóneas en función del tipo de datos y de su uso previsto.
- Creadas por el usuario, cuando por la naturaleza de los datos y su calidad, la lista anterior es insuficiente y es necesario definir otras adicionales, siempre a partir de las medidas básicas de la calidad de datos (descritas en el anexo G) y aportando la descripción completa de las mismas a través de los atributos indicados en la norma.

Las medidas se describen a partir de un conjunto numeroso de atributos por lo que es recomendable que su descripción completa se almacene en un **catálogo** de forma que, a la hora de emplearlas sean mencionadas únicamente con su identificador o su nombre y se remita a dicho catálogo para obtener la descripción completa de las mismas.

b) *Métodos de Evaluación*

Este descriptor indica los procedimientos y métodos que se aplican para evaluar la medida con la que se obtendrá un resultado, es decir, define **cómo** se va a evaluar. Para cada medida se debe proporcionar un método de evaluación.

Los métodos de evaluación pueden ser:

- Directos o métodos basados en la inspección de ítems que determinan la calidad mediante la comparación de los datos con información de referencia. Cuando la información de referencia que utilizan es del propio conjunto de datos se denominan directos internos**.** Si precisan de datos de referencia externos al conjunto de datos a evaluar, entonces se identifican como métodos directos externos. En ambos casos se puede aplicar realizando una inspección completa (se evalúa la totalidad de los datos) o un muestreo (se aplican a subconjuntos de datos).
- Indirectos, infieren o estiman la calidad a partir de información de los propios datos, como por ejemplo el linaje. Se basan en conocimiento o experiencia externa sobre el producto y pueden ser subjetivo. Cuando se informa de la calidad de forma indirecta es importante complementarlo con la descripción de cómo se ha determinado.

c) *Resultado*

Es obligatorio proporcionar al menos un resultado para cada elemento de calidad. Según la naturaleza de los datos (tienen diferentes procedencias, métodos de captura, etc.) puede ser recomendable ofrecer varios resultados de la misma medida aplicada en ámbitos individuales o subconjuntos del ámbito de estudio total.

Pueden ser:

- Cuantitativos, hay que expresarlos en conformidad con la descripción de la medida y deberían ofrecerse junto con la unidad de medida que se trate.
- De conformidad, derivado al comparar el resultado de la medida con el nivel de conformidad previamente especificado. Se puede proporcionar un resultado de conformidad por cada medida.
- Descriptivo o declaración textual, de carácter subjetivo, que se proporciona cuando no es posible realizar una evaluación cuantitativa.
- De cobertura, que resulta de una evaluación de la calidad y está organizado como una cobertura. Su descripción se facilita en la modificación 1 de la norma (ISO 19157:2013/Amd 1:2018): *describir la calidad de los datos utilizando coberturas.*

Ilustración 4-3: Síntesis de los elementos de calidad de los datos y sus descriptores

4.2.3.4. Elementos de metacalidad

Estos elementos son declaraciones cuantitativas y cualitativas sobre la evaluación de la calidad realizada y su resultado. Los elementos que se pueden emplear para su descripción son:

- Confianza: fiabilidad del resultado de calidad obtenido.

- Representatividad: grado en que la muestra empleada en la evaluación ha producido un resultado representativo de todos los datos del ámbito.
- Homogeneidad: uniformidad de los resultados obtenidos al evaluar varios subconjuntos de datos del ámbito de estudio. Es recomendable emplear este elemento cuando distintas personas realizan una misma tarea en distintas zonas, como por ejemplo cuando los datos han sido capturados por diferentes operadores.

Al igual que ocurre con los elementos de calidad, cada uno de los componentes de la metacalidad debe ser descrito por una medida, un método de evaluación y un resultado, y, adicionalmente, también puede usarse el descriptor "elemento de calidad relacionado", con el que indicar el elemento de calidad sobre el que se aplica el elemento de metacalidad.

4.2.4. Guía práctica para su aplicación

Este apartado facilita una serie de recomendaciones a seguir a la hora de implementar las fases que componen este proceso, incorporando referencias a la experiencia del IGN de España en relación al producto de la Información Geográfica de Referencia de Redes de Transportes (IGR-RT): conjunto de datos correspondiente a una red tridimensional del transporte, de cobertura nacional, compuesta por cinco modos: red viaria, por raíl, por vías marítimas, aérea y por cable, junto con sus respectivas conexiones intermodales y las infraestructuras asociadas a cada modo.

4.2.4.1. Fase preliminar: Especificaciones de producto

Para poder realizar el proceso de evaluación de la calidad, el conjunto de datos debería estar claramente descrito en las Especificaciones de producto (ISO 19131) de forma que se pueda entender correctamente el universo de discurso interpretado. En el modelo de datos y el catálogo de objetos geográficos se identifican los tipos de objetos geográficos, sus atributos y relaciones, y se describen en detalle a partir de una denominación, descripción, tipo de valor y dominio de valores. Es importante también conocer las normas de captura pues también serán condicionantes a la hora de estimar la idoneidad de las medidas aplicables en cada elemento de calidad, así como, en caso de existir, los umbrales de aceptación especificados en la sección de calidad de las especificaciones.

Como un ejemplo práctico de este tipo de documentación, se pueden consultar las Especificaciones de la IGR-RT, en cuyo **anexo de calidad** se encuentra la información detallada sobre cómo se a planteado la evaluación de la calidad de este conjunto de datos.

4.2.4.2. Fase 1: Definir la unidad de calidad de los datos.

La unidad de calidad se define a partir del ámbito de datos y de los elementos de calidad que se van a evaluar sobre éste.

a) *Ámbito*

Se debe definir al menos un ámbito por cada elemento de calidad a evaluar y su delimitación tendrá en cuenta el hecho de que la calidad sea homogénea dentro de él. Si no lo es, se pueden definir distintos ámbitos y acotar en ellos los resultados de la evaluación. Si no es posible concretar el ámbito, entonces éste será la totalidad del conjunto de datos.

En el caso de referencia de la IGR-RT los ámbitos se definieron según los modos de transporte, y dentro de ellos, por subconjuntos de datos, siguiendo el criterio de conjunto de características comunes. Por ejemplo, dentro del modo de red viaria, se aplican elementos de calidad a los

distintos componentes de la red: trazado lineal (tramos), la información identificativa de las vías (vial), los portales y puntos kilométricos, las infraestructuras. En otros casos puede ser más acertado definirlo por extensión espacial o incluso temporal.

b) *Elementos de calidad de los datos*

A la hora de seleccionar los elementos a valorar en cada ámbito, hay que tener en cuenta que a veces varios elementos permiten valorar una misma característica. Las indicaciones pueden resultar útiles a la hora de seleccionar los elementos.

Sobre los elementos:

- El elemento de usabilidad únicamente debe emplearse cuando con el resto de los elementos no es posible evaluar los requisitos de usuario.
- La consistencia lógica es el único que puede evaluarse completamente sin conocimiento de la verdad terreno, pues valora el modo en que los datos se ajustan a sus relaciones internas y de modelado.
- La compleción y la exactitud posicional y temática se emplean para describir cómo se relaciona el conjunto de datos con el universo del discurso. Sin embargo, la calidad temporal implica una mezcla de elementos ligados a la consistencia lógica pero también otros que precisan el conocimiento de la verdad terreno.
- Se pueden definir nuevos elementos cuando el aspecto de calidad que se quiere analizar no esté completamente recogido en los especificados en la norma.

Sobre las relaciones entre elementos:

- **Omisión_Consistencia lógica**. La ausencia de valores en principio se corresponde con un error de omisión, pero cuando se trata de un valor obligatorio según el modelo, entonces se corresponden con un error de consistencia conceptual.
- **Compleción_Exactitud**: al comparar los datos con la fuente de referencia, si se identifica el objeto homólogo del dato en la fuente (ej.: el trazado de un camino frente a su imagen en una ortofoto), entonces se puede evaluar la exactitud posicional y temática, pero si no se llega ni a establecer esa relación, entonces se estaría evaluando un error de compleción.
- **Consistencia conceptual_consistencia temporal**: si en el esquema conceptual se contemplan las reglas que describen la validez temporal de los datos y la corrección del orden de los eventos, se podría valorar estos aspectos como consistencia conceptual en lugar de temporal.

En el anexo de calidad de la IGR-RT se documentan los elementos de calidad que se evalúan en cada modo de transporte y su aplicación concreta a los diferentes objetos geográficos o subconjuntos de datos. Además de considerar los elementos de calidad de la norma, también se creó el elemento de “fidelidad geométrica” para evaluar el grado de idoneidad de la forma del trazado de las geometrías de las carreteras respecto a la verdad terreno de la ortofotografía. Se trata de un elemento que complementa la información de la exactitud posicional porque evalúa el grado de parecido del trazado al de la realidad. Es especialmente interesante en los carriles de incorporación/salida donde resulta un tanto subjetivo determinar el punto en el que el eje del carril secundario (aceleración/deceleración) debe unirse al eje troncal de la carretera.

- 1 Calidad de los datos
  - 1.1 Comisión
    - 1.1.1 Red Viaria
      - 1.1.1.1 Comisión (Se han representado tramos en RT que no existen en la realidad)
      - 1.1.1.2 Comisión (Se han representado portales en RT que no existen en la realidad)
      - 1.1.1.3 Comisión (Se han representado áreas de infraestructura en RT que no existen en la realidad)
      - 1.1.1.4 Incoherencia portales: elementos de RT sin correspondencia en Catastro en un rango de 100 m. (Portal d...
      - 1.1.1.5 Nombres viales urbanos duplicados (Existe más de un nombre de vial urbano duplicado en un municipio)
      - 1.1.1.6 Geometría duplicada (Existen elementos duplicados geométricamente)
    - 1.1.2 Red Ferroviaria
    - 1.1.3 Red Aérea
    - 1.1.4 Red Marítima
  - 1.2 Omisión
    - 1.2.1 Red Viaria
    - 1.2.2 Red Ferroviaria
    - 1.2.3 Red Aérea
    - 1.2.4 Red Marítima
  - 1.3 Consistencia conceptual
  - 1.4 Consistencia de dominio
  - 1.5 Consistencia de formato
  - 1.6 Consistencia topológica
  - 1.7 Exactitud posicional
  - 1.8 Corrección de la clasificación
  - 1.9 Fidelidad geométrica

Ilustración 4-4: Resumen del listado de los elementos de calidad aplicados en los ámbitos de datos. Extracto del índice del anexo de calidad de la IGR-RT

4.2.4.3. Fase 2: Definir la metodología de evaluación

Esta fase comprende la definición de la medida y la metodología con la que se va a realizar, incluido el ámbito de inspección, cuyo resultado permitirá evaluar al elemento de calidad.

Método de evaluación: se recomienda emplear métodos de evaluación directos, tanto sobre fuentes de referencia interna (el propio conjunto de datos) y externa (ej. ortofotografía), y reportar la información general de la calidad (más propia de los indirectos) a través de los metadatos de uso, propósito y linaje. A la hora de documentar el método de evaluación es importante especificar su tipología, aportar la descripción del procedimiento por el que se realiza la comparación del dato con la fuente de referencia y especificar la fecha o fechas en las que se ha realizado.

Medida: hay que emplear las medidas normalizadas recogidas en la norma y si se crea alguna adicional, hacerlo a partir de las medidas básicas. Las medidas deben permitir cuantificar el aspecto a valorar y deben ser descritas a partir de un identificador, nombre y definición entre otros atributos. Cuando sea preciso, bien por ámbito de los datos o bien por sus múltiples aplicaciones, se pueden aplicar diferentes medidas para evaluar un mismo elemento de calidad.

Ámbito de inspección: hay que especificar si el método se aplica sobre el conjunto completo de datos o si realiza un muestreo.

Ejemplo. Para ilustrar este contenido utilizaremos un ejemplo realizado en la IGR-RT.

DATOS. La red viaria contiene el conjunto de datos puntuales de portales y de puntos kilométricos de todo el territorio, vinculados tanto al trazado geométrico (tramo) como a la identificación alfanumérica (vial), ambos almacenados en la tabla rt_portalpk_p, tal y como se muestra en este extracto de la documentación:

Tabla rt_portalpk_p					
Atributo	**Obligatorio/ Condicional/ OPcional**	**Descripción**	**Tipo de dato**	**Código**	**Tamaño**
[id_vial] FK	O	Identificador del vial	Entero largo		12
[id_tramo] FK	O	Identificador del tramo	Entero largo		12
id_porpk	O	Identificador del portal o PK [El identificador se genera conforme a la regla INEMUN * 10M + secuencial, siendo INEMUN el código INE de provincia y municipio sobre el que se	Entero largo		12
numero	O	Número de portal o de PK	Integer		4
tipo_porpk	O	Especifica si el valor es un portal o un PK	*ValorTipoPortalPK (Int.)*		
			Portal	1	1
			PK	2	
fuente	O	Procedencia de la geometría	*ValorFuente (Int.)*		4
geometry	O	Geometría del portal o PK	GM_Point		

Ilustración 4-5: Extracto del modelo de datos de IGR-R en relación a los datos de portales y puntos kilométricos.

UNIDAD DE CALIDAD. Comisión, omisión, consistencia lógica (conceptual, dominio, formato, exactitud posicional, consistencia temática (corrección de clasificación) aplicada al subconjunto de datos de portales y puntos kilométricos.

METODOLOGÍA DE EVALUACIÓN: en el anexo de calidad se documentan todas las medidas y los métodos de evaluación definidos para cada elemento de calidad comprendidos en la unidad de calidad. A continuación, a título de ejemplo, se incorpora únicamente la descripción de la medida y el método por el que se evalúa la consistencia de dominio, al determinar si existen valores de atributos que no se ajustan a las listas controladas del modelo de datos de IGR-RT para la tabla rt_portalpk_p:

	Elemento de calidad	Consistencia de Dominio
Ámbito	Modo(s) de transporte	Red viaria
	Capa(s)	rt_portalpk_p
Medida[1]	Identificador	16
	Tabla	D.16
	Nombre	Número de items no conformes en dominio - valor
	Definición	Recuento de todos los items del conjunto de datos que no son conformes con su dominio de valores.
	Ejemplo	
Método de evaluación	Descripción	Los valores de los atributos han de ajustarse a las listas controladas o los valores esperados, no se admiten nulos o ceros
	Tipo	001: Interno directo. Inspección completa
	Procedimiento	Se comprueba automáticamente que los valores de los atributos no sean distintos de los valores permitidos por las listas de valores. En los atributos cuyo valor pueda ser un texto o numeración abierta, se verifica que los valores están en el rango esperado. Los atributos que se comprueban son: id_porpk, id_tramo, id_vial (distinto de desconocido y

		no aplicable y mayor que 10010000000); fuente (solo admite como valor el IGN, DGC, o la CA); numero (menor de 10,000 y distinto de no aplicable); tipo_porpk (valores de la lista controlada)
	Fuente	N/A
	Muestreo	N/A
	Nivel de aceptación	0
Observaciones		

Tabla 4-1: Documentación de la medida y el método aplicado para evaluar la consistencia de dominio de los datos de portales y puntos kilométricos

Por último, indicar que en la definición de la metodología, dado que un error en los datos puede impactar en varios elementos, es recomendable establecer el orden en el que se van a evaluar los elementos, siendo recomendable emplear este **orden usual**:

1º. Consistencia de formato, pues si los datos no son legibles no se continuaría con la evaluación.

2º. Resto de componentes de la consistencia lógica, para valorar su proximidad al modelo de datos y resto de reglas internas del producto. Los datos no conformes no continuarían el proceso.

3º. Compleción, para determinar las comisiones y omisiones de los datos respecto a la fuente de referencia o verdad terreno.

4º. Exactitud posicional, temática y temporal, es decir, sobre los datos legibles, conformes a modelo y disponibles en el conjunto de datos, se valora su precisión o grado de proximidad a la fuente de referencia o verdad terreno.

4.2.4.4. Fase 3: Definir el nivel de conformidad

El nivel de conformidad es el valor o conjunto de valores que determinan el grado de adecuación de los datos a los criterios recogidos en las Especificaciones o en los requisitos de usuario. Para cada resultado que se obtenga en la fase anterior se define un nivel de aceptación (límite de errores admitido) bajo el cual se considera que el resultado es admisible y por tanto conforme.

4.2.4.5. Fase 4: Resultado de la evaluación

Se distinguen dos fases:

1º. Cálculo de los **valores cuantitativos,** resultado de aplicar las medidas y métodos de evaluación de los elementos de cada unidad de calidad. Es importante especificar tanto el valor como la unidad.

2º. Determinación de la conformidad: resultado derivado procedente de la comparación de los resultados cuantitativos anteriores con los niveles de aceptación definidos. Si el valor es aceptable (el número de errores están dentro del umbral permitido), se dice que la evaluación ha sido **conforme**; en caso contrario el resultado de la evaluación es **no conforme**. Continuando con el ejemplo anterior, el nivel de aceptación de errores en la consistencia de dominio de los datos de portales y puntos kilométricos es cero, es decir los datos serán conformes cuando no exista ningún error en este sentido.

4.2.4.6. Fase 5: Informe de resultados

Como ya se ha indicado anteriormente, los resultados deben facilitarse en los metadatos y, si es viable, en el informe de calidad correspondiente donde se describa con detalle todos los componentes implicados en la evaluación de la calidad. El informe no tiene una estructura fija y a título de ejemplo la norma en su anexo E facilita ejemplos sobre cómo informar de la evaluación como metadatos.

4.3. Automatización de los controles de calidad

4.3.1. Importancia de las especificaciones del producto en el control de calidad

Parece lógico pensar que no se puede obtener un buen producto sin unas buenas especificaciones. Estas permitirán:

- en primer lugar, desarrollar el producto de modo homogéneo y en base a unos criterios sólidos, y facilitarán el aprendizaje de quienes participen en el proceso de generación del producto.
- en segundo lugar, las especificaciones ayudan a los usuarios a saber qué pueden esperar del producto, y en algunos casos incluso cómo usarlo.
- finalmente, el aspecto más importante en lo que respecta a este capítulo, está en el papel que cumplen las especificaciones como base para la comprobación de si un producto cartográfico cumple los requisitos que se le presuponen.

Es por ello que los controles a los que sometamos nuestro conjunto de datos deben venir refrendados por las especificaciones del producto. Es recomendable emplear un diccionario, tabla o catálogo de elementos asociado al control de calidad, donde se definan claramente, y de forma coherente con las especificaciones:

- los distintos elementos que pueden existir en el conjunto de datos que se está produciendo,
- y las características que estos elementos deben cumplir, especificadas en términos de la ISO 19157.

4.3.2. Definición de variables sujetas a control

El primer problema al que nos enfrentaremos al definir el proceso de control de calidad radica en identificar las variables que tienen un mayor efecto en la calidad del conjunto de datos que estamos desarrollando. Una vez identificadas, deberemos fijar unos indicadores operativos (límites de error permitido, dentro de los cuales dichas variables se consideran aceptables).

Estos indicadores deberían estar relacionados con los elementos de calidad y medidas que se abordan en la ISO 19157 y que hayamos considerado relevantes en nuestras especificaciones de producto. Además, dado que numerosos conjuntos de datos son utilizados por los usuarios para aplicaciones muy específicas, y posiblemente diferentes entre sí, es aconsejable la inclusión de variables o medidas relativas a la calidad funcional requerida. La especificación de calidad funcional da paso a un enfoque más amplio, en el que la calidad del producto no viene dada solo por las características del dato (por ejemplo, exactitud posicional horizontal absoluta de 1 metro

al 99%), sino que se define en términos de una aplicación concreta (si decimos, por ejemplo, que el 95% de los tramos de la red viaria deben ser adecuados para el cálculo de rutas).

Una vez identificadas las variables relevantes y sus indicadores, será necesario definir la metodología de control, esto es: la forma de monitorización y de medición de dicha variable, y las acciones correctivas que habrán de implementarse en caso de que la variable presente valores no aceptables.

Algunos de estos procesos (monitorización, medición, corrección de errores) podrán automatizarse. La automatización de procesos permitirá reducir las tareas manuales y repetitivas, logrando así un aumento de la eficiencia del proceso de revisión y una reducción de los errores humanos que se puedan cometer durante el mismo.

4.3.3. Posibilidad de automatizar el control de calidad

Ahora bien, debemos ser conscientes de que no todas las tareas de revisión son susceptibles de automatizarse: por un lado, no resulta conveniente automatizar tareas en las que el operador humano aporta un valor añadido específico, o que tienen una componente subjetiva, ya que existirá el riesgo de que el resultado obtenido no cumpla los requisitos mínimos establecidos. Por otro lado, no siempre es posible automatizar el proceso, debido a múltiples aspectos de los datos con los que se trabaja, como veremos a continuación.

4.3.3.1. Compleción, exactitud posicional, exactitud temática, calidad temporal

En lo que respecta a estos elementos de calidad, el control de cualquier conjunto de datos deberá realizarse respecto a otro/s conjunto/s de datos de referencia, que consideraremos completos y correctos, en su componente posicional, temática y temporal.

Esos conjuntos de datos de referencia pueden tener diversos formatos. Si el formato es tal que la información de interés está estructurada o es comprensible de forma directa por las máquinas, la revisión de los elementos mencionados será, al menos en parte, automatizable.

Veamos esto mediante un ejemplo: supongamos que deseamos incorporar las carreteras existentes en nuestro ámbito de trabajo al conjunto de datos que estamos produciendo. La información de que disponemos de dichas carreteras puede ser muy variada y estar disponible en distintos formatos (por ejemplo, datos vectoriales procedentes de Openstreetmap, y datos ráster procedentes de imágenes de satélite):

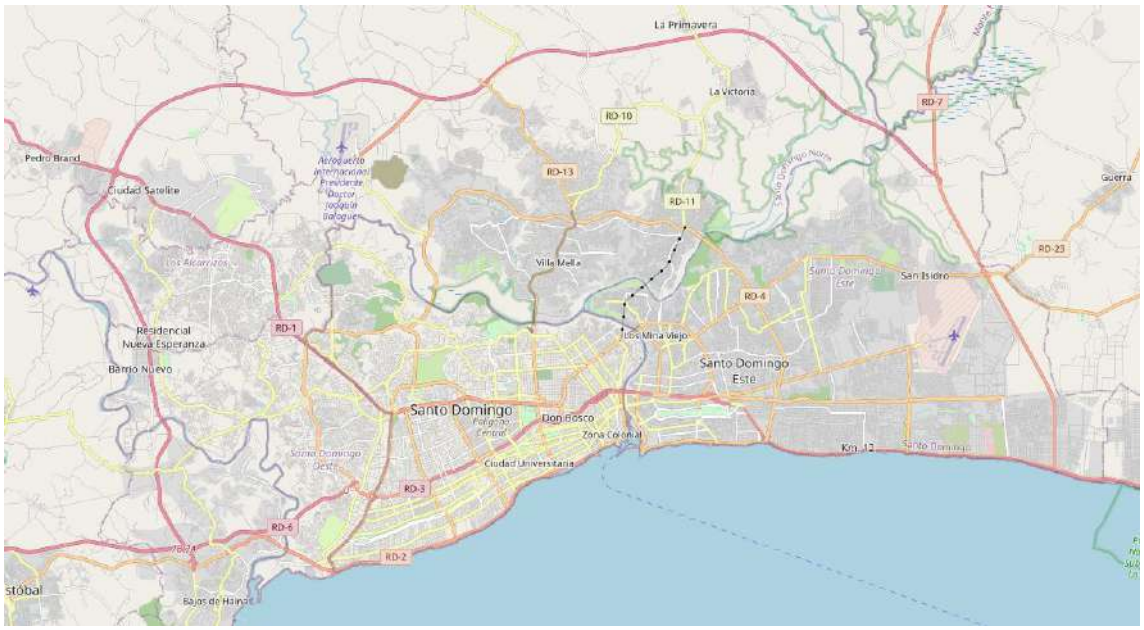


Ilustración 4-6: Datos vectoriales de Santo Domingo procedentes de Openstreetmap

En el caso de la información vectorial de Openstreetmap, las carreteras son elementos geográficos con unas características propias que permiten diferenciarlas de los elementos que no son carreteras, sin tener que llevar a cabo ningún análisis adicional. Tienen una geometría lineal, distinta de los elementos puntuales o poligonales que representan otros fenómenos geográficos, y una serie de atributos o etiquetas que indican que se trata de carreteras, y todo ello es directamente comprensible no solo para el operador humano, sino también para el software GIS o ETL.

Ilustración 4-7: Imagen de satélite de la zona de Santo Domingo

En el caso de la imagen de satélite, vemos que las carreteras son las mismas. El cerebro humano, al ver esta imagen, sabe dónde hay una carretera porque aplica su experiencia y su conocimiento acumulado, y entiende que un conjunto de píxeles de determinado color, agrupados linealmente a lo largo de un recorrido es una carretera. Sin embargo, para una máquina, esta imagen no es más que un conjunto de píxeles, dispuestos en una matriz rectangular, cada uno de ellos de un color determinado. No puede saber qué pixel pertenece a una carretera y cuál no (a menos que antes se lleve a cabo algún proceso para extraer dicha información, proceso que excede del alcance de la presente guía).

Por lo tanto, si deseamos automatizar el proceso de inspección de calidad, necesitamos partir de datos de referencia directamente comprensibles para una máquina (o, en su defecto, llevar a cabo algún proceso intermedio que los haga comprensibles).

4.3.3.2. Consistencia lógica

Este elemento de calidad hace referencia a la coherencia interna de nuestro conjunto de datos, en la medida en que estos **cumplen las reglas lógicas** definidas en el modelo de datos conceptual y las especificaciones técnicas. Como se ha visto en el apartado 4.2.3, engloba los siguientes subelementos, ordenados de menor a mayor complejidad:

- Consistencia de formato: grado en que los datos se almacenan de acuerdo con la estructura física prevista.
- Consistencia de dominio: adherencia de los valores a su dominio.
- Consistencia topológica: corrección de las características topológicas.
- Consistencia conceptual: adherencia a las reglas del modelo de datos que representa la realidad geográfica.

Este elemento es el único cuyo control de calidad puede (y debería) automatizarse completamente. Ahora bien, en ocasiones será necesario que un operador experto revise los resultados, sobre todo en lo que respecta a la consistencia conceptual y a la consistencia topológica. Esto se debe a que, al definir nuestro modelo de datos, estamos simplificando la realidad, eligiendo los aspectos de la misma que nos interesan para poder representarla (si no lo hiciésemos, sería generar cartografía sería una tarea inabarcable). Al simplificar el mundo real, podemos estar obviando determinados aspectos o dejando de recoger determinados casos. Esos casos, por lo tanto, aparecerán clasificados como errores en nuestro control de calidad; sin embargo, lo cierto es que no se tratará de información incorrectamente reflejada en nuestro producto (lo que denominamos falso positivo), sino de información que refleja supuestos no recogidos en nuestro modelo. En tal caso, puede ser conveniente modificar el modelo de datos para recoger dicho supuesto, o bien simplemente tener en cuenta que esos falsos positivos podrán aparecer, y catalogarlos como tales. Hacer una cosa u otra dependerá de los recursos disponibles, las necesidades que debe cubrir nuestro conjunto de datos y la frecuencia con que aparecen dichos supuestos que no se tuvieron en cuenta.

4.3.4. Herramientas para la automatización

La automatización de procesos utiliza cualquier herramienta informática a nuestro alcance. En el ámbito de la información geográfica, las herramientas fundamentales serán el software (GIS y ETL) y el desarrollo de scripts propios.

4.3.4.1. Software específico GIS (Geographic Information Systems)

En la actualidad, los paquetes de software GIS más extendidos permiten configurar el entorno de trabajo de tal manera que desde la fase de producción de los datos se minimicen los errores. Además, cuentan con herramientas tanto de edición como de análisis que nos permitirán abordar las tareas de inspección de los datos y de su corrección, además de automatizarlas parcial o totalmente.

Como ejemplos de este tipo de software, podemos mencionar ArcGIS de Esri en el ámbito del software propietario, y QGIS en el ámbito del software libre. Ambos son de uso generalizado a nivel global.

4.3.4.2. Software específico ETL (Extract Transform Load)

Este tipo de software nos permite leer (Extract) y escribir (Load) datos en una gran variedad de formatos, realizar sobre ellos múltiples operaciones de edición, análisis e integración, y encadenar estas operaciones unas con otras (Transform).

Como ejemplos de este tipo de software, cabe destacar FME en el ámbito del software propietario. En el ámbito de software libre también se han desarrollado aplicaciones de este tipo (Talend Open Studio, Pentaho Kettle, Logstash, etc.), si bien en la mayor parte de los casos no se trata de software totalmente libre, sino que determinadas herramientas o volúmenes de datos solo se permiten haciendo uso de versiones de pago del software.

4.3.4.3. Desarrollo de scripts propios

En el caso de que las herramientas disponibles en el mercado no se ajusten a las necesidades de nuestro trabajo, siempre cabe la posibilidad de desarrollar nuestros propios scripts, ya sea para ejecutarlos dentro del software GIS o ETL, ya sea para ejecutarlos de modo autónomo.

En la actualidad (febrero de 2022) existen numerosas bibliotecas (conjuntos de funciones relacionadas entre sí, que por lo general hacen referencia a un ámbito específico) que permiten trabajar con datos geoespaciales. Además, se han desarrollado bibliotecas tanto en ArcGIS como en QGIS, en los lenguajes de programación Python y C, para que el usuario extienda las funcionalidades del software de acuerdo a sus necesidades.

4.3.5. Control de procesos

4.3.5.1. Por qué es necesario

Definimos un **proceso** como un conjunto lógico de actividades interrelacionadas que toma datos de entrada, les añade valor, y tiene como resultado unos productos determinados para un usuario dado. Las características del proceso determinarán, en gran medida, la calidad de los productos y servicios que ofrece cualquier organismo. Si la cadena de procesos es eficiente y eficaz, el resultado será coherente con las expectativas.

El proceso de mejora continua que se presentaba en el apartado 4.1.1 busca, por tanto, mejorar la calidad de los procesos, lo que conducirá a:

a) Mejorar la calidad del producto final.

b) Tener un conocimiento más profundo de los procesos necesarios para producir nuestros datos, lo que a su vez permitirá:

b.1. Aumentar la eficiencia de los procesos (reducir tiempo y costes).

b.2. Eliminar problemas (causas potenciales de error) en su origen.

b.3. Mejorar la capacidad de adaptación a los cambios que sean necesarios dentro de los procesos establecidos, ya sea por la aparición de nuevos requisitos, ya sea por la necesidad de evolucionar los modelos de datos o las formas de producción por causas técnicas, institucionales, operativas etc.

c) Integrar más fácilmente las funciones de los distintos agentes que tengan algún papel en la producción de los datos.

4.3.5.2. Elementos de calidad durante la etapa de producción

En este apartado se dan algunas recomendaciones relativas al control de calidad de los datos durante el proceso de producción. Se trata de recomendaciones genéricas, y por tanto aplicables en la mayoría de ocasiones. No obstante, en determinados casos podrían no ser aplicables, o ser insuficientes para garantizar la calidad del proceso.

Una vez más, conviene subrayar que los parámetros de cada uno de los procesos relativos al control de calidad deberán ser coherentes con las especificaciones del producto y las variables definidas como relevantes.

a) *Compleción*

Por definición, este elemento de calidad debería revisarse no durante el proceso de producción, sino al final del mismo (apartado 4.3.6), ya que el conjunto de datos sobre el que estamos trabajando únicamente estará completo al finalizar el trabajo. Sin embargo, en este apartado se darán algunos consejos que ayudarán a evitar errores de este tipo.

Barrido exhaustivo del ámbito espacial

La edición de elementos geográficos en un territorio requiere, por lo general, que el operador revise toda el área de trabajo, que en muchos casos puede ocupar cientos e incluso miles de kilómetros. Para asegurarnos de que no queda ninguna zona sin revisar, es aconsejable compartimentar el territorio en subunidades de tamaño menor, y abordar la revisión de dichas unidades siguiendo un orden espacial previamente fijado. Para ello existen múltiples opciones:

- Crear una cuadrícula que abarque todo el ámbito espacial (los paquetes de software GIS tienen una herramienta específica para hacer capas de este tipo), e ir marcando las zonas de la cuadrícula a medida que su revisión se vaya completando.
- Usar otro tipo de división del terreno ya existente (por ejemplo, la organización del territorio en unidades administrativas).
- Definir zonas homogéneas (por ejemplo, áreas urbanas, áreas rústicas, zonas de bosque, etc.)

En cualquier caso, a la hora de definir estas unidades debe tenerse en cuenta:

- Que el tamaño de la unidad elegida debería ser lo suficientemente pequeño como para que un solo operador pueda realizar la revisión en un tiempo aceptable a la escala adecuada.
- Que el tamaño dependerá del tipo de elemento a revisar. Cuanto mayor sea la densidad de elementos en una zona, menor ha de ser el tamaño de cada unidad de revisión.

— Que la definición de las unidades debe poder hacerse de forma rápida. En este sentido, es mucho más sencillo utilizar cuadrículas o divisiones ya existentes, que definir zonas homogéneas de trabajo. No obstante, en determinadas ocasiones esto puede ser necesario.

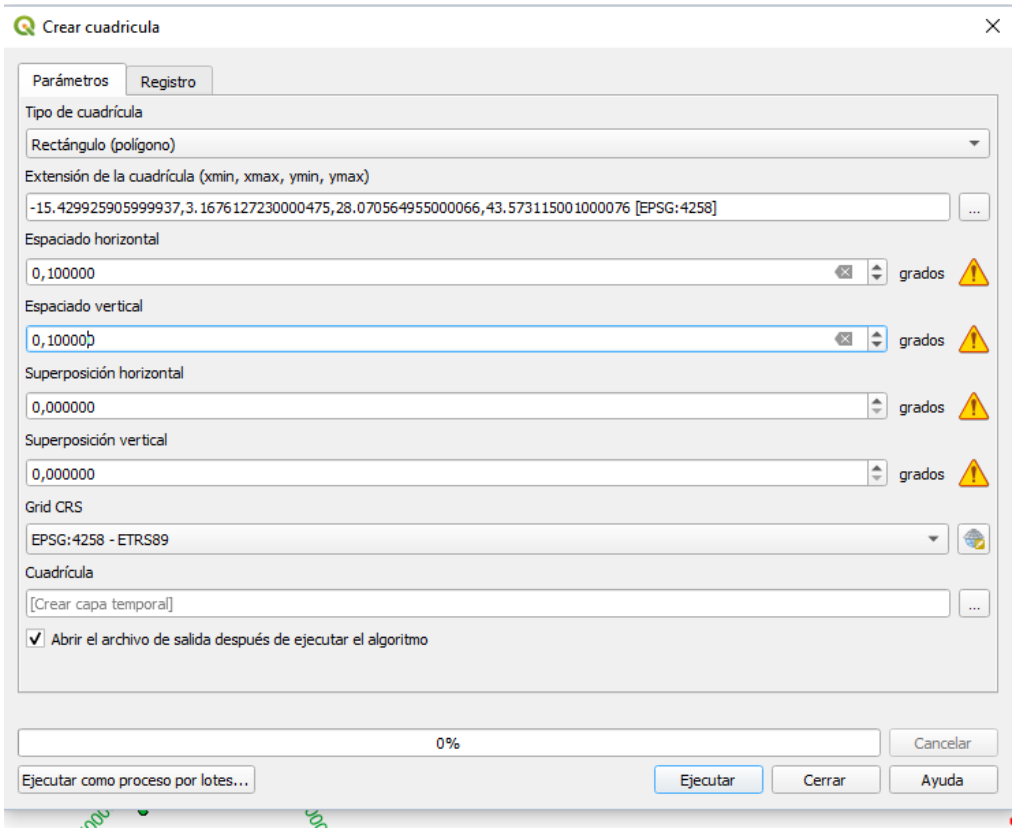


Ilustración 4-8: Cuadro de diálogo para creación de cuadrícula de revisión en QGIS

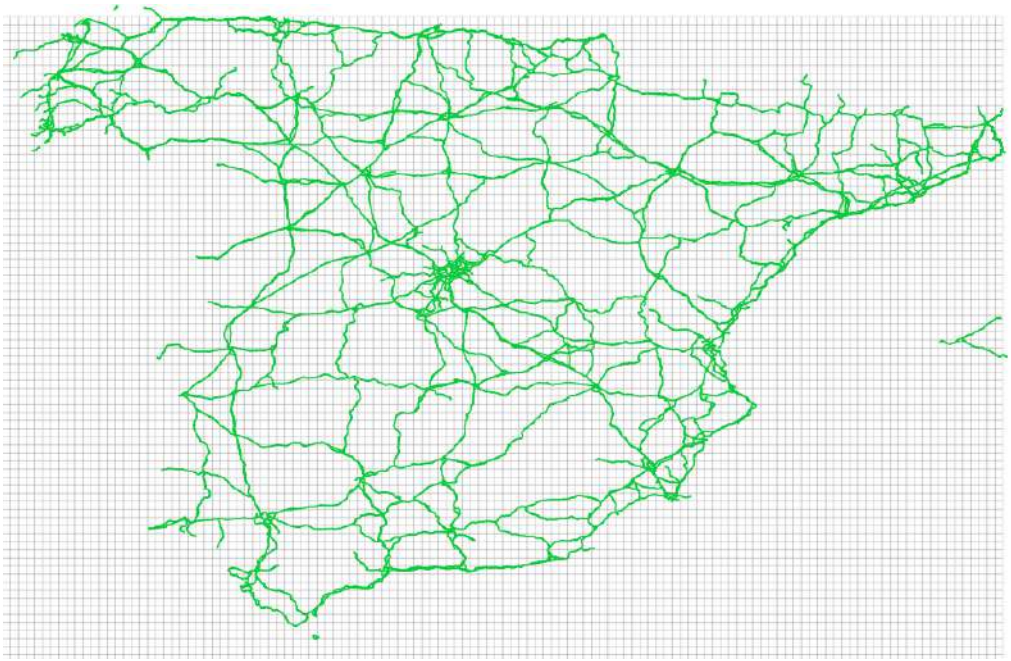


*Ilustración 4-9: Cuadrícula formada por elementos de 10 km*10km de lado para revisión de la red viaria de alta capacidad en Península Ibérica e Islas Baleares*

Listado de elementos

En muchas ocasiones se revisan los datos existentes en el territorio para incluir nuevos datos procedentes de catálogos de terceros (por ejemplo, un listado gasolineras). En este caso, para controlar la inclusión de todos los elementos del catálogo, y para facilitar su posterior revisión, se aconseja incluir como atributo en nuestra tabla el identificador que cada elemento tuviera en origen. Esto hará, además, que, si posteriormente tenemos que incorporar actualizaciones de dichos catálogos, la identificación de modificaciones respecto al catálogo inicial sea mucho más sencilla.

b) *Consistencia lógica*

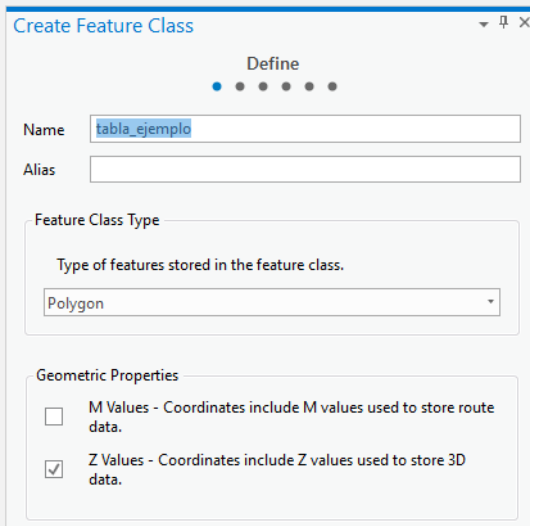
El software GIS nos da diversas herramientas para garantizar la consistencia lógica de los datos que estamos editando.

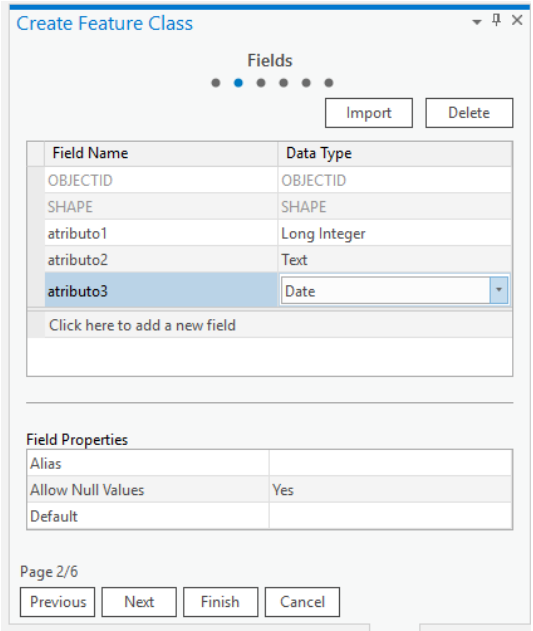
Consistencia de formato

En primer lugar, respecto a la consistencia de formato, deberemos garantizar que los siguientes aspectos son coherentes con lo establecido en las especificaciones:

- Sistema de referencia geográfico (GRS) y proyección empleada.
- Formato de los datos.
- Tipo de geometría: puntual, lineal, superficial o sin geometría; y dentro de las tres primeras, geometrías simples o multigeometrías.
- Nombre de las tablas
- Nombre y tipo de los atributos.
- Codificación de los atributos de tipo texto.

Para evitar errores de este tipo en el producto final, todos estos aspectos deberán definirse en la etapa de creación de las capas del proyecto. Por lo general, todos los paquetes de software GIS guían al usuario en la definición de estos parámetros cada vez que se crea una nueva capa en el proyecto.





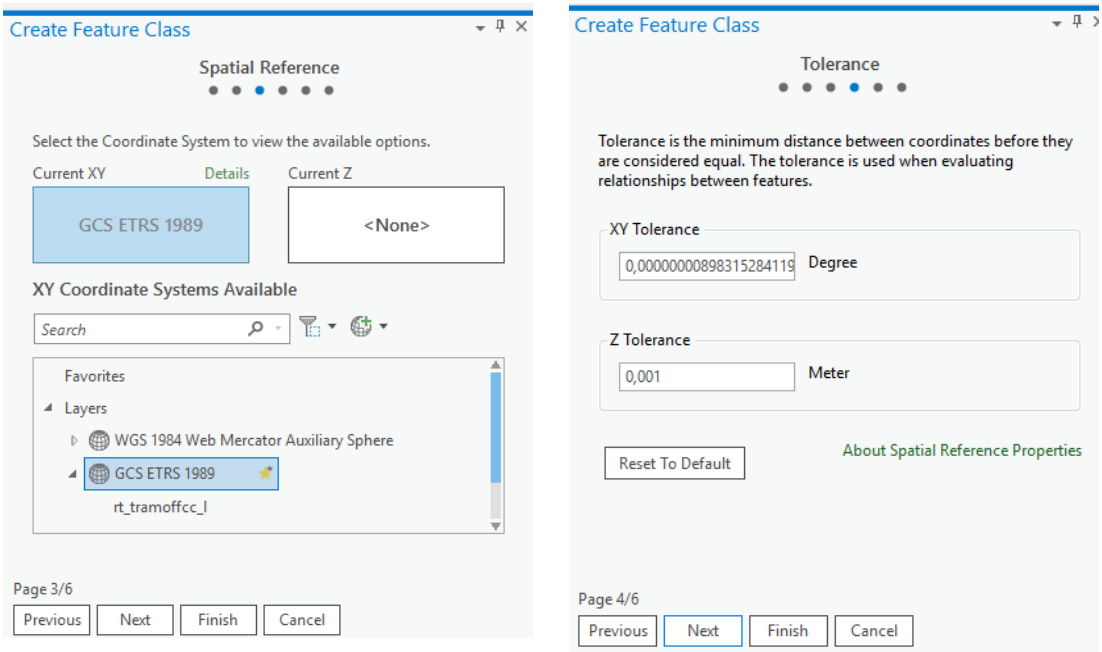


Ilustración 4-10: Cuadro de diálogo de ArcGIS Pro para creación de nueva capa

Llegados a este punto, y aunque no es estrictamente parte de lo que podríamos considerar el control de calidad, consideramos necesario llamar la atención sobre diversos puntos que han de tenerse en cuenta a la hora de definir el formato de los datos en las especificaciones del producto:

- A la hora de elegir el tipo de geometría con la que se va a trabajar, es importante tener en cuenta que no todas las herramientas y funciones existentes en los paquetes GIS y ETL admiten geometrías complejas (superficies poliédricas, curvas compuestas, etc.), ni multigeometrías. Por lo tanto, el tipo de operaciones que haya que realizar sobre los datos que desarrollamos será uno de los aspectos que definan el formato de nuestros datos.
- Del mismo modo, cada formato tiene una serie de características que lo hacen más o menos adecuado a nuestras necesidades. Las limitaciones y ventajas de cada formato deberán ser tenidas en cuenta a la hora de definirlo en las especificaciones (por ejemplo, un archivo shapefile no admite más de 2,000,000 de registros, los nombres de los campos no pueden tener más de 10 caracteres, y determinadas funciones de ArcGIS, como las relativa a la topología, o al análisis de redes, no se pueden ejecutar sobre shapefiles, sino que han de realizarse sobre geodatabases; como contrapartida, el shapefile es un estándar de facto que puede ser leído y modificado por prácticamente cualquier software GIS).
- La codificación de los atributos de tipo texto es uno de los aspectos más problemáticos, ya que elegir una codificación incorrecta puede dar lugar a que el contenido de la tabla no sea legible o presente problemas en determinados caracteres, sobre todo cuando se trasvasa información de un sistema a otro o entre distintas aplicaciones. En este sentido, lo habitual si se va a trabajar en lengua española es elegir la codificación UTF-8 o bien la ISO-8859-1.

Consistencia de dominio

En segundo lugar, deberemos garantizar que el valor asignado a cada atributo se encuentra dentro de las opciones recogidas en el modelo de datos. Para ello, tanto el software GIS como las bases de datos que permiten gestionar datos espaciales disponen de mecanismos para definir los dominios de cada atributo. La definición de los dominios imposibilita la introducción de valores de atributos no contemplados en el modelo de datos, evitando desde su origen este tipo de error.

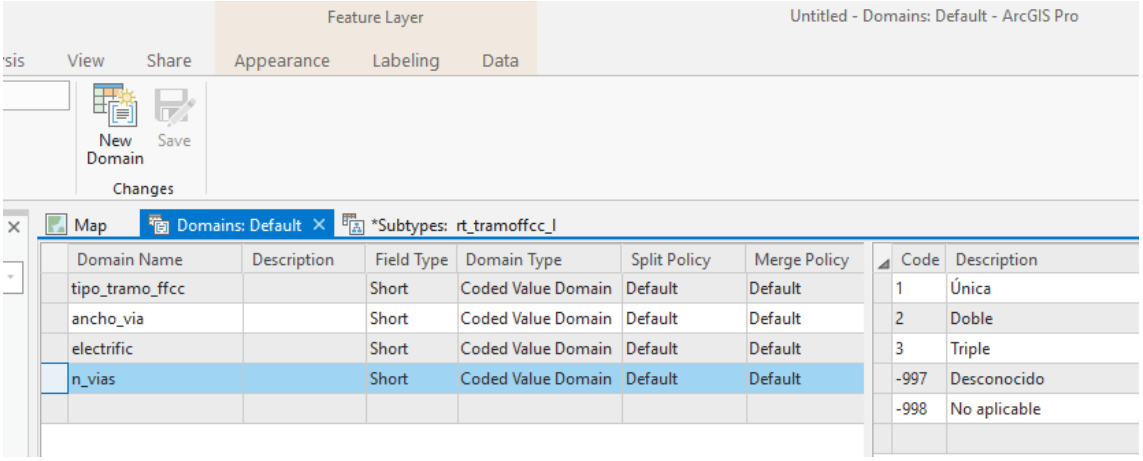


Ilustración 4-11: Tabla de definición de dominios en ArcGIS Pro

Consistencia conceptual

En tercer lugar, debemos garantizar el cumplimiento de las reglas de coherencia conceptual, que determinan cómo se relacionan entre sí los distintos elementos, en lo que respecta a los valores de sus atributos, o bien los valores de un mismo elemento unos respecto de otros.

Por ejemplo, veamos la imagen que sigue, que corresponde a un extracto del modelo de datos de la red ferroviaria recogida en la Información Geográfica de Referencia de Redes de Transporte del IGN de España. En esta tabla, que refleja parte del modelo conceptual de los tramos (geometrías lineales) que representan las vías ferroviarias, se puede ver la dependencia de algunos de sus atributos respecto a otros atributos o a otras tablas (si el atributo tipo_tramo indica que se trata de una playa de vías, los valores ancho_via, electrific y n_vias serán -998).

1.2.1. *Tabla: rt_tramoffcc_l*

Tabla rt_tramoffcc_l				
Atributo	**Obligatorio/ Condicional/ Opcional**	**Descripción**	**Tipo de dato**	**Código**
id_tramo	O	Identificador del tramo. Éste se genera conforme a la regla id_provincia * 10000M + secuencial, siendo id_provincia el código INE de provincia sobre el que se encuentra el tramo en cada caso.	Entero largo	
codigo	C	Código de tramo asignado por gestor de la línea de tipo Tren: - Cuando se trate de tramos de líneas de ADIF: EELLLTTTT (E - eje, L – línea, T - tramo) - Si no se conoce el valor del atributo: -997 - Este atributo no es aplicable a tramos de tipo "playa de vías". En estos casos el valor será -998.	CharacterString	
tipo_tramo	O	Especifica si pertenece a una playa de vías o no: - Atributo solo aplicable a los tramos de líneas de tipo Tren - En las líneas de otros tipos distintos de Tren: -998 - No es válido el valor: -997	***ValorTipoRed (Integer)***	
			Troncal	1
			Playa de vías	2
ancho_via	O	Tipo de ancho de vía según la clasificación de ADIF: - Atributo solo aplicable a los tramos de líneas de tipo Tren - Si no se conoce el valor del atributo en los tramos de líneas de tipo Tren: -997 - En las líneas de otros tipos distintos de Tren y en playas de vías: -998	***ValorAnchoVia (Integer)***	
			Ibérico	1
			UIC	2
			Métrico	3
			Mixto	4
electrific	O	Tramo electrificado o no: - Si no se conoce el valor del atributo: -997 - En playas de vías: -998 - En tramos troncales no se admite el valor: -998	***ValorElectrificado (Integer)***	
			Sí	1
			No	2
n_vias	O	Número de vías - Si no se conoce el valor del atributo: -997 - En playas de vías: -998 - En tramos troncales no se admite el valor: -998	***ValorNVias (Integer)***	
			Única	1
			Doble	2
			Triple	3

Ilustración 4-12: Extracto de las especificaciones de la Información Geográfica de Referencia (tramos de red ferroviaria)

En casos como este, deberíamos establecer los mecanismos necesarios para garantizar que el proceso de creación de una estación implementa la lógica de la relación identificador-municipio.

Para ello, existen diversas estrategias que dependen del formato o las herramientas con las que estemos trabajando. En el caso de que se trabaje con ArcGIS, podemos crear subtipos (siguiendo con el ejemplo anterior, para la tabla rt_tramoffcc_l tendría dos subtipos: troncal y playa de vías) de modo que determinados atributos de cada subtipo adquieran, cuando se crean, los valores por defecto que le correspondan de acuerdo con el modelo de datos).

Map | Domains: Default | *Subtypes: rt_tramoffcc_l

Subtype Name >		Troncal		Playa de vías	
Field Name	Data Type	Domain	Default Value	Domain	Default Value
OBJECTID	Object ID				
Shape	Geometry				
id_tramo	Double				
codigo	Text				
***tipo_tramo**	Long		1		2
ancho_via	Long				-998
electrific	Long				-998
estadofis	Long				
n_vias	Long				-998

Ilustración 4-13: Tabla de definición de subtipos en ArcGIS Pro

En el caso de que nuestros datos estén en Bases de Datos Espaciales, este aspecto se puede gestionar, por un lado, mediante la creación de claves externas que relacionan tablas entre sí, y mediante la creación de triggers.

Las claves externas (foreign key) se utilizan para garantizar la integridad referencial entre tablas. Una clave externa es un campo que relaciona cada registro de una tabla (tabla hija) con un registro de una tabla principal, a partir de otro campo en esta tabla. Un ejemplo de esto puede ser la tabla alfanumérica rrt_nodoctra_tramo, que relaciona los nodos de la red viaria (por ejemplo, gasolineras) con los tramos lineales mediante los cuales se conectan a la red viaria. en este caso, la tabla rrt_nodoctra_tramo sería la tabla hija, cuyas claves externas serían los identificadores de las tablas nodo y tramo, que son sus tablas principales. Gracias a que se ha establecido una restricción de clave externa, no se pueden crear registros en la tabla hija cuyas claves foráneas no estén en las tablas principales, y además si se borra un elemento de la tabla principal, se borrarán igualmente los hagan referencia a él en la tabla hija.

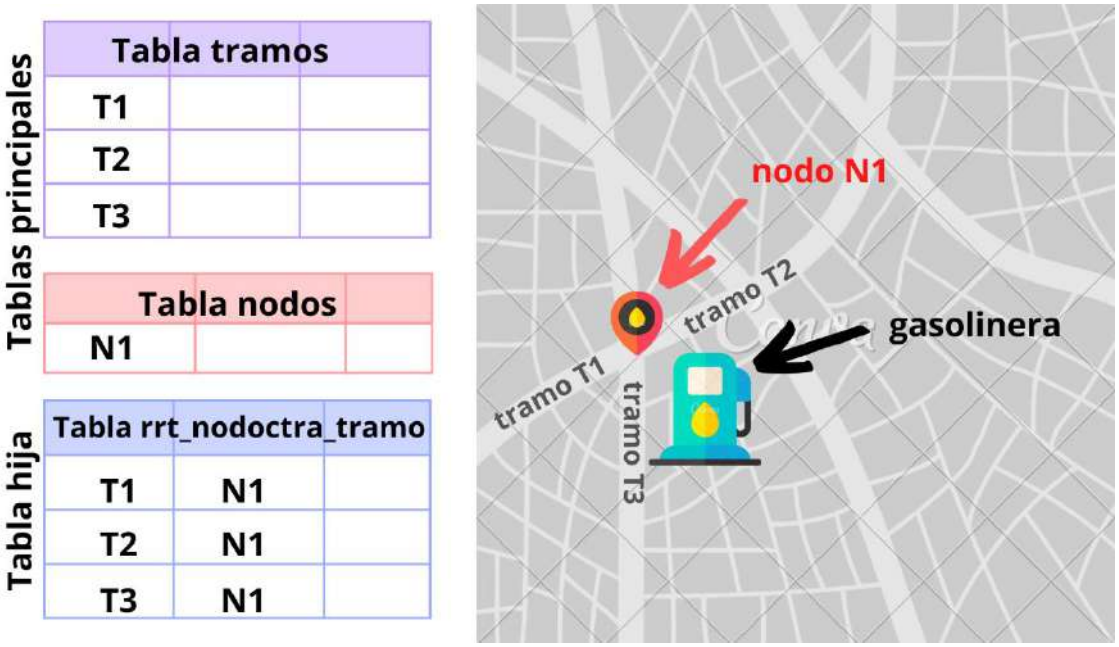


Ilustración 4-14:Ejemplo de uso de clave externa

Por otro lado, el segundo mecanismo que hemos mencionado (triggers), es una función que se acciona cuando un registro se crea, modifica o elimina, o cuando se modifica alguno de sus atributos. Con este mecanismo, podemos hacer que, si a un registro de la tabla rt_tramoffcc_l se le asigna el tipo_tramo=2, automáticamente los valores de los atributos ancho_via, electrific y n_vias pasen a ser -998, tal y como se define en el modelo de datos.

Consistencia topológica

Por último, debemos garantizar el cumplimiento de las reglas de coherencia topológica, que determinan cómo se relacionan espacialmente entre sí los distintos elementos. Los diversos paquetes de GIS disponen de herramientas de edición topológica que permiten asegurar ciertos aspectos (conectividad entre líneas, preservación de contigüidad entre polígonos colindantes si se mueve algún vértice o lado de alguno de ellos, etc.)

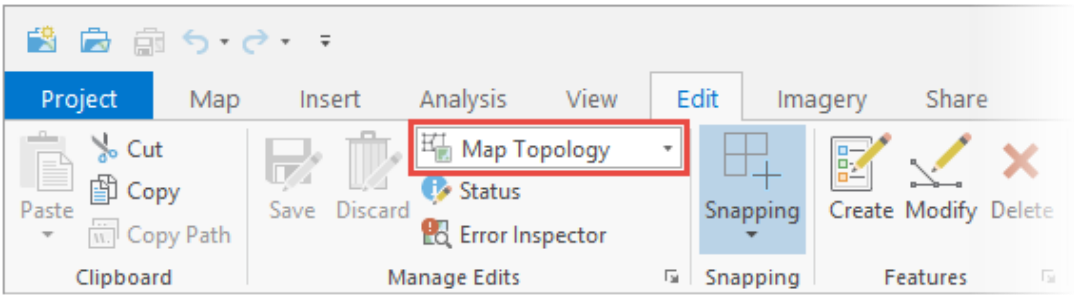


Ilustración 4-15: Herramienta de edición topológica de ArcGIS Pro. Fuente: https://pro.arcgis.com/es/pro-app/latest/help/editing/move-topology-edges-and-nodes.htm

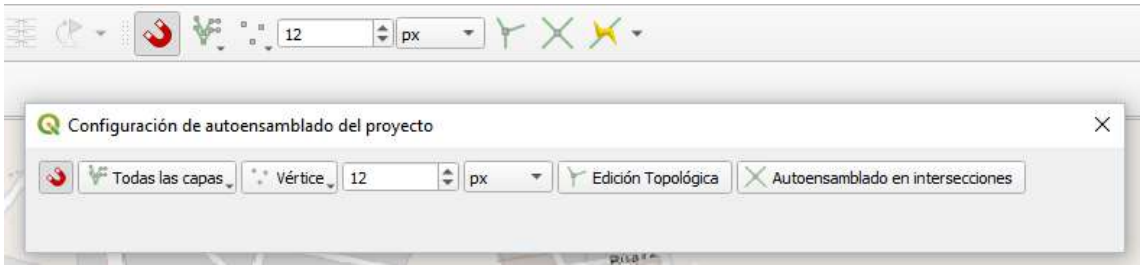


Ilustración 4-16: Herramienta de edición topológica de QGIS

c) *Exactitud posicional*

Exactitud posicional absoluta

En lo que respecta a la exactitud planimétrica, la única manera de garantizar que el resultado final de nuestro trabajo no presenta errores importantes será **establecer una escala de trabajo adecuada**. Ha de tenerse en cuenta que la escala necesaria para digitalizar ha de ser mucho mayor que la escala que se utiliza para hacer barridos visuales del ámbito de trabajo. El objetivo del barrido visual el detectar elementos, mientras que para digitalizarlos tendremos que utilizar una escala suficientemente grande como para reproducir el elemento de la forma más fiel posible a la realidad (tan fiel como sea necesario de acuerdo con los requisitos establecidos). Como ejemplo, en la captura de red viaria urbana y caminos, puede ser adecuada una escala 10,000 o 5,000 para detectar la existencia de elementos sobre una imagen aérea o de satélite, pero una vez detectados, para digitalizarlos correctamente deberemos aumentar considerablemente la escala.

Ilustración 4-17: Comparación entre dos escalas de visualización (a la izquierda, 5,000; a la derecha, 1,000)

En lo que respecta a la exactitud posicional absoluta altimétrica, puede optarse por dotar de cota a los elementos mientras se trabaja en ellos o al final del trabajo. En ambos casos, los paquetes de software GIS ofrecen herramientas adecuadas. Lo que en este caso condicionará el resultado final será la calidad de la información altimétrica de la que partimos. Por lo general, se asignarán cotas a partir de Modelos Digitales del Terreno, y será la resolución de dichos modelos la que condicione la calidad de nuestro producto.

Exactitud posicional relativa o interna

Para garantizar que las posiciones relativas de los fenómenos sean correctas, será deseable utilizar una **escala de trabajo adecuada** (al igual que en el caso anterior, ya que la exactitud absoluta garantiza la exactitud relativa) y establecer adecuadamente las **relaciones topológicas entre elementos**.

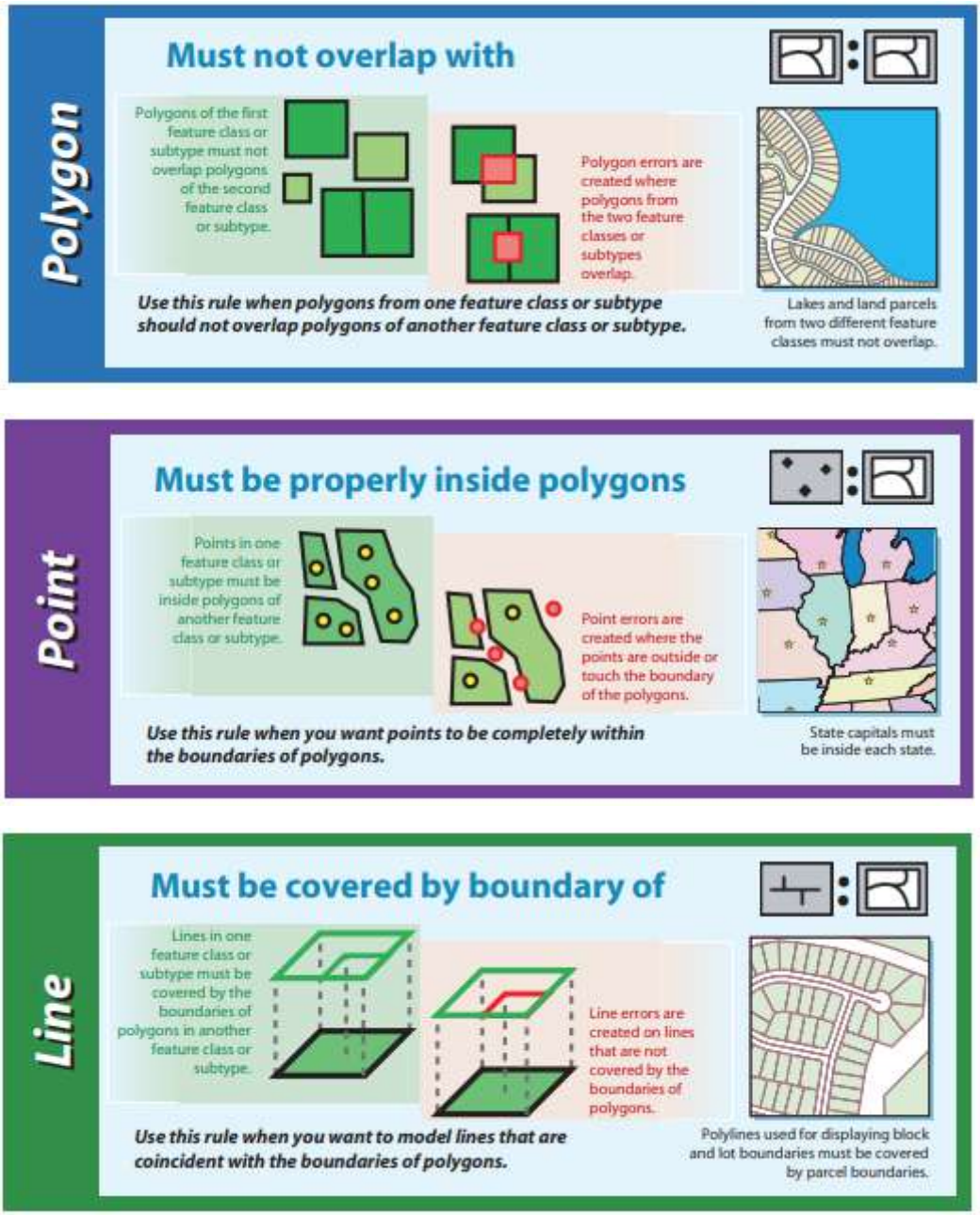


Ilustración 4-18: Ejemplos de reglas topológicas que establecen relaciones entre elementos de distintas capas (fuente: http://webhelp.esri.com/arcgisdesktop/9.2/pdf/topology_rules_poster.pdf)

d) *Exactitud temática*

Cuando se está editando una capa, en muchos casos resulta complicado garantizar la exactitud temática, especialmente cuando las tablas en las que trabajamos tienen gran número de atributos, que debemos rellenar cada vez que creamos un elemento. A continuación, se mencionan algunas estrategias recomendables para lograr una buena exactitud temática durante la etapa de edición:

Simbolización y etiquetado de elementos de acuerdo a sus atributos relevantes

De esta manera, a medida que vayamos editando iremos visualizando no solo la geometría de cada elemento, sino también su componente semántico, lo cual ayudará a detectar errores en los atributos que definen la visualización.

Definición de dominios y subtipos

Como ya se vio en el apartado b) Consistencia lógica, es posible definir dominios que eviten incluir valores no acordes a las especificaciones, y subtipos que automáticamente asignen una serie de valores por defecto a nuestros elementos. Esta funcionalidad no solo es importante para garantizar la consistencia lógica, sino también para evitar errores importantes de exactitud temática.

Definición de sesiones de trabajo distintas para los distintos elementos

Establecer un orden de edición en función de los distintos tipos de elementos puede ayudar a mejorar la exactitud temática final, además de mejorar la coherencia a la hora de aplicar criterios. Resulta bastante evidente que es más fácil lograr dicha coherencia si nos dedicamos específicamente a un elemento (como por ejemplo las carreteras), ya que ello nos permitirá ir descubriendo su problemática específica, y resolviendo los aspectos que puedan dar lugar a duda dentro de las especificaciones. Otra ventaja de definir sesiones de trabajo distintas es que podemos dar prioridad a los elementos más relevantes, y obtener cartografía adecuada para determinados usos antes que si abordáramos la digitalización o revisión de todos los elementos al mismo tiempo. Como contrapartida, esta forma de trabajar hará que probablemente tengamos que revisar una misma zona más de una vez, con las posibles demoras que ello pueda causar.

Ilustración 4-19: Ejemplo de simbolización de red viaria con distintos estilos en función de si se trata de un camino, una vía urbana, una carretera convencional u una autovía

4.3.6. Control de calidad del producto final

Independientemente del alcance de los controles realizados durante la fase de producción, al finalizar la misma debe llevarse a cabo una inspección del producto final. El objetivo de esta inspección será no solo identificar los errores detectados (los elementos que no cumplen con las características definidas en las especificaciones técnicas correspondientes), sino también localizarlos y clasificarlos para agilizar su corrección y facilitar el desarrollo de acciones de mejora.

El análisis del grado de calidad de un conjunto de datos se realiza, muy a menudo, comparándolo con otro conjunto de datos de referencia, que en teoría contendrá la totalidad de los datos que deseamos integrar. Dado que en este apartado se abordará la automatización de los procesos de inspección (validación) de los datos y corrección de las incidencias detectadas, se asumirá que los datos con los que trabajamos cumplen los requisitos establecidos en el apartado 4.3.3. En los casos en los que la inspección de calidad no sea automatizable, y el volumen de los datos a revisar no sea asumible (por costes o tiempo), deberán realizarse inspecciones manuales sobre una muestra representativa de los datos, tal y como se explica en posteriores apartados de este capítulo.

Finalmente, es necesario recordar que, al igual que en anteriores apartados, los parámetros de cada uno de los procesos relativos al control de calidad deberán ser coherentes con las especificaciones del producto y las variables definidas como relevantes.

4.3.6.1. Metodología para la creación de informes de errores

Como se indica al comienzo de este subapartado, una inspección de calidad de producto tiene como objetivo prioritario identificar y localizar errores. La razón primera de ello es la de proceder a su corrección. Pero existe otra razón de peso, y es la de realizar un análisis de los errores para identificar posibles sistematismos, problemas metodológicos o problemas de falta de formación en el personal que estén ocasionando que determinados errores sucedan con especial frecuencia o gravedad. Realizar este análisis será la única manera de resolver problemas latentes que nos conduzcan a una mejora continua del proceso de generación de datos espaciales.

Para que el resultado del proceso de inspección facilite este tipo de información, es aconsejable realizar a priori una clasificación de los posibles errores, agrupándolos en función de los aspectos que nos interesen y dándoles un código concreto y una descripción concisa.

Como ejemplo, la siguiente imagen muestra un extracto de la tabla en base a la cual se identifican los errores durante la Inspección de Calidad de los elementos de la red viaria de la Información Geográfica de Referencia de Redes de Transporte. En dicha tabla se facilita:

- El identificador único de cada tipo de error, que permitirá realizar, a posteriori, análisis estadísticos sobre aquellos errores más recurrentes.
- El objeto geográfico en el que se ha detectado el error.
- La descripción del error correspondiente.
- Un resumen de la metodología utilizada en la inspección, indicando si se trata de un control automático, semiautomático o manual.
- El subelemento de calidad al que hace referencia

Id	Objeto geográfico/Tabla	Descripcion	Metodologia	SubelementoCalidad	TipoControl
1297	Tramo	No pueden existir lazos o bucles en los tramos	Análisis topológico de lazos o bucles	Consistencia Topológica	Semiautomático
1298	Tramo	Los tramos no deberían tener vértices separados menos de 0,5 metros	Se calcula la distancia entre vértices y se detectan aquellos cuya distancia es menor de 0,5 metros	Consistencia Conceptual	Semiautomático
1299	Tramo	Hay muy pocos tramos tramos en el municipio	Se cuentan los elementos de la capa y se verifica si existen menos de 50	Omisión	Semiautomático
1301	Portalpk	No pueden existir registros con identificadores duplicados	Se detectan los registros que tengan el mismo valor para id	Comisión	Automático
1302	Portalpk	No pueden existir elementos duplicados geométricamente	Se detectan elementos duplicados geométricamente	Comisión	Automático
1303	Portalpk	El identificador único del elemento ha de comenzar por los cinco dígitos de su código INE de provincia y	Se reproyectan las capas de municipio y del elemento. A continuación se diferencian los municipios cuya provincia sea menor o mayor de 9, para así tomar los cuatro o los	Consistencia Conceptual	Semiautomático

Ilustración 4-20: Extracto de la tabla en base a la cual se identifican los errores de la red viaria de la IGR-RT.

Los elementos erróneos deberán identificarse con su correspondiente código de error y localizarse dentro de nuestro conjunto de datos. Para localizarlos, podemos añadir un campo de error en los datos, o bien crear archivos adicionales que contengan lo elementos erróneos o hagan referencia a ellos. Esta segunda opción nos da más flexibilidad a la hora de añadir

información de interés en la gestión del error (equipo o persona responsable del error, si se trata de un falso positivo, o cualquier otra circunstancia o aclaración que sea conveniente hacer).

4.3.6.2. Validación de datos

En este apartado se abordarán las distintas estrategias de validación del producto final (inspección del producto para comprobar que cumple con los requisitos especificados) para cada uno de los elementos de calidad recogidos por la ISO 19157.

a) *Compleción*

El análisis de compleción de un conjunto de datos debe realizarse siempre respecto a otro/s conjunto/s de datos fuente, que consideraremos que contiene/n la totalidad de los datos que deseamos integrar. La comparación entre los datos de referencia y los nuestros puede llevarse a cabo en base a criterios espaciales y/o alfanuméricos.

Detección de errores de omisión en base a criterios espaciales

Se buscará, para cada uno de los elementos presentes en los datos de referencia, un elemento homólogo en nuestro conjunto de datos, utilizando para ello funciones de análisis espacial. En el caso de que no se encuentre correspondencia con ningún elemento, se catalogará como error de omisión.

Detección de errores de comisión en base a criterios espaciales

Se buscará, para cada uno de los elementos presentes en nuestro conjunto de datos, un elemento homólogo en los datos de referencia, utilizando para ello funciones de análisis espacial. En el caso de que no se encuentre correspondencia con ningún elemento, se catalogará como error de comisión.

Detección de errores de omisión en base a criterios alfanuméricos

Se buscará, para cada uno de los elementos presentes en los datos de referencia, un elemento homólogo en nuestro conjunto de datos, utilizando para ello funciones de selección en base al valor de uno o varios atributos. En el caso de que no se encuentre correspondencia con ningún elemento, se catalogará como error de omisión.

Detección de errores de comisión en base a criterios alfanuméricos

Se buscará, para cada uno de los elementos presentes en nuestro conjunto de datos, un elemento homólogo en los datos de referencia. En el caso de que no se encuentre correspondencia con ningún elemento, se catalogará como error de comisión.

Consideraciones a tener en cuenta

Al hacer referencia a las funciones de análisis espacial en base a las cuales identificamos elementos homólogos en dos conjuntos de datos, podemos estar aludiendo a diversas herramientas:

- Selección de elementos iguales (mismo tamaño y ubicación).
- Selección de elementos parcial o totalmente coincidentes en extensión.
- Selección del vecino más próximo, dentro de una distancia máxima previamente establecida.

La utilización de una función u otra dependerá tanto del tamaño, de la forma, y la naturaleza de los objetos, porque cada uno de estos aspectos conlleva una problemática distinta:

- no tiene la misma problemática comparar áreas (por ejemplo, dos edificios de conjuntos de datos distintos), elementos lineales (por ejemplo, dos carreteras), y elementos puntuales, sobre todo si los dos conjuntos de datos comparados tienen distinta resolución.
- las comparaciones, en muchos casos, puede ser necesario realizarlas entre dos tipos de datos de distinta geometría (lo que en nuestro conjunto de datos se modela como un elemento superficial, en una fuente de datos de referencia y de menor resolución puede estar modelado como un elemento puntual)
- no cabe esperar los mismos resultados de comparar la ubicación de elementos superficiales de reducida extensión que de otros que ocupen grandes áreas,
- comparar el área que ocupan los elementos artificiales (edificios, canales, carreteras, parcelas de cultivo), por lo general es más sencillo que comparar el área ocupada por elementos naturales (bosques, pastos, ríos), ya que los primeros tienen límites bien definidos, mientras los límites de los segundos son más difusos y cambiantes.

Un ejemplo de esto se puede ver en la siguiente imagen, donde se comparan los datos de autopistas y autovías de IGR-RT (en azul) con los del Mapa Oficial de Carreteras del Ministerio de Infraestructuras de Transporte, Movilidad y Agenda Urbana (en rojo). Los primeros tienen una resolución espacial pensada para ser visualizada en web y GIS (hasta 1:1,000); los segundos tienen una resolución espacial determinada por el requisito de generar un mapa de carreteras de pequeña escala (1:1,000,000). La comparación espacial debe tener este aspecto en cuenta, ya que existen diferencias ostensibles entre las geometrías de ambas capas, a pesar de que hacen referencia a las mismas carreteras. Es necesario señalar que este problema es especialmente relevante cuando se trabaja con datos de escalas muy pequeña y que han sido desarrollados en proyecciones distintas a la de nuestro conjunto de datos.

Ilustración 4-21: Comparación entre dos conjuntos de datos lineales (IGR-RT y MOCI)

En lo que respecta a la búsqueda de datos homólogos entre los dos conjuntos de datos en base a determinados atributos, es evidente que dichos atributos deben estar presentes tanto en el conjunto de datos de referencia como en el conjunto que estamos sometiendo a inspección. Podemos utilizar uno o varios atributos, pero siempre deben ser únicos para cada elemento del conjunto. Conservar los identificadores de los datos de referencia en nuestro conjunto de datos es una buena práctica, ya que no solo facilitará la comparación, sino también el mantenimiento de los datos (esto es, seremos capaces de detectar rápidamente las modificaciones que se hagan en el futuro en los datos de referencia).

Corrección automática de errores

Una vez se obtienen los resultados de la inspección de calidad, puede llevarse a cabo una corrección automática (borrado de elementos que supongan error de comisión o incorporación de los elementos que identificados como errores de omisión). La opción de hacerlo dependerá de la fiabilidad que otorguemos a la fuente de referencia y la completitud de la misma.

Por ejemplo: si partimos de los datos oficiales del organismo que controla la navegación aérea del país, que informa acerca de la ubicación de los aeródromos existentes, podemos suponer que todos los aeródromos acerca de los cuales tiene datos existen en la realidad. De modo que, si en nuestro conjunto de datos falta alguno de los aeródromos que los responsables de navegación aérea tienen registrados, habrá altas posibilidades de que hayamos cometido un error de omisión, cuya corrección podríamos automatizar (copiando directamente todos los aeródromos que nos falten respecto a esa fuente). Sin embargo, puede que haya determinados aeródromos de carácter temporal (por ejemplo, aeródromos temporales para las brigadas forestales que operan durante verano para la prevención de incendios). Estos aeródromos, por su carácter temporal, podrían no estar recogidos en la fuente mencionada. Sin embargo, ello no implica que no existan ni que no sean de interés para nuestros datos. Por ello, si los borramos automáticamente al ver que no están en la fuente de referencia, estaríamos cometiendo un error. Debido a ello, si el número de errores lo permite, es aconsejable no realizar correcciones automáticas, sino revisar los errores detectados utilizando para ello otra fuente (por lo general imágenes aéreas).

b) *Consistencia lógica*

Tal y como se indicó en el apartado 4.3.5.2b), parte del éxito en la obtención de la calidad buscada, en lo que a los subelementos de consistencia lógica se refiere, reside en la correcta configuración de los archivosde trabajo y del entorno de trabajo., que ayudará a reducir al máximo los errores detectados durante la fase de inspección.

Detección de errores de consistencia de formato

- **Sistema de referencia geográfico (GRS), proyección empleada**: tanto el sistema de referencia como la proyección empleada son parte de los metadatos de los archivos de trabajo. Por ejemplo, si trabajamos con shapefiles, los archivos con formato .prj son los que indican el sistema y proyección en el que están definidos los datos. En este caso, crear un script o utilizar cualquier herramienta que sirva para leer recursivamente todos los archivos con esta extensión nos permitirá verificar que están en el sistema y proyección adecuados. En las geodatabases de Esri, por el contrario, la proyección no está definida en archivos aparte para cada capa; en este caso, tendremos que acceder a esta información a través de las funciones desarrolladas al efecto en la biblioteca de Python arcpy. Por último, en las bases de datos (por ejemplo, PostGIS o geopackage),

esta información está centralizada en una única tabla, que recoge dichos datos para la totalidad de las tablas).

- **Ubicación geográfica de los datos**: por lo general, conocemos bien los límites de nuestra zona de trabajo. Para comprobar si existen elementos fuera de ella, que en principio serán elementos erróneos, basta con realizar un análisis espacial (intersección, o selección por ubicación) de nuestros datos respecto a una capa que contenga las áreas que definen nuestro ámbito de interés. Lo que quede fuera de él habrá de revisarse (puede que se trate de elementos desplazados debido a la utilización de una proyección errónea o una reproyección poco adecuada; puede que el área que hayamos identificado haya de extenderse, a la vista de los resultados del control de calidad; o puede que, efectivamente, se trate de elementos erróneos que deberán ser corregidos).
- **Formato y nombre de los archivos de trabajo**: aunque el formato de los archivos puede comprobarse simplemente revisando las extensiones de los mismo en las carpetas de trabajo, podemos desarrollar otras metodologías que nos ayuden a agilizar esta comprobación. Suele ser efectivo crear una plantilla de proyecto en cualquier software GIS, en la que se importen capas con el formato y estructura deseada (aunque estén vacíos), apuntando a un directorio genérico. Luego se utilizará esa plantilla para cargar los datos que estamos desarrollando, cambiando la ubicación de los datos de prueba por la de nuestro directorio de trabajo. Si detectamos problemas en la lectura de alguna de las capas, deberemos revisarla para ver si efectivamente se debe a un problema con el formato o el nombre. También se pueden utilizar scripts que vayan leyendo todos los directorios, subdirectorios y archivos desde un archivo raíz que especifiquemos, para ir comparando los formatos y nombres de archivos con los que deberían aparecer.
- **Tipo de geometría (puntual, lineal o superficial, simples o multigeometrías)**: es altamente improbable que, si se ha definido correctamente la capa en el momento de su creación, se puedan dar errores de este tipo. Los errores más habituales suceden cuando se ha definido la capa con un tipo de geometría genérico (que acepta todo tipo de geometrías), cuando debido a procesos de análisis espacial surgen elementos nulos (es decir, sin geometría) y multigeometrías, o cuando la capa se crea de modo dinámico a partir de un proceso GIS o ETL, y no somos nosotros quienes definimos la geometría de la capa, sino que esa definición la da el software en función de determinadas variables. Para revisar este aspecto, tanto el software GIS como el software ETL tienen desarrolladas funcionalidades específicas.
- **Nombre y tipo de los atributos**: lo más habitual es utilizar software ETL o scripts personalizados para ir recorriendo todos los atributos de cada una de las tablas y comprobar su nombre y tipo respecto a las especificaciones.
- **Codificación de los atributos de tipo texto**: cuando este tipo de atributos no tienen asignada la codificación correcta, lo habitual es que aparezcan caracteres ilegibles en lugar de determinados caracteres específicos del idioma (las letras con acento o cedilla, la ñ, etc.). Este problema es difícil de evitar, sobre todo si importamos datos de otras fuentes que han sido desarrolladas con una codificación distinta. Para detectar este tipo de errores, lo más habitual es realizar una búsqueda de caracteres no permitidos (asterisco, guion bajo, etc.), que muestre registros susceptibles de tener este error.

Corrección automática de errores de consistencia de formato

- Los **errores relativos al Sistema de referencia geográfico (GRS) y proyección** asignados a nuestros datos pueden ser debidos a dos motivos: el primero es que no se haya definido ningún tipo de GRS ni proyección. El segundo es que el GRS y la proyección no sean los establecidos en las especificaciones. Todos los paquetes de software GIS, ETL y bases de datos con soporte espacial tienen funcionalidades específicas para resolver cualquiera de los dos problemas. En el primer caso, simplemente habrá que definir el GRS y proyección deseados (se parte del supuesto de que la ubicación era correcta, simplemente había una omisión de información). En el segundo caso, habrá que reproyectar. La reproyección supone un cambio de la ubicación y en ocasiones de la forma de los elementos (ya que cada proyección produce unas deformaciones distintas9, que puede llegar a afectar a la consistencia topológica de nuestros datos.
- Los **errores de formato y nombre de los archivos, y estructura de las tablas** son fácilmente subsanables con el software ETL, que permiten hacer transformaciones en una amplia variedad de formatos, renombrar atributos y cambiar el tipo de los mismos.
- En la mayor parte de los casos, los **errores relativos al tipo de geometría** no se pueden subsanar, por lo general, en el mismo archivo que contiene el error. Ni un shapefile ni una capa de una geodatabase pueden cambiar su tipo de geometría una vez creadas, de modo que deberá obtenerse un archivo nuevo, utilizando las funcionalidades presentes en cualquier paquete GIS o ETL para convertir las geometrías compuestas en simples (o las geometrías lineales en áreas o cualquier transformación relativa que requiera la capa para ser conforme a las especificaciones), y asegurándonos en este caso de que la geometría asignada es correcta.
- Para el caso de **geometrías nulas o que presenten determinados errores** (bucles, vértices duplicados), existen diversas herramientas que permiten su corrección automática (por ejemplo, la herramienta Repair Geometry de ArcGIS).

Detección de errores de consistencia de dominio

Como sabemos, los valores de los atributos han de ajustarse a las listas controladas o rangos de valores definidos en las especificaciones. Después de definir dominios en el proyecto, en la geodatabase o en las tablas de una base de datos, es imposible que se incluir valores fuera de dominio. Sin embargo, en algunos casos los dominios se definen posteriormente a empezar a trabajar con los datos, y los registros creados antes de definir las restricciones de dominio pueden tener este tipo de error.

Si los dominios se han definido, tanto los paquetes GIS como las bases de datos tienen funcionalidades propias para hacer esta comprobación.

Si no se han definido, se recomienda realizar la comprobación mediante herramientas ETL, que además nos permitirá realizar correcciones automáticas sobre los datos erróneos.

Detección de errores de consistencia conceptual

Este tipo de errores engloban, por lo general, varios atributos de una tabla o varias, y tienen que ver con cómo han representado en nuestro modelo de datos los aspectos y objetos del mundo real que son relevantes para nuestros requisitos funcionales.

Este tipo de comprobaciones conllevan la revisión simultánea de varias tablas y de las relaciones entre ellas. Un error de consistencia conceptual implica por lo general dos o más atributos, uno

de los cuales será necesariamente erróneo, pero a priori no podemos decir cuál (lo cual dificulta las correcciones automáticas).

Ejemplos de errores de consistencia conceptual

En la Tabla 4-2 se muestran, como ejemplo, varios errores de tipo conceptual que afectan a una única tabla, de acuerdo con los errores definidos para la red viaria de la Información Geográfica de Referencia del IGN.

El error con id 1192 nos indica que los identificadores dependen del tipo de vial, esto es: conforme a las especificaciones del producto, los identificadores de carretera se construyen de forma determinada, distinta que los de vía urbana o los de caminos. Si al realizar la comprobación (mediante software ETL por lo general) se detecta que el identificador no se ha construido conforme a las normas (el atributo identificador está subordinado al atributo tipo_vial), sabremos que es el identificador lo que debemos modificar, ya el tipo_vial representa una realidad concreta que no podemos falsear, mientras el identificador es un número dado únicamente en el contexto de nuestro trabajo. En casos como este, conocemos cuál de los dos atributos es el erróneo y podemos establecer correcciones automáticas.

Id	Objeto geográfico/Tabla	Descripción
1192	Vial	Incoherencia entre identificador del vial y el atributo tipo_vial.
1193	Tramo	Tramo de carretera con atributo de firme sin pavimentar
1213	Tramo	Tramo de tipo Rotonda, con sentido Doble

Tabla 4-2: Ejemplo de errores de consistencia conceptual que afectan a varios atributos de una misma tabla

El resto de filas de la Tabla 4-2 (errores 1193 y 1213) muestran elementos con atributos incoherentes dentro de una misma tabla: de acuerdo a las especificaciones de IGR-RT, no se considera carretera un tramo que no esté pavimentado, y en las rotondas la circulación siempre es de sentido único. En estos casos, sin mirar el dato, no podemos saber a priori qué atributo es incorrecto: en el primer caso, desconocemos si se trata de una carretera (estaría mal el atributo “firme”) o de un tramo sin asfaltar (estaría mal la clase “carretera”, ya que debería catalogarse como “camino”); en el segundo caso, tampoco se sabe si se trata de una rotonda (estaría mal el atributo “sentido”, que debería ser único) o de un tramo de otro tipo y sentido doble (estaría mal el tipo “rotonda”).

En la Tabla 4-3 se muestran, también como ejemplo, errores de tipo conceptual que afectan a varias tablas, de acuerdo a los errores definidos para la red viaria de la Información Geográfica de Referencia del IGN. Puede tratarse de tablas externas a nuestro modelo (como la relativa al error 1203) o a tablas internas.

Id	Objeto geográfico/Tabla	Descripción
1203	Tramo	El identificador único del elemento ha de comenzar por los cinco dígitos de su código INE de provincia y municipio
1313	Portalpk, Tramo	Si el sentido del tramo es Doble, el sentido de los puntos kilométricos ha de ser necesariamente Ambos sentidos y viceversa
1508	Puntoctra, Areactra	El elemento puntual de infraestructura viaria ha de estar siempre en el interior de un recinto de área de infraestructura viaria

Tabla 4-3: Ejemplo de errores de consistencia conceptual que afectan a varios atributos de distintas tablas

Detección de errores de consistencia topológica

Por lo general, las comprobaciones topológicas se llevan a cabo en varios pasos: en primer lugar, ha de crearse la topología de las capas, que en función del software empleado podrá crear o no una serie de capas adicionales con información de apoyo para agilizar el análisis; en segundo lugar, es necesario establecer las reglas que definen las relaciones permitidas entre elementos de una misma capa o entre elementos de cada capa respecto a las demás; y, finalmente, se llevan a cabo las comprobaciones basadas en lo anterior.

La Tabla 4-4 muestra algunas de las reglas definidas para el modelo de red viaria de la IGR-RT de IGN.

Id	Objeto geográfico/Tabla	Descripción	
1121	Nodoctra, Tramo	Los nodos han de estar físicamente entre dos tramos de red viaria	Topo
1221	Tramo	No pueden existir solapes en entidades de tipo lineal a menos que tengan distinto valor del atributo situación	Topo

Tabla 4-4: Ejemplo de errores de consistencia topológica

Las herramientas de software GIS que llevan a cabo análisis de topología dan por lo general la opción de ir corrigiendo los errores unan vez ha terminado el análisis, y para cada tipo de error suele ofrecer una corrección determinada, lo que agiliza el proceso. También da la opción de marcar el error como excepción, si se comprueba que trata simplemente de un falso positivo.

c) *Exactitud posicional*

La **inspección automática de la exactitud posicional relativa o interna** (recordemos que este aspecto se refiere a la proximidad entre las posiciones relativas de los fenómenos y sus respectivas posiciones relativas verdaderas) debería basarse en las reglas topológicas que hayamos establecido, de acuerdo a lo establecido en el apartado anterior.

La **inspección automática de la exactitud absoluta o externa** se basará en la comparación de las geometrías de nuestro conjunto de datos con las geometrías de un conjunto de datos de referencia que tomamos como verdadero.

Consideraciones a tener en cuenta

Antes de automatizar este tipo de inspección, es necesario tener en cuenta varios aspectos:

- En primer lugar, debemos saber que determinadas funciones de análisis espacial de los paquetes de software GIS y ETL pueden no funcionar adecuadamente con

multigeometrías (por lo tanto, si nuestro conjunto de datos o el de referencia admiten multigeometrías, en muchos casos será necesario realizar procesos de desagregación de dichas geometrías antes de comenzar la inspección).

- En segundo lugar, la comparación entre las geometrías de dos o más objetos geográficos llega a ser un proceso muy intensivo en el uso recursos de procesamiento, ya que puede ser necesario realizar comparaciones, vértice a vértice, de todas las geometrías de una capa respecto a las de otra. Por ello, es aconsejable limitar el ámbito de comparación de cada elemento de un conjunto respecto al otro (por ejemplo, delimitar el ámbito máximo de comparación de las geometrías, obligar a que las comparaciones se hagan entre elementos con un determinado atributo común, o identificar de forma previa a la inspección elementos homólogos entre los dos conjuntos de datos para que la comparación se haga exclusivamente entre los elementos homólogos).
- En tercer lugar, un exceso de vértices en nuestras geometrías hará que el proceso de comparación sea más lento, ya que implicará necesariamente un mayor número de comprobaciones (entendemos como exceso de vértices en un elemento la existencia de vértices que no influyen en la definición de la forma de un elemento a la escala de definida en nuestras especificaciones). Debido a esto, en ocasiones es deseable aplicar algún algoritmo de generalización de las geometrías antes de llevar a cabo este tipo de inspección.

Estrategias de inspección de exactitud posicional absoluta

El algoritmo de inspección diferirá en función del tipo de geometría con el que estemos trabajando.

- Solape (overlap) o intersección: si un elemento de nuestro conjunto de datos es correcto posicionalmente respecto a otro de un conjunto de datos de referencia, debería coincidir espacialmente con él. Es decir, el solape entre ambos debería ser perfecto. Sin embargo, esto es difícil que ocurra si la fuente de referencia ha tenido que reproyectarse (siempre existen deformaciones en los conjuntos de datos reproyectados, aunque sean pequeñas), si ambos datos están a escalas distintas o si el tipo de geometría no es la misma. En cualquier caso, este tipo de comprobación funciona de modo óptimo con elementos de tipo área (no es necesarios que tanto el conjunto de datos de referencia como el inspeccionado tengan geometrías superficiales, con que uno de ellos lo tenga es suficiente). Por ello, si ninguno de los dos conjuntos tiene geometrías de este tipo, es aconsejable realizar un buffer de alguno, o bien utilizar otro método de comparación.
- Distancia entre vértices de elementos: si un elemento de nuestro conjunto de datos es correcto posicionalmente respecto a otro de un conjunto de datos de referencia, los vértices que definen ambos elementos deberían estar muy próximos. La cuantificación de la variable proximidad (es decir, cuál cerca deben estar dos vértices para que la exactitud posicional se considere correcta) deberá hacerse en función de la escala de trabajo y las características de los datos. Este tipo de algoritmo es el más adecuado cuando trabajamos con elementos lineales, si bien debe tenerse en cuenta los que se ha indicado anteriormente respecto al uso intensivo de recursos en el análisis espacial basado en los vértices de los elementos.
- Vecino más próximo (nearest neighbor): si un elemento de nuestro conjunto de datos es correcto posicionalmente respecto a otro de un conjunto de datos de referencia, la distancia máxima entre ambos elementos deberá ser muy pequeña o nula. Este

algoritmo está especialmente indicado si tenemos geometrías puntuales. En cambio, si lo utilizamos con geometrías lineales, da resultados poco significativos (el hecho de que una línea esté cerca de otra, e incluso a distancia 0 si intersecan, no quiere decir que la línea del conjunto de datos inspeccionado sea posicionalmente correcta).

d) *Calidad temporal y exactitud temática*

En este epígrafe se abordan conjuntamente ambos apartados, ya que se aborda la exactitud de una medida de tiempo de la misma manera que la exactitud temática del resto de atributos.

Para realizar este tipo de inspección es necesario llevar a cabo varios pasos:

- En primer lugar, identificar, para cada uno de los elementos presentes en nuestro conjunto de datos, un elemento homólogo en los datos de referencia. Esta identificación puede estar basada en atributos identificadores o en criterios espaciales. Sin embargo, suele prevalecer la utilización de atributos identificadores, ya que en múltiples casos los datos temáticos de referencia pueden venir en formatos sin geometría (archivos de texto o csv, Excel, e incluso PDF).
- En segundo lugar, y una vez realizada la identificación de elementos homólogos, es necesario comprobar si los atributos de interés tienen el mismo dominio que los del conjunto de datos inspeccionado (por ejemplo, los tramos de carreteras reciben el código 1003 en la IGR-RT del IGN; cuando utilizamos datos de otros organismos para realizar la inspección de calidad, en ningún caso vienen con ese código asignado, sino con cualquier otro; si hiciésemos una comparación sin tener esto en cuenta, los resultados de la comparación serían engañosos). Por eso, si los dominios de ambos conjuntos de datos no son comparables, será necesario realizar un mapeo de alguno de los dos para hacer viable la comparación.
- Llevados a cabo estos pasos, ya se podría realizar la comparación de los atributos de interés, utilizando herramientas ETL, GIS o scripts propios.

4.4. Aspectos generales de calidad posicional

La exactitud posicional es, quizás, el elemento por excelencia de la calidad de la información geográfica. No significa que sea el más importante, por ejemplo, en un mapa de líneas de metro, posiblemente sea más importante la compleción y exactitud temática que la posición, pero sí es un elemento específico de los datos geoespaciales. Véase que la consistencia lógica, la compleción, la exactitud temática y la exactitud temporal podrían aplicarse a datos no espaciales, por ejemplo, a una base de datos que contiene el conjunto de personas que trabaja en una organización. Además, el carácter métrico que normalmente se asume que tienen los mapas, hace que sea un aspecto de la calidad al que siempre se le ha prestado especial atención. En este breve apartado se apuntan algunas indicaciones o comentarios a tener en cuenta en el control de la calidad posicional.

Tradicionalmente, y con carácter general, suele hablarse de que la exactitud posicional de los elementos que se muestran en un mapa debe estar por debajo de 0,2 milímetros a la escala del mapa. Así, los errores posicionales de un mapa a escala 1:10,000 deberían estar por debajo de 2 metros; en un 1:25,000, por debajo de 5 metros, y, en un 1:200,000, por debajo de 40 metros. No obstante, a medida que la escala es más pequeña, debe tenerse más en consideración la diferencia entre la posición real de los objetos geográficos y su representación en mapa que, en

muchas ocasiones, debe alterar deliberadamente la posición y simplificar las formas de los elementos para conseguir una correcta lectura del mapa.

La mayoría de los test y metodologías de control posicional que se realizan es en base a puntos. Esto aporta cierta simplicidad en cuanto a la medición de los elementos de control (lo que se considera verdad terreno) y al cálculo de la diferencia entre el punto en el mapa y en el control. Sin embargo, entraña una dificultad en cuanto a la propia definición de la población. Si los elementos son lineales o superficiales, en teoría habría infinitos puntos a lo largo de las líneas o de los contornos de los polígonos, si bien es verdad que los puntos intermedios a lo largo de un segmento (entre dos vértices de la línea) son muy difíciles de evaluar, ya que hay una indeterminación (a lo largo de la propia línea) de cuál es ese punto en la verdad terreno.

Por estos motivos, habitualmente se reduce la población a los vértices que conforman esos elementos lineales y superficiales, incluso se suelen descartar vértices que estén muy alineados con los adyacentes, por el mismo problema que se ha señalado anteriormente de determinar cuál es el punto en la verdad terreno. Sin embargo, reducir la población a vértices en los que hay un cambio dirección importante (por ejemplo, la esquina de una casa, un cruce de caminos o una curva acusada en un río) hace que escapen del control partes de las líneas o de los contornos que pueden ser importantes (por ejemplo, si se ha digitalizado con un número muy reducido de vértices).

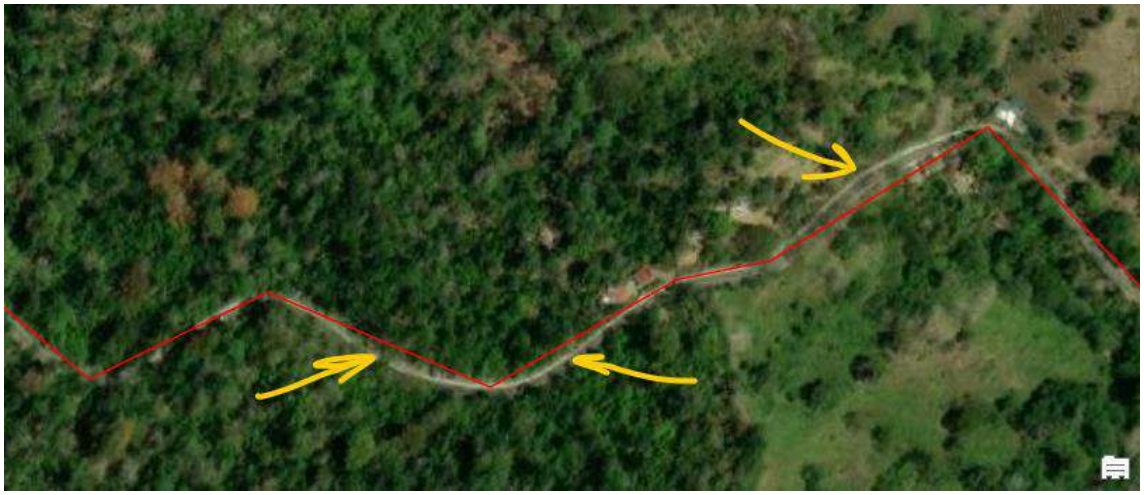


Ilustración 4-22: Camino digitalizado con número reducido de vértices. Véase que los vértices se encuentran bien posicionados, pero hay partes de la línea que se alejan notablemente de la verdad terreno.

Una alternativa al control por puntos es realizar un control posicional por líneas. Es decir, tomar como verdad terreno la propia línea medida en campo o digitalizada sobre una fuente de mayor exactitud. La comparación entre la línea del mapa y la considerada como verdad terreno se complica, ya que no vale la simple diferencia de coordenadas, como puede hacerse en el control por puntos. No obstante, se pueden aplicar distintas técnicas como, por ejemplo, la creación de búferes a varias distancias y determinación del porcentaje de línea que entra dentro de cada búfer.

Ilustración 4-23: Comparativa de camino a evaluar (línea roja) con búferes a 3 y 8 metros de la línea considerada como verdad terreno. Se puede determinar qué porcentaje de la línea se encuentra dentro de cada búfer.

Por otro lado, independientemente de que se utilicen puntos o líneas para realizar el control, es importante tener en cuenta el grado de incertidumbre que en ocasiones se da en la propia determinación de la verdad terreno. Esto sucede especialmente en escalas grandes, donde se buscan exactitudes mejores. Por ejemplo, si en las especificaciones del producto se indica que los ríos se capturan por el eje, este eje presenta cierto grado de indeterminación que puede ser, en ocasiones, de varios metros. Si se toma como punto para evaluar la exactitud posicional la confluencia de dos ríos, el punto exacto donde se produce esta confluencia es muy indefinido.

Ilustración 4-24: Punto de confluencia de los ejes de dos ríos (líneas azules). En amarillo se marca una zona de incertidumbre en la que se podría haber hecho la confluencia de los ejes.

Por último, se reseñan en este apartado algunos test creados por distintos organismos para evaluar la calidad posicional de sus productos y que han sido utilizados por diferentes organizaciones. Todos ellos se basan en puntos.

- Test NMAS (*National Map Accuracy Standards*).

  Se trata de un test muy sencillo utilizado desde 1947 por las administraciones americanas productoras de cartografía hasta hace pocos años. El resultado del test indica si la cartografía cumple o no cumple.

  Descripción:

  - Se comprueba si el 10% de los puntos de la muestra tiene un error horizontal mayor de 0,085 cm (1/30’’) en mapas de escala mayor de 1:20,000 o de 0,05 cm (1/50’’) en mapas de escala menor o igual a 1:20,000.

- También se puede aplicar sobre la componente vertical, comprobando si el 10% de los puntos de la muestra tiene un error vertical mayor de la mitad de la equidistancia de las curvas de nivel.
- Conviene verificar previamente que se cumplen las condiciones de aleatoriedad y normalidad.

Procedimiento:

- Seleccionar una muestra de al menos 20 puntos.
- Determinar la componente horizontal y vertical (en su caso) de error de cada punto, *i*:

$$e_{H_i} = \sqrt{\left(x_{t_i} - x_{m_i}\right)^2 + \left(y_{t_i} - y_{m_i}\right)^2}$$

$$e_{z_i} = z_{t_i} - z_{m_i}$$

Donde

$x_{t_i},\ y_{t_i},\ z_{t_i}$ son las coordenadas consideradas como verdad terreno.

$x_{m_i},\ y_{m_i},\ z_{m_i}$ son las coordenadas mapa.

- Determinar máximos tolerables. Para la componente horizontal, 0,085 cm (1/30'') en mapas de escala mayor de 1:20,000 y 0,05 cm (1/50'') en mapas de escala menor o igual a 1:20,000). Para la componente vertical, la mitad de la equidistancia.
- Contar cuántos puntos exceden los máximos tolerables.
- Si más del 10% de los puntos lo superan, no se pasa el test. En caso contrario, sí se pasa.

- Test EMAS (*Engineering Map Accuracy Standard*).

Especifica la exactitud de los mapas topográficos a gran escala. Fue desarrollado por la ASPRS (*American Society of Photogrammetry and Remote Sensing*) junto con la ASCE (*American Society of Civil Engineers*) y el *American Congress on Surveying and Mapping*.

Permite definir umbrales de aceptación de error a un nivel de confianza. Está basado en hipótesis sobre límites aceptables para los errores aleatorios y sistemáticos.

Descripción:

- Se evalúan las componentes X, Y, Z de forma separada.
- Efectúa test sobre la media, bajo supuesto de varianza poblacional desconocida, para determinar si el error sistemático es aceptable, y test sobre la varianza para determinar si el error aleatorio es aceptable, respecto a una varianza preestablecida.
- Conviene verificar previamente que se cumplen las condiciones de aleatoriedad y normalidad.

Procedimiento:

- Seleccionar una muestra de al menos 20 puntos.
- Calcular el error de cada punto, *i*, en cada componente:

$$e_{x_i} = x_{t_i} - x_{m_i} \qquad e_{y_i} = y_{t_i} - y_{m_i} \qquad e_{z_i} = z_{t_i} - z_{m_i}$$

Donde

$x_{t_i},\ y_{t_i},\ z_{t_i}$ son las coordenadas consideradas como verdad terreno.

$x_{m_i},\ y_{m_i},\ z_{m_i}$ son las coordenadas mapa.

- Calcular el error medio de cada componente:

$$\overline{e}_x = \frac{1}{n}\sum_{i=1}^{n} e_{x_i} \qquad \overline{e}_y = \frac{1}{n}\sum_{i=1}^{n} e_{y_i} \qquad \overline{e}_z = \frac{1}{n}\sum_{i=1}^{n} e_{z_i}$$

- Calcular la desviación típica muestral en cada componente:

$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{x_i} - \overline{e}_x\right)^2} \quad S_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{y_i} - \overline{e}_y\right)^2} \quad S_z = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(e_{z_i} - \overline{e}_z\right)^2}$$

- Realizar <u>test sobre la media</u>, bajo supuesto de varianza poblacional desconocida, para determinar si el error sistemático es aceptable. Se establecen las siguientes hipótesis:

$$H_0: \mu_p = 0 \qquad H_1: \mu_p \neq 0$$

Se determinan los siguientes estadísticos:

$$t_x = \frac{\overline{e}_x\sqrt{n}}{S_x} \qquad t_y = \frac{\overline{e}_y\sqrt{n}}{S_y} \qquad t_z = \frac{\overline{e}_z\sqrt{n}}{S_z}$$

Se toma el valor de la distribución *t de Student* con n-1 grados de libertad y nivel de significación α: $t_{n-1,\alpha/2}$

El mapa, o la IG, superará el test (se cumple la hipótesis H_0) con un nivel de confianza α si se cumple que:

$$|t_x| \leq t_{n-1,\alpha/2} \qquad |t_y| \leq t_{n-1,\alpha/2} \qquad |t_z| \leq t_{n-1,\alpha/2}$$

- Realizar <u>test sobre la varianza</u> para determinar si el error aleatorio se encuentra dentro de los límites aceptables. Se establecen las siguientes hipótesis en relación a un valor máximo de varianza, a^2, preestablecido:

$$H_0: \sigma_p^2 \leq a^2 \qquad H_1: \sigma_p^2 > a^2$$

El valor máximo de varianza, a^2, se establece en función de la exactitud posicional que se exija. Por ejemplo, para una cartografía 1:25,000, donde los errores máximos (al 95%) en las componentes planimétricas deberían estar acotados en 5 metros, se puede suponer una varianza máxima, a priori, de $2{,}5^2$.

Se determinan los siguientes estadísticos:

$$\chi_x^2 = \frac{S_x^2(n-1)}{a_x^2} \quad \chi_y^2 = \frac{S_y^2(n-1)}{a_y^2} \quad \chi_z^2 = \frac{S_z^2(n-1)}{a_z^2}$$

Se toma el valor teórico de la distribución χ^2 con n-1 grados de libertad y nivel de significación α: $\chi^2_{n-1,\alpha}$

El mapa, o la IG, superará el test (se cumple la hipótesis H_0) con un nivel de confianza α si se cumple que:

$$\chi_x^2 \leq \chi_{n-1,\alpha}^2 \qquad \chi_y^2 \leq \chi_{n-1,\alpha}^2 \qquad \chi_z^2 \leq \chi_{n-1,\alpha}^2$$

- Test NSSDA (*National Standard for Spatial Data Accuracy*)

Propuesto por el *Federal Geographic Data Committee* (FGDC, 1998), vino a sustituir al NMAS y al EMAS.

Se basa en el error medio cuadrático (EMC) de la muestra y es el usuario, o la agencia productora de cartografía, quien decide los umbrales de aceptación.

Descripción:

- Se evalúa la componente horizontal (X, Y de forma conjunta) y la componente vertical.
- Se fundamenta en el EMC horizontal, por un lado, y vertical, por otro, y calcula el error de la muestra analizada en función de un determinado nivel de confianza impuesto por el usuario (generalmente, del 95%).
- Conviene verificar previamente que se cumplen las condiciones de aleatoriedad y normalidad.

Procedimiento:

- Seleccionar una muestra de al menos 20 puntos.
- Calcular el error de cada punto, *i*, en cada componente:

$$e_{x_i} = x_{t_i} - x_{m_i} \qquad e_{y_i} = y_{t_i} - y_{m_i} \qquad e_{z_i} = z_{t_i} - z_{m_i}$$

Donde

$x_{t_i},\ y_{t_i},\ z_{t_i}$ son las coordenadas consideradas como verdad terreno.

$x_{m_i},\ y_{m_i},\ z_{m_i}$ son las coordenadas mapa.

- Calcular el error medio cuadrático (EMC) de cada componente:

$$EMC_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_{x_i}}^2} \quad EMC_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_{y_i}}^2} \quad EMC_z = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_{z_i}}^2}$$

- Obtener el valor $NSSDA_H$ (horizontal):
  - Si EMCx ≈ EMCy (se puede considerar si 0,6 < EMC_{min}/EMC_{max})

$$NSSDA_H = 2{,}4477\sqrt{\frac{EMC_x^2 + EMC_y^2}{2}}$$

  - Si EMCx ≠ EMCy

$$NSSDA_H = 2{,}4477 * 0{,}5\left(EMC_x + EMC_y\right)$$

(2,4477 es el valor de la distribución Normal circular al 95%)

- Obtener el valor $NSSDA_Z$ (vertical):

$$NSSDA_Z = 1{,}96\ EMC_z$$

(1,96 es el valor de la distribución Normal al 95%)

- Incluir, junto a la leyenda del mapa, las sentencias:

  "Se ha verificado una exactitud horizontal de [$NSSDA_H$] metros al 95% de confianza"

  "Se ha verificado una exactitud vertical de [$NSSDA_Z$] metros al 95% de confianza"

4.5. Muestreos

4.5.1. Necesidad de muestreo y marco normativo

La norma ISO 19157:2013, dentro del método de evaluación directa, define básicamente dos métodos de inspección de los elementos a la hora de realizar el control de calidad:

- Inspección completa
- Muestreo

La inspección completa comprueba cada ítem de la población especificada por el ámbito de la calidad de los datos y tiene sentido cuando la población es pequeña o las pruebas pueden realizarse de manera automática. Sin embargo, en muchas ocasiones no se cumple ninguna de estas dos premisas, es decir, la población es grande y no es posible la automatización. De hecho, las pruebas relacionadas con los elementos de la calidad de exactitud posicional, exactitud temática y compleción en muy pocos casos se pueden automatizar. Por tanto, es necesario tomar una muestra representativa de la población para poder inferir la calidad de esta.

El resto de alusiones que se dan en la norma referidas a los muestreos es en anexos informativos, salvo la referencia que se hace en el anexo C (normativo), Diccionario de datos para la calidad de datos, al definir el modelo UML de evaluación de la calidad de datos, donde hay una entrada para el *samplingScheme*, que simplemente debe ser rellenada con un texto libre indicando información sobre el tipo de esquema de muestreo y descripción del procedimiento del mismo. En los anexos E, Evaluación e informe de la calidad de datos, y F, métodos de muestreo para la evaluación, se proporcionan pautas que pueden resultar muy útiles a la hora de definir las muestras y documentar el proceso realizado.

Pero se debe reseñar que ambos anexos (el E y el F) son de carácter informativo, por lo que no constriñen cómo se deben realizar los muestreos. Es decir, la única obligación que se marca en la norma es la de informar cómo se ha hecho.

4.5.2. Definiciones y estrategia de muestreo

En el anexo F de la ISO 19157:2013 se proporciona un esquema con los componentes de la estrategia de muestreo que puede resultar muy útil y que se reproduce a continuación.

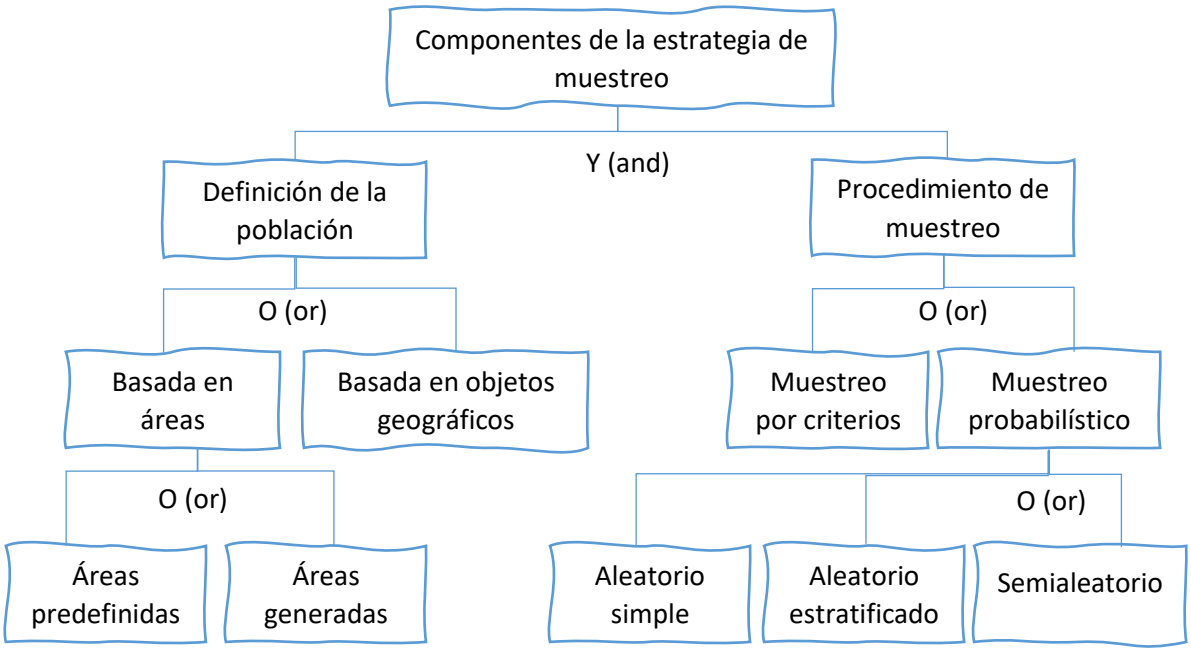


Ilustración 4-25: Componentes de la estrategia de muestreo según la ISO 19157:2013.

Así pues, el primer paso en el proceso de muestreo es definir la población, que principalmente se puede distinguir entre las basadas en áreas o basadas en objetos geográficos. También se pueden hacer otros supuestos, como decidir que la población está definida por los vértices que componen todos los elementos. Por ejemplo, en un control de la exactitud posicional por puntos, se puede decidir que lo que se va a evaluar es la posición de los vértices que componen los objetos geográficos de un tema concreto.

Dentro de las poblaciones basadas en áreas, en ocasiones puede interesar emplear áreas ya definidas, como pueden ser los municipios de un territorio, o generar áreas siguiendo un patrón, por ejemplo, una cuadrícula con un tamaño de 2 km x 2 km.

También es posible definir una primera etapa de muestreo basada en áreas, seguida de un muestreo basado en objetos geográficos que están dentro de las áreas seleccionadas. En cualquier caso, se debe establecer un criterio sobre los elementos que se encuentran a caballo entre dos áreas.

Por otro lado, se debe definir el procedimiento para tomar las muestras. Se puede distinguir entre muestreo por criterios y muestreo probabilístico. En el primer caso, las muestras se toman en base a un conocimiento experto o juicio profesional. Puede resultar útil, por ejemplo, cuando se quiere inspeccionar elementos que, por algún motivo, son susceptibles de presentar algún error. Sin embargo, los resultados de un muestreo basado en criterios no se pueden extrapolar al conjunto de la población desde un punto de vista estadístico.

En el segundo caso, se aplica la teoría de muestreos, por lo que cada elemento de la población tiene una probabilidad conocida de ser seleccionado y los resultados se pueden extrapolar al conjunto de la población con un error de estimación acotado. A continuación, se indican tres tipos de muestreos probabilísticos habituales:

- Aleatorio simple: se emplean número aleatorios para seleccionar los elementos de la población, por lo que cada uno tiene la misma probabilidad de ser seleccionado que el

resto. Resulta útil cuando la población es relativamente homogénea (en el aspecto que se quiere controlar), es decir, no presenta patrones o agrupamientos notables.

- Aleatorio estratificado: si la población presenta patrones o agrupamientos, se puede dividir en estratos o subpoblaciones, de manera que los elementos de un estrato mantengan cierta homogeneidad y sean diferentes a los de los otros estratos. Es importante que no haya solapamiento entre los diferentes estratos, ya que en cada uno se toma una muestra aleatoria y un mismo elemento de la población global no debe ser candidato en dos de las subpoblaciones.

  Una de las desventajas de los muestreos aleatorios es que puede ocurrir que no cubran geográficamente todo el ámbito de la población, es decir, en una muestra tomada por número aleatorios, los elementos seleccionados pueden quedar más o menos concentrados en una zona.

- Semialeatorio: en este caso, se realiza una selección aleatoria de los elementos iniciales de la muestra seguida de una regla para la selección de los restantes. Puede resultar útil, por ejemplo, para definir una regla o un patrón que, después de la selección del elemento o elementos iniciales, asegure que los restantes se distribuyen a lo largo de toda el área que cubre la población.

4.5.3. Muestreos probabilísticos

Como se ha comentado, los muestreos de base probabilística permiten extrapolar los resultados obtenidos al conjunto de la población acotando, con un cierto nivel de confianza, el margen de error que se asume por el hecho de inspeccionar sólo una muestra y no el total de la población.

4.5.3.1. Estimación de parámetros e intervalo de confianza

Por ejemplo, si se toma una **muestra aleatoria simple**, la estimación del parámetro de interés se puede determinar cómo se señala a continuación, según el tipo de parámetro:

- Para la media

$$\widehat{\overline{X}} = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

  Es decir, la estimación de la media de la población es la media de la muestra.

  Por ejemplo, se quiere conocer el tiempo medio que va a requerirse para la edición de cada hoja de la serie del mapa topográfico 1:25,000 de un determinado país. Para ello, se ha realizado este trabajo en una muestra de hojas y se ha contabilizado el tiempo requerido en cada una de ellas. La media de tiempos de la muestra es la estimación del tiempo medio para cada hoja de la población.

- Para el total

$$\hat{X} = N\overline{x}$$

  Es decir, la estimación del total de la población es igual a la media de la muestra multiplicada por el número de elementos.

Siguiendo con el ejemplo anterior, la estimación del tiempo total de edición que se requiere para la edición de toda la serie del mapa topográfico nacional se obtiene multiplicando el tiempo medio de la muestra por el número de elementos de la población.

- Para la proporción

$$\hat{P} = p = \frac{1}{n}\sum_{i=1}^{n} A_i$$

Donde $A_i = \begin{cases} 1 \; si \; se \; cumple \; la \; cualidad \\ 0 \; en \; otro \; caso \end{cases}$

Es decir, la estimación de la proporción de una clase o de una cualidad en la población es igual a la proporción que se da de dicha cualidad en la muestra.

Por ejemplo, se quiere conocer la proporción de elementos cuyo atributo nombre se ha rellenado de forma incorrecta (no contiene el nombre real o considerado como verdadero). Para ello, se ha revisado este atributo en una muestra de elementos, comparando su valor con fuentes de información fiables o con visitas de campo. La estimación de la proporción de nombres incorrectos en la población es la propia proporción de elementos de la muestra en los que esto ocurre.

- Para el total de la clase

$$\hat{A} = N\hat{P}$$

Siguiendo con el ejemplo anterior, la estimación del número total de elementos con atributo nombre incorrecto es igual a la estimación de la proporción de elementos con nombre incorrecto multiplicada por el número de elementos.

Estas estimaciones tienen una varianza, que depende de las características de la población y son, por tanto, desconocidas. Se recurre a estimar las varianzas a partir de la información que proporciona una muestra:

- Para la media

$$\hat{V}\left(\widehat{\overline{X}}\right) = \frac{N-n}{Nn}s^2$$

Donde s^2 representa la cuasi-varianza muestral $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$

- Para el total

$$\hat{V}(\hat{X}) = N^2\hat{V}\left(\widehat{\overline{X}}\right)$$

Es decir, la estimación de la varianza del total es igual a la de la media por el número de elementos de la población al cuadrado.

- Para la proporción

$$\hat{V}(\hat{P}) = \frac{N-n}{N(n-1)}p(1-p)$$

- Para el total de la clase

$$\hat{V}(\hat{A}) = N^2\hat{V}(\hat{P})$$

Véase que, en todos los casos, conforme *n* es más grande y se acerca a *N*, esta varianza disminuye. Esto es lógico, ya que, si se inspecciona toda la población, la estimación coincide con el propio parámetro.

A partir de la estimación del parámetro y de su varianza, se puede construir un intervalo de confianza que indica en qué rango se encuentra, con un nivel de confianza determinado, la variable que se desea conocer.

$$\theta \in \left(\hat{\theta} - Z_{1-\alpha/2}\sqrt{\hat{V}(\hat{\theta})} \ , \ \theta + Z_{1-\alpha/2}\sqrt{\hat{V}(\hat{\theta})}\right)$$

4.5.3.2. Tamaño de muestra a partir de error de estimación

Se puede realizar el camino inverso para tratar de determinar el tamaño de muestra necesario para no exceder un error de estimación prefijado. Este error de estimación serían los términos que suman y restan en el intervalo de confianza.

$$\varepsilon = Z_{1-\alpha/2}\sqrt{\hat{V}(\hat{\theta})}$$

Se puede despejar el tamaño de muestra, *n*, de la expresión anterior.

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Donde n_0 tiene una expresión diferente según los cuatro parámetros, θ, vistos anteriormente:

- Para la media, $\theta = \overline{X}$

$$n_0 = \frac{Z_{1-\alpha/2}^2 s^2}{\varepsilon^2}$$

- Para el total, $\theta = X$

$$n_0 = \frac{Z_{1-\alpha/2}^2 N^2 s^2}{\varepsilon^2}$$

- Para la proporción, $\theta = P$

$$n_0 = \frac{Z_{1-\alpha/2}^2 p(1-p)}{\varepsilon^2}$$

- Para el total de la clase, $\theta = A$

$$n_0 = \frac{Z_{1-\alpha/2}^2 N^2 p(1-p)}{\varepsilon^2}$$

  Donde *p* y s^2 son estimaciones de la proporción y de la varianza que pueden ser obtenidas a partir de un estudio piloto. En el caso de la proporción, se puede establecer el valor más desfavorable, 0,5, que maximiza el valor de $p(1-p)$

  Véase también que, en caso de poblaciones muy grandes, $n = n_0$

En resumen, esta formulación permite determinar qué tamaño debe tener la muestra, como mínimo, para que la estimación que se realiza a partir de ella no exceda de un cierto valor, que se denomina error de estimación, ε, y con un cierto nivel de confianza, α. Como es de esperar, cuanto menor se quiera que sea el error de estimación, más tamaño de muestra se necesita.

Igualmente, cuanta mayor confianza se desee en que no se supere ese error de estimación, el valor de $Z^2_{1-\alpha/2}$ es más elevado y obliga a aumentar también el tamaño de la muestra.

Véase también que el error de estimación, ε, va expresado en las mismas unidades que el parámetro que se quiere estimar. Por ejemplo, si se quiere estimar el tiempo medio en horas de edición de cada hoja 1:25,000, el error de estimación también se debe expresar en horas; pero si el objetivo es conocer la proporción de elementos con atributo nombre erróneo, el error de estimación se expresa en la misma unidad que la proporción.

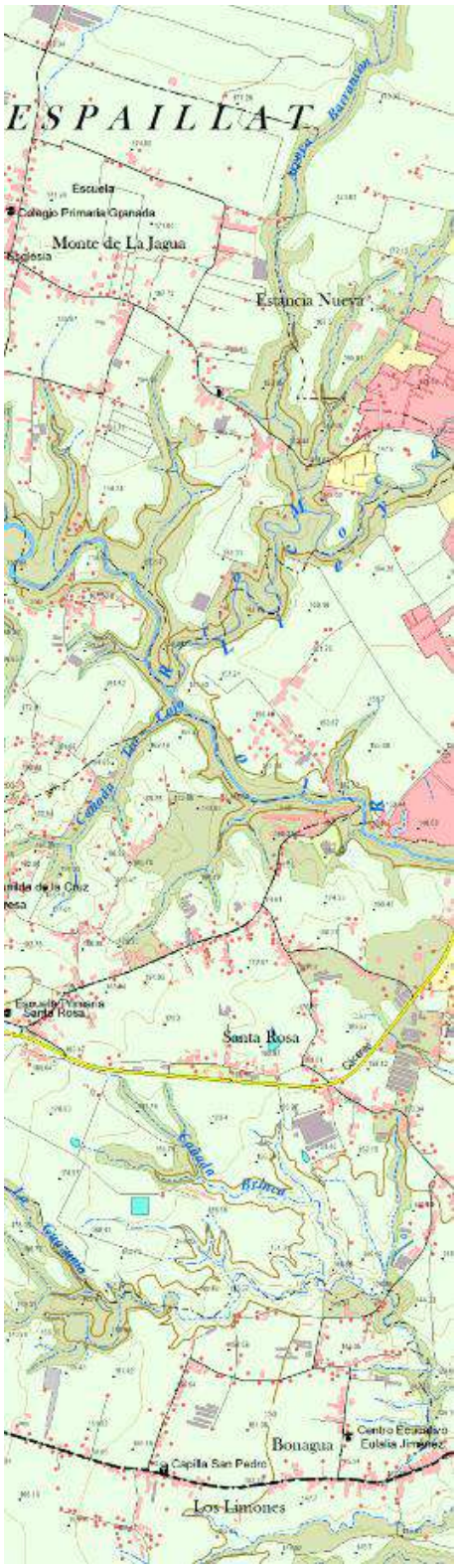
4.5.3.3. Introducción a las normas ISO 2859 y 3951

Por su parte, la norma ISO19157:2013, hace referencia y se apoya en otras normas internacionales existentes para la inspección por muestreo, en particular en las normas ISO 2859, con sus diferentes partes, y la ISO 3951. Se trata de normas que provienen del ámbito industrial y son la base de lo que se denominan los procesos de aceptación por muestreo. La aceptación o rechazo se realiza bajo un paradigma estadístico en el que se asumen ciertos riesgos (riesgos de usuario y de productor), en beneficio de no inspeccionar el 100% de los elementos.

La ISO 2859 hace referencia a los procedimientos de muestreo para la inspección por atributos, esto es, características cuya consideración hace que una unidad pueda ser clasificada como “buena” o como “defectuosa”; mientras que la 3951 lo hace para la inspección por variables, entendiendo como tal, características que pueden ser medidas sobre una escala continua, por ejemplo, una discrepancia posicional media en metros. Ambas permiten determinar planes de muestreo eficaces que garanticen que la calidad no sea inferior a ciertos niveles previamente establecidos.

Al provenir del ámbito de la industria, en ellas es importante el concepto de lote, mínima unidad para la que la calidad puede evaluarse, e ítem, mínima unidad a inspeccionar. Como se indica en el anexo F de la ISO 19157:2013, la ISO 2859-1 se aplica principalmente a la inspección de series continuas de lotes, la 2859-2 puede aplicarse a lotes individuales o aislados, y la 2859-3 se aplica a procedimientos de muestreo por lotes salteados. Por ejemplo, en un proceso industrial por el que se reciben lotes periódicos de un determinado producto, tiene sentido la aplicación de la 2859-1.

Para la utilización de estas normas, se debe definir un Límite de Calidad Aceptable (LCA), que es quien va a especificar el nivel de conformidad de la calidad de un conjunto de datos. Así, básicamente en función de este LCA y del tamaño de los lotes, se proporcionan una serie de tablas que indican cuál debe ser el tamaño de la muestra y cuáles son los límites para aceptar o rechazar los lotes (por ejemplo, cuantos rechazos puede haber como máximo en la muestra).



CAPÍTULO 5: PRODUCCIÓN DEL MAPA I. GENERACIÓN DE LOS DATOS CARTOGRÁFICOS

Métodos de producción: Proceso automático y Edición simplificada

Descripción breve

Descripción de conceptos básicos de cartografía y exposición de los procesos y metodologías de producción, automática y con edición simplificada, aplicables sobre la información geográfica almacenada para la generación de los datos cartográficos de un Mapa Topográfico 1:25,000

Índice de contenidos

5.1. Introducción a la Cartografía. Conceptos básicos 1

5.1.1. El mapa como representación de la realidad: parámetros matemáticos 2

Sistema geodésico de referencia 2

Proyección cartográfica 4

Escala 6

Representación del relieve 7

Declinación y convergencia 8

5.1.2. Cartografía Topográfica 9

5.1.2.1. Contenido de un Mapa Topográfico 1:25,000 9

Según la tipología de la representación del objeto geográfico 9

Según las temáticas que componen al mapa 10

5.1.2.2. Contenidos del MTN25 10

Orografía 10

Hidrografía 10

Comunicaciones 11

Construcciones 11

Coberturas y Usos del suelo 12

Líneas límite administrativas 13

Lugares o áreas de especial protección 13

Vértices geodésicos 14

Toponimia 14

Exteriores 15

5.1.2.3. La simbología y la jerarquía de capas en la estructuración de contenidos 15

5.1.2.4. El mapa topográfico comparado con la mera fotografía aérea 17

5.2. Producción automática de cartografía: Mapa Automático 19

5.2.1. Antecedentes: documentación del proceso tradicional 19

5.2.2. Del mapa tradicional al mapa automático. Motivación 21

5.2.3. Fuentes origen 24

5.2.3.1. Selección de las fuentes de datos 24

5.2.3.2. Extracción por unidades de producción 25

5.2.3.3. Traducción al modelo cartográfico 26

5.2.4. Procesos automáticos de redacción cartográfica 28
5.2.4.1. Redacción cartográfica: generalización y edición cartográfica 28
5.2.4.2. Automatización del proceso 29
(1) Datos origen 30
(2) Análisis 31
(3) Selección de estrategia de actuación 31
(4) Aplicación de estrategia de actuación 31
(5) Refinamiento 32
(6) Salida Cartográfica 32
5.2.4.3. Operadores utilizados 32
Filtrado de elementos (Eliminación/Selección) 32
Simplificación de formas 33
Suavizado 34
Amalgamado 34
Colapso 34
Exageración/Ensalzamiento 35
Desplazamientos 35
Tipificación 36
Ajuste de contornos 36
Refinamiento estético 36
5.2.5. Etiquetado automático 37
5.2.5.1. Clasificación de elementos. Asignación de propiedades de texto 37
5.2.5.2. Selección de elementos a etiquetar 38
5.2.5.3. Asignación de jerarquías 39
5.2.5.4. Reglas de etiquetado. Elementos puntuales, lineales y superficiales 40
5.2.5.5. Detección de solapes. Resolución de conflictos 41
5.2.6. Salidas Intermedias 42
5.2.6.1. Planimetría 43
5.2.6.2. Coberturas y usos del suelo 44
5.2.7. Almacenamiento en Base de Datos Cartográfica 45
5.2.8. Diseño e implantación del flujo de automatización 46
5.2.8.1. Componentes generales 47
Entrada del proceso 47
Salida del proceso 47
Software utilizado 47
Parámetros 48

5.2.8.2. Flujo de operaciones 49
1. Proceso preliminar 49
2. Traducción del modelo 49
3. Tratamiento de elementos lineales 49
4. Tratamiento de edificaciones 50
5. Altimetría 53
6. Tratamiento de símbolos puntuales 53
7. Proceso automático de rotulación 53
8. Incorporación de otras fuentes de origen 54
9. Almacenamiento en la base de datos cartográfica 55
5.3. Mapa automático con edición adicional 57
5.3.1. Objetivo 57
5.3.2. Proceso de producción 58
5.3.2.1. Archivos de partida y finales 58
5.3.2.2. Protocolo (check-list) 59
5.3.2.3. Desarrollo, actividades y mapa del proceso 61
Planificación y gestión del trabajo en el IGN 63
Edición del contenido de la hoja 63
Generación del FFS del contenido de la hoja 69
Generación del PDF con exteriores y sombreado 69
Edición de los exteriores en el pdf 69
Revisión del PDF final 72
Control de calidad de la entrega por muestreo 72
Revisión del pdf final en talleres 72
Registro 73
5.3.3. Diferencias con el proceso tradicional 73

Índice de ilustraciones

Ilustración 5 - 1: El modelo Ratajski 1
Ilustración 5 - 2: Perfil de la superficie terrestre, el geoide y el elipsoide 2
Ilustración 5 - 3: Ejemplo de ondulación del geoide expresada en metros 3
Ilustración 5 - 4: Clasificación geométrica de las proyecciones. Fuente: Mapping for a sustainable world. UN & ICA 4

Ilustración 5 - 5: Proyecciones tangentes y secantes 5
Ilustración 5 - 6: Ámbito territorial de cada huso U.T.M 5
Ilustración 5 - 7: Dos ejemplos de escalas gráficas lineales 6
Ilustración 5 - 8: Gráfico conceptual de la equidistancia 8
Ilustración 5 - 9: Valores de declinación y convergencia para una hoja del MTN25 9
Ilustración 5 - 10: Orografía de MTN25 10
Ilustración 5 - 11: Hidrografía de MTN25 11
Ilustración 5 - 12: Comunicaciones de MTN25 11
Ilustración 5 - 13: Construcciones de MTN25 12
Ilustración 5 - 14: Usos del suelo de MTN25 12
Ilustración 5 - 15: Representación de las líneas límite jurisdiccionales 13
Ilustración 5 - 16: Representación de los límites de las zonas de especial protección 13
Ilustración 5 - 17: Representación de un vértice geodésico 14
Ilustración 5 - 18: Toponimia y Líneas límite administrativas de MTN25 14
Ilustración 5 - 19: Exteriores de una hoja de MTN25 15
Ilustración 5 - 20: Cuadro de símbolos cartográficos de MTN25 16
Ilustración 5 - 21: Fragmento de tabla de simbolización por atribución semántica 16
Ilustración 5 - 22: Ejemplo de estructura ordenada de capas para su correcta visualización 17
Ilustración 5 - 23: Detalle del Mapa Topográfico Nacional a escala 1:25,000 (MTN25) 18
Ilustración 5 - 24: Composición de diferentes ediciones de MTN, Ortofoto y LiDAR 19
Ilustración 5 - 25: Cambio de paradigma en la producción de cartografía. 22
Ilustración 5 - 26: Diferentes ámbitos de actuación: por hojas, por provincias, por capas. 22
Ilustración 5 - 27: Visualizador de cartografía e imágenes 23
Ilustración 5 - 28: Aplicación de reglas de edición para generar el Mapa 24
Ilustración 5 - 29: Fuentes de referencia para el MTN25 25
Ilustración 5 - 30: Extracción desde una base de datos continua, por hoja de MTN25 26
Ilustración 5 - 31: Extracción de elementos superficiales por intersección con hoja de MTN25 26
Ilustración 5 - 32: Esquema del proceso de traducción del modeloBTN-MTN25 27
Ilustración 5 - 33: Correspondencia a través de combinaciones de atributos. 27
Ilustración 5 - 34: Traducción del modelo. Atributos asignados a cada combinación 28
Ilustración 5 - 35: Traducción del modelo por análisis espacial. 28
Ilustración 5 - 36: Redacción Cartográfica, contexto y objetivo 29
Ilustración 5 - 37: Esquema conceptual del flujo de automatización 30
Ilustración 5 - 38: Proceso de filtrado aplicado a piscinas y puntos de cota 33
Ilustración 5 - 39: Selección de caminos a eliminar en función de su longitud y ubicación 33
Ilustración 5 - 40: Simplificación de formas aplicada al contorno de un edificio 33
Ilustración 5 - 41: Suavizado de caminos dentro del perímetro de un parque 34
Ilustración 5 - 42: Amalgamado de edificaciones que se encuentran próximas 34
Ilustración 5 - 43: Colapso de las dos calzadas de una carretera en una sola 35
Ilustración 5 - 44: Exageración de distancias entre edificios para acentuar esqueleto urbano . 35
Ilustración 5 - 45: Desplazamientos aplicados a edificaciones y vías de comunicación 35
Ilustración 5 - 46: Tipificación aplicada a la capa de edificaciones de una urbanización 36
Ilustración 5 - 47: Ajuste de contornos en bordes de cultivos y edificios 36
Ilustración 5 - 48: Aplicación de símbolo con un patrón en un área de central solar 37
Ilustración 5 - 49: : Categorización de los textos de entidades de población en la leyenda del MTN25 37
Ilustración 5 - 50: Categorización para los rótulos de los elementos río y embalse en la leyenda del MTN25 38

Ilustración 5 - 51: Atributo Jerarquía de la tabla 0302L_RIO en el modelo de datos de BTN..... 38
Ilustración 5 - 52: Tabla de decisión sobre rotulación según el atributo 'Tipo' 39
Ilustración 5 - 53: Generalización de textos por selección dentro de una partición 39
Ilustración 5 - 54: Asignación de jerarquías a textos mediante comparación con ediciones históricas de MTN25 y MTN50 en FME..... 40
Ilustración 5 - 55: Reglas de edición de elementos superficiales y lineales. Módulo MapTextLabeller en FME 41
Ilustración 5 - 56: Esquema del proceso de eliminación de solapes de textos en FME..... 42
Ilustración 5 - 57: Salida en formato DGN simbolizada..... 43
Ilustración 5 - 58: Salida de planimetría en formato DGN 43
Ilustración 5 - 59: Salida de coberturas del suelo en formato DGN..... 44
Ilustración 5 - 60: Salidas intermedias en el flujo de producción 44
Ilustración 5 - 61: Ubicación de la BBDD Cartográfica dentro del flujo de automatización 45
Figura 5 - 62: Procesos de redacción cartográfica, ubicación dentro del flujo general..... 47
Ilustración 5 - 63: Ilustración de implementación mediante FME 48
Ilustración 5 - 64: Flujo implementado para automatizar la redacción cartográfica..... 49
Ilustración 5 - 65: Salida cartográfica de edificaciones en zona de densidad media..... 51
Ilustración 5 - 66: salida cartográfica de edificaciones en casco urbano..... 51
Ilustración 5 - 67: Árbol de decisión implementado para gestionar las actuaciones sobre edificaciones..... 52
Ilustración 5 - 68: Clasificación de edificios según particiones de edificación..... 52
Ilustración 5 - 69: Parametrización de rotulado mediante el módulo MaptextLabeller de FME 54
Ilustración 5 - 70: Detalle de la estructura de tablas almacenada en Postgis..... 55
Ilustración 5 - 71: Ejemplo de ítems y roles que participan en la Edición Simplificada 60
Ilustración 5 - 72: Calificación de los ítems incumplidos 60
Ilustración 5 - 73: Extracto de la pestaña de contenido_exteriores 60
Ilustración 5 - 74: Extracto de la pestaña imprenta 61
Ilustración 5 - 75: Extracto de la pestaña de resultados 61
Ilustración 5 - 76: Mapa de proceso de la Edición Simplificada..... 62
Ilustración 5 - 77: Extracto de ítems relativos a errores en la información origen..... 63
Ilustración 5 - 78: Extracto de ítems relativos a composición de la hoja 64
Ilustración 5 - 79: Extracto de ítems relativos al proceso previo 64
Ilustración 5 - 80: Antes y después de ejecutar los ítems del proceso previo 64
Ilustración 5 - 81: Extracto de ítems relativos a la toponimia..... 66
Ilustración 5 - 82: Extracto de ítems relativos a inconsistencias..... 67
Ilustración 5 - 83: Ejemplos de inconsistencias entre la hidrografía y las comunicaciones en la fuente origen..... 67
Ilustración 5 - 84: Extracto de ítems relativos a simbologías..... 67
Ilustración 5 - 85: Símbolo de cementerio invertido..... 67
Ilustración 5 - 86: Extracto de ítems relativos a vías de comunicación 68
Ilustración 5 - 87: Ítem relativo a cortafuegos 68
Ilustración 5 - 88: Extracto de ítems relativos a hidrografía 68
Ilustración 5 - 89: Extracto de ítems relativos a controles a efectuar en la hoja..... 69
Ilustración 5 - 90: Extracto de ítems relativos a revisión del contenido y exteriores 70
Ilustración 5 - 91: Flujo del proceso de Edición Simplificada 73

Índice de tablas

Tabla 5-1: Clasificación de mapas 1
Tabla 5-2: Listado de requisitos para el diseño de la Base de Datos Cartográfica..... 46
Tabla 5-3: Denominación de las tablas en función del Tema y Tipo de geometría 55
Tabla 5-4: Atributos almacenados para cada geometría 56
Tabla 5-5: Atributos almacenados para la información de cada hoja..... 56

5.1. Introducción a la Cartografía. Conceptos básicos

Se puede definir la Cartografía como el conjunto de estudios y operaciones científicas y técnicas que intervienen en la formación o análisis de mapas, modelos en relieve o globos, que representen la Tierra, o parte de ella, o cualquier parte del Universo. La Asociación Cartográfica Internacional define el concepto de mapa como *la representación convencional gráfica de fenómenos concretos o abstractos, localizados en la Tierra o en cualquier parte del Universo.*

Una clasificación posible de los mapas en función de su escala y contenido sería:

Mapas generales	Mapas especiales	
- Mapas **topográficos** a escala grande, que describen accidentes naturales o artificiales del terreno - Mapas cartográficos que representan grandes regiones - Mapas del mundo entero	- Mapas políticos - Mapas urbanos - Mapas de comunicaciones - Mapas científicos de diversas clases - Mapas económicos y estadísticos	- Mapas turísticos - Mapas artísticos o publicitarios - Cartas para la navegación marítima y aérea - Mapas catastrales a gran escala

Tabla 5-1: Clasificación de mapas

En cualquier caso, un mapa aparece como un conjunto de dibujos, palabras escritas y signos de muy variado aspecto, a veces muy sencillo, como un croquis dibujado en un papel; otras muy complejo, impreso en muchos colores. Pero un mapa es siempre el esquema de una realidad, y su formación obedece a numerosos acuerdos y convenciones que deben conocerse para su correcta interpretación.

Los elementos que configuran la elaboración de un mapa de una determinada realidad georreferenciable y como el posible usuario percibe esa realidad se pueden sintetizar, según el **Modelo de Ratajski**, en siguiente. Pudiéndose decir que el objetivo que persigue la cartografía es que el ángulo δ que aparece en dicha figura sea lo más pequeño posible.

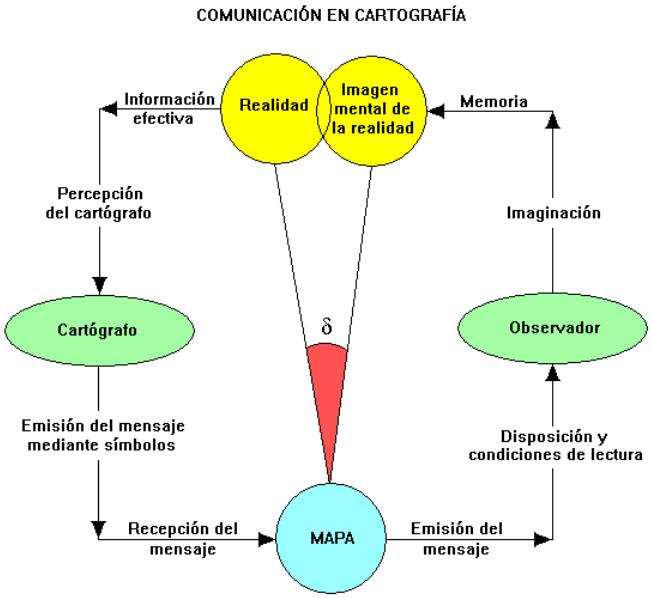


Ilustración 5 - 1: El modelo Ratajski

5.1.1. El mapa como representación de la realidad: parámetros matemáticos

Un mapa es un documento **métrico**, es decir, el tamaño de lo cartografiado se corresponde con sus dimensiones reales, salvo las alteraciones que se introducen en él con objeto de permitir una adecuada legibilidad.

Los problemas esenciales y evidentes que afectan a la representación de la superficie terrestre son dos. El primero, está causado por la dimensión de la superficie terrestre, que es mucho mayor que la que puede emplearse para su representación. El segundo consiste en que la superficie de representación, si excluimos globos y maquetas, generalmente es plana, mientras que la superficie terrestre no lo es. El primer problema conduce al concepto de escala; el segundo se desdobla en dos aproximaciones: los métodos de representar el relieve y los métodos de proyección cartográfica.

A la hora de realizar la representación gráfica de la superficie terrestre sobre un plano intervienen una serie de factores, algunos de naturaleza matemática, que es fundamental tener en cuenta y que son los que a continuación se exponen brevemente.

Sistema geodésico de referencia

Todo el conjunto de medidas efectuadas para la confección de un mapa que se llevan a cabo sobre la superficie de la Tierra, están influidas por la irregular distribución de la masa terrestre que afecta a la dirección de la gravedad y por tanto a la determinación de la vertical de cada lugar. Debido a esto, todas las observaciones terrestres se efectúan sobre una figura irregular llamada **geoide**, que se define como *la superficie equipotencial respecto a la fuerza de gravedad al nivel medio del mar, y en la que la dirección de la gravedad es perpendicular en todos los lugares*. El geoide coincide con la superficie media del mar en sus partes oceánicas y está por debajo de la corteza terrestre en las zonas continentales. La diferencia entre la superficie terrestre y el geoide es consecuencia de la irregular distribución de montañas y depresiones.

Ilustración 5 - 2: Perfil de la superficie terrestre, el geoide y el elipsoide

El geoide también está afectado por el movimiento de rotación de la Tierra, que le produce un abombamiento en la región ecuatorial y un achatamiento en las regiones polares. Para determinar la forma del geoide es necesario medir la gravedad en numerosos puntos de la Tierra. Las secciones del geoide son curvas irregulares, aunque aproximadamente circulares.

Sin embargo, para elaborar los mapas con precisión deberá utilizarse una superficie geométrica regular. Las observaciones realizadas sobre el geoide habrá que transferirlas, pues, a una figura regular que más se aproxime a él. Siendo ésta el **elipsoide**, figura geométrica engendrada por la revolución de una elipse sobre su eje menor. En geodesia clásica, para

efectuar los cálculos de paso del geoide al elipsoide, se hace coincidir tangencialmente a éste con el geoide en un punto llamado **datum**. A medida que sucesivas observaciones van proporcionando un conocimiento más exacto del geoide, se calculan elipsoides que se adaptan mejor al geoide. Hay dos opciones a la hora de elegir un elipsoide: optar por uno que se ajuste lo máximo posible al geoide en la zona que queremos representar en un mapa; u optar por un elipsoide que, aunque no se ajuste óptimamente al geoide en la zona que nos interesa, sea el que mejor se ajusta al geoide en su totalidad. Está claro que la primera opción es más precisa, pero localista; mientras que la segunda opción sacrifica la precisión en aras de una homogeneidad internacional.

Las observaciones basadas en satélites artificiales **GNSS** dan como resultado coordenadas geocéntricas, sobre un elipsoide, de los puntos en que se hace la observación. En la actualidad se utiliza como sistema de referencia el elipsoide WGS84 (World General System referido al año 1984) por acuerdo de la UGGI (Unión Geodésica y Geofísica Internacional), como es el caso de la República Dominicana es el WGS84. En otros países se utilizan otros sistemas de referencia como es el caso de España, que emplea el ETRS89, salvo en las Islas Canarias que se emplea el REGCAN95, ambos basados en el elipsoide GRS80.

Las distancias entre el geoide y los elipsoides no son grandes. En ocasiones estas diferencias son positivas, y en otras negativas, siendo lo común que el elipsoide esté por debajo del geoide en las zonas continentales y por encima en los océanos, según muestra la ilustración anterior. Ahora vemos, mediante curvas de nivel, la **ondulación del geoide** (diferencias entre el geoide y el elipsoide) de la Unión Astronómica Internacional.

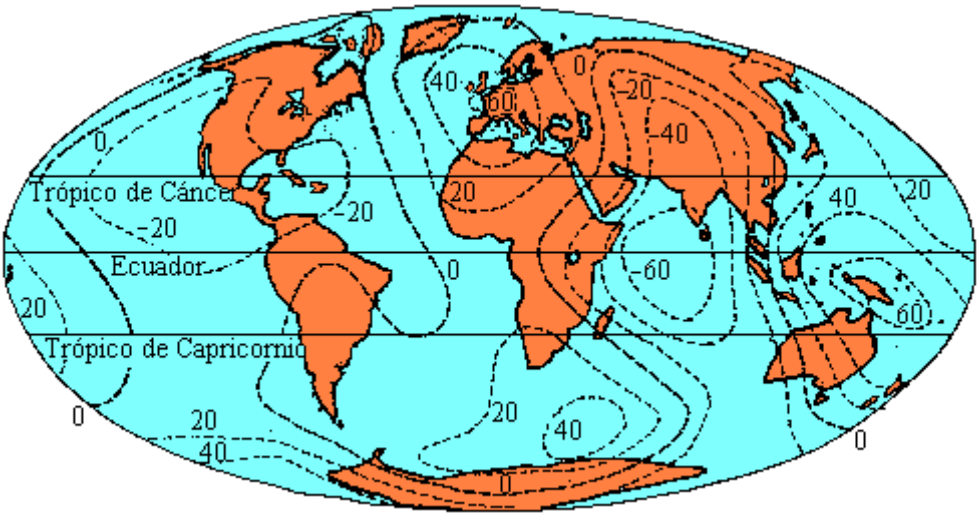


Ilustración 5 - 3: Ejemplo de ondulación del geoide expresada en metros

Una vez que se tengan todos los datos situados sobre un elipsoide, se podrán referir a un sistema geográfico de localización. Está internacionalmente aceptado el sistema formado por una red de paralelos y meridianos que permiten determinar la posición de un punto en la Tierra mediante sus **coordenadas geográficas**: **latitud y longitud**. Utilizando como origen de latitudes el Ecuador del elipsoide y como origen de longitudes el meridiano de Greenwich o meridiano 0° del elipsoide.

Proyección cartográfica

Un elipsoide no es una figura desarrollable en un plano. Para poder representar una zona de la superficie terrestre en un plano con las menores deformaciones posibles se acude al auxilio de las **proyecciones cartográficas** que establecen una correspondencia biunívoca entre los puntos del terreno proyectados sobre el elipsoide y los del mapa.

Cualquier proyección o sistema de representación cartográfica introduce deformaciones a la hora de representar el terreno. Aunque estas deformaciones pueden afectar a infinidad de características, como por ejemplo el paralelismo de los paralelos, la convergencia de los meridianos, las intersecciones perpendiculares de paralelos y meridianos, los polos representados como elementos puntuales, etc., las alteraciones más importantes son aquellas relacionadas con ángulos, superficies y distancias.

Las proyecciones que no deforman los ángulos se denominan **conformes**. Las proyecciones conformes mantienen los ángulos alrededor de puntos, o lo que es lo mismo, las líneas en el elipsoide forman al cortarse el mismo ángulo que sus representaciones en el plano (el ángulo de dos curvas que se cortan es el ángulo de sus tangentes). Las proyecciones no conformes tendrán **anamorfosis angular** (diferencia entre el ángulo en la proyección y el ángulo real).

Otras proyecciones conservan la superficie de cualquier figura en todas las regiones, recibiendo el nombre de **equivalentes**. Las proyecciones no equivalentes presentan **anamorfosis superficial** (cociente entre una superficie en el mapa y la superficie real). Por su propia definición una proyección no puede ser a la vez equivalente y conforme. Así, todas las proyecciones conformes presentarán regiones similares con tamaños desiguales y todas las equivalentes deformarán la mayoría de los ángulos.

La conservación de las distancias en las proyecciones cartográficas sólo se puede mantener a lo largo de determinadas líneas que se denominan **automecoícas**, o en direcciones que parten de determinados puntos diciéndose entonces que la proyección es **equidistante**. La deformación de las distancias en zonas en que estas no se conservan se denomina **anamorfosis lineal**.

Una clasificación geométrica de las proyecciones nos lleva a considerar si el paso del elipsoide al plano se hace directamente o por intermedio de un cono o un cilindro. En el primer caso se habla de proyecciones planas o perspectivas; en el segundo, de desarrollos.

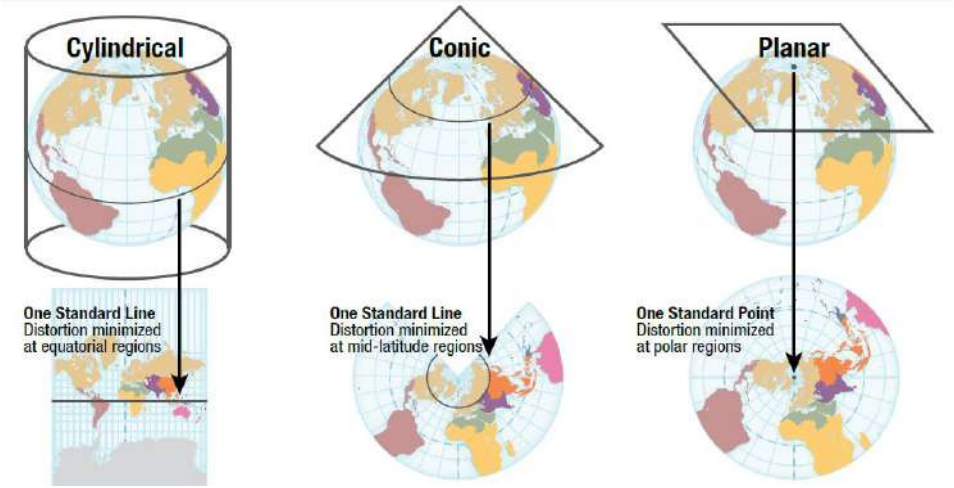


Ilustración 5 - 4: Clasificación geométrica de las proyecciones. Fuente: Mapping for a sustainable world. UN & ICA

Dentro de cada grupo se puede distinguir si se ha proyectado geométricamente según rectas paralelas o rectas concurrentes en un punto. También puede considerarse en el caso de las proyecciones sobre un plano, si éste es tangente o secante en un polo del elipsoide, en un punto de su ecuador o en un punto cualquiera. En los casos de proyecciones cónicas o cilíndricas se podrá distinguir si el cono o cilindro es tangente o secante a la superficie del elipsoide, como se muestra en la figura anterior. Análogamente, según la posición del eje del cono o del cilindro se les llama polares, transversas y oblicuas.

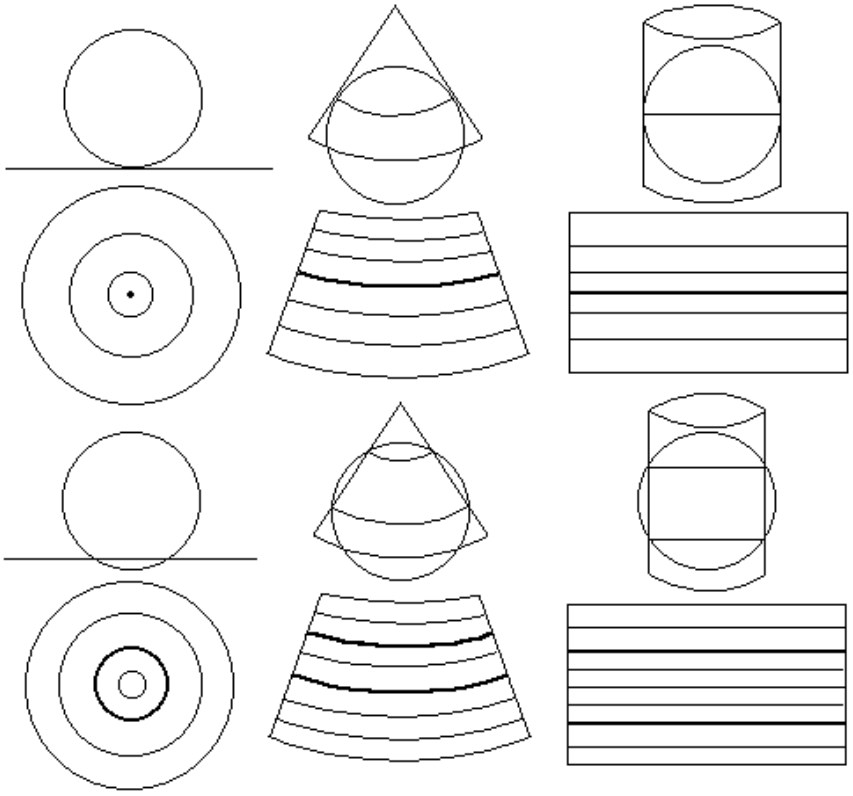


Ilustración 5 - 5: Proyecciones tangentes y secantes

En mapas topográficos se suelen utilizar proyecciones conformes, ya que los ángulos y las formas de la realidad se conservan al pasar al plano. En una gran cantidad de países se utiliza la proyección UTM (Universal Transverse Mercator), que divide la Tierra en 60 husos de 6º de longitud para evitar que se dispare el factor de escala, que en el meridiano central es de 0,9996 y va creciendo al alejarse del mismo. Los casquetes polares no se representan en UTM, quedando limitado el empleo de este sistema de representación a latitudes menores de 80°, como puede apreciarse la ilustración 5-6.

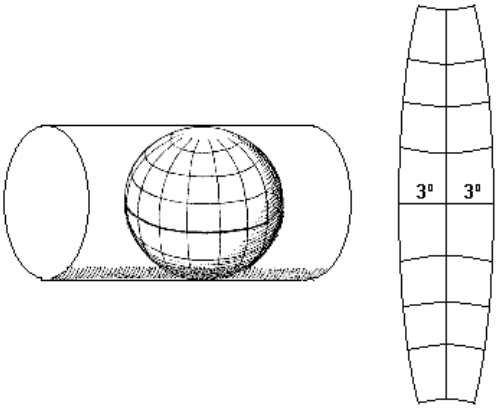


Ilustración 5 - 6: Ámbito territorial de cada huso U.T.M

En República Dominicana se utiliza UTM en el huso 19 (meridiano central, -69º), mientras que en España, la proyección cartográfica empleada en la serie del Mapa Topográfico Nacional 1:25,000 es la UTM, en los husos 28, 29, 30 y 31, según la zona.

Cuando se tiene una zona de la Tierra representada mediante una proyección cartográfica, cualquier punto de la proyección se puede localizar fácilmente por su longitud y latitud, pero como las fórmulas de la geometría plana son mucho más sencillas que las de la geometría esférica, resulta más sencillo localizar puntos en los mapas mediante un sistema de coordenadas cartesianas, entendiendo que estas coordenadas se aplican a la superficie plana de la proyección y no a la superficie sobre el elipsoide.

Cada proyección tiene su sistema de **coordenadas cartesianas: "X", "Y"** con un determinado origen, que varía según la zona representada. Por ejemplo, la proyección UTM, dentro de cada huso de 6° (que se numeran a partir del antimeridiano de Greenwich, en sentido de Oeste a Este), sitúa el origen de las ordenadas en el Ecuador, y el origen de las abscisas en una recta paralela al meridiano central del huso situada a 500 kilómetros al Oeste. De esta manera, para facilitar la localización de puntos, los mapas suelen llevar impresa la cuadrícula correspondiente a su sistema de representación cartográfico; se habla, así, de cuadrícula Lambert, o cuadrícula UTM.

Escala

Como el territorio a representar puede ser muy extenso y la carta debe tener un tamaño manejable, es imprescindible dibujar los objetos representados a una determinada escala. Este parámetro es justamente la relación matemática entre la distancia medida en el mapa y su correspondiente medida en el mundo real, es decir, la distancia correspondiente en la Tierra proyectada sobre una superficie de referencia, normalmente un elipsoide. La distancia en el mapa es considerada siempre como la unidad. La escala de un mapa puede expresarse de distintas formas, siendo las más comunes:

- **Escala numérica**: es una simple fracción o razón. Puede indicarse como 1:200,000 o 1/200,000. Ello significa que siguiendo líneas concretas 1 mm o 1 cm sobre el mapa representa 200,000 mm o 200,000 cm respectivamente sobre el terreno (proyectado sobre una superficie de referencia).
- **Escala gráfica lineal**: es una línea o barra situada en el mapa, a menudo en la carátula explicativa o en el margen de la hoja, que se ha subdividido para indicar las longitudes sobre el mapa de las unidades que se utilicen para medir distancias terrestres. Normalmente un extremo de esta barra está mucho más dividido, de modo que el usuario pueda medir las distancias con mayor precisión, como se muestra en la escala superior de la siguiente ilustración.

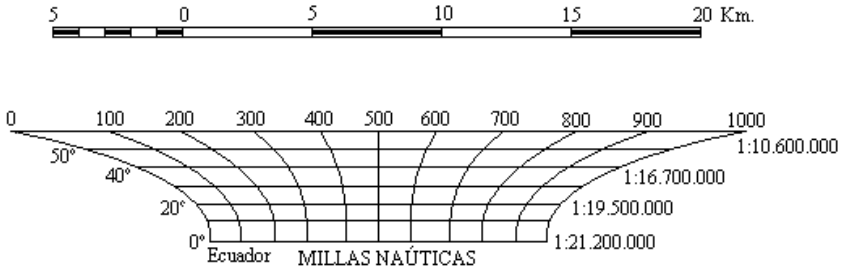


Ilustración 5 - 7: Dos ejemplos de escalas gráficas lineales

Pero, como debido a la representación plana de la superficie terrestre mediante una proyección cartográfica, la escala de un mapa no es la misma en todas las direcciones, a veces se representan, en determinados mapas, unas escalas gráficas de tipo múltiple, que, por ejemplo, van variando según la latitud que se considere, como se puede apreciar en la escala inferior de la anterior ilustración.

Al no ser posible transformar la superficie terrestre en un plano sin que se produzcan deformaciones, la escala numérica de un mapa sólo se ajusta a la realidad en puntos concretos o a lo largo de determinadas líneas; en los demás lugares la escala real del mapa será mayor o menor que la escala numérica. La expresión de la relación entre la escala real y la escala numérica en un punto se denomina **factor de escala**. En mapas a gran escala (cuanto más pequeño es el denominador de la escala numérica mayor es la escala del mapa) los factores de escala se aproximan mucho a la unidad, y varían muy poco de un lugar a otro. No obstante, en muchos mapas de gran escala se suele indicar en su carátula explicativa o en un margen el factor de escala para el centro de la hoja. Multiplicando el factor de escala de una determinada zona por la escala numérica del mapa se obtiene la escala real de esa zona.

Por tanto, la escala determina, con rigor geométrico, la resolución de los datos que van a ser representados, así como la cantidad de los mismos. Así, un mapa a escala 1:25,000 indica que cada milímetro medido en la carta equivale a 25,000 milímetros de terreno, es decir, 25 metros de la realidad. En un mapa de estas características un campo de futbol cuyas dimensiones fueran 100 metros x 50 metros, quedaría representado por un rectángulo exacto de 4 mm x 2mm. Por tanto, a la hora de hacer el mapa será necesario elegir adecuadamente la precisión de la información de la que se parte para que esté en consonancia con la escala del mapa, y generalizar si hay un exceso de información. Además, señalar que, aunque la tecnología y los soportes digitales en los que se encuentran los datos permitan generar un mapa a escala mayor a partir de un mapa de escala menor (ej. 1:10,000 a partir de 1:25,000), realmente no es correcto realizarlo, según la explicación anterior sobre las constricciones que la escala genera en los objetos a representar.

Generalmente, cuando se emplea esta escala como referencia para la representación topográfica oficial de un país, se genera lo que se denomina una **serie cartográfica** que estará compuesta por el número de hojas o mapas a dicha escala, necesario para la representación completa del territorio. A título de ejemplo, se puede mencionar que, en el caso de España, el Instituto Geográfico Nacional tiene asignada, por ley, la misión de actualizar el Mapa Topográfico Nacional a escala 1:25,000 de toda España (MTN25). Con esta serie cartográfica son necesarias algo más de 4,000 hojas para cubrir completamente el territorio peninsular; los archipiélagos balear y canario, y las ciudades autónomas de Ceuta y Melilla.

Representación del relieve

Las altitudes que figuran en los mapas están referidas al geoide y tienen como cota cero la del nivel medio del mar en un punto concreto. Cada país ha adoptado un determinado punto para fijar la superficie de referencia de las altitudes, siendo éste un dato que obligatoriamente tiene que figurar en la información marginal del mapa. A lo largo de la historia de la Cartografía se han empleado numerosos sistemas para representar el relieve del terreno (normales, tintas hipsométricas, sombreados, etc.), pero está universalmente admitido el sistema de puntos acotados con **curvas de nivel**. Son líneas que, a igual distancia vertical entre cada dos consecutivas, unen puntos de igual cota. La separación vertical existente entre dos curvas de

nivel consecutivas es la **equidistancia**, que se adoptará, para un determinado mapa, en función de su escala y de la importancia del relieve (a menor escala mayor equidistancia, a mayores pendientes mayor equidistancia, y viceversa). La equidistancia es un dato que deberá figurar en los márgenes del mapa.

Las curvas de nivel proporcionan información cuantitativa (indicación métrica de las altitudes) e información cualitativa (apreciación visual sobre las variaciones de pendiente y la morfología del terreno). Sin embargo, se requiere una cierta soltura en la lectura de mapas para obtener claramente de un vistazo la información cualitativa facilitada por las curvas de nivel, soltura que no se puede exigir a los potenciales usuarios no profesionales del mapa, que constituyen la inmensa mayoría. Por ello, es muy recomendable imprimir sobre las curvas de nivel un sombreado, que da una sensación muy plástica de las formas del relieve y suministra información intuitiva sobre las pendientes del terreno (no sobre su altura).

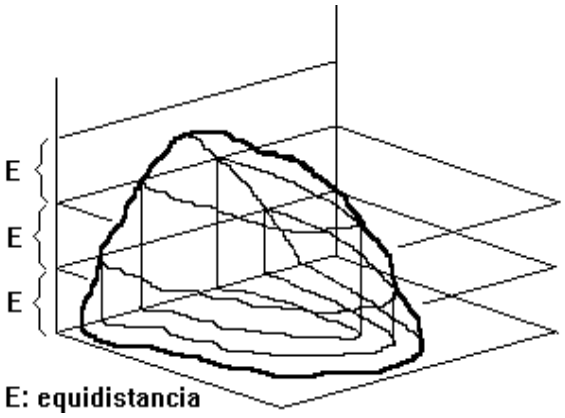


Ilustración 5 - 8: Gráfico conceptual de la equidistancia

A la representación numérica (digital) del territorio se la conoce como **Modelo Digital del Terreno (MDT)** pudiéndosele definir como el *conjunto de datos numéricos que describe la distribución espacial continua de una característica cuantitativa del territorio* (Doyle, 1978). Esta variable continua puede ser la altitud, la temperatura, el índice de precipitaciones, etc., denominándose **Modelo Digital de Elevaciones (MDE**) cuando representan la altitud.

En cualquier caso, la decisión de incluir un sombreado en un mapa siempre estará supeditada a su escala y propósito.

Declinación y convergencia

Por último, los mapas a gran escala también llevarán como información en uno de sus márgenes, o en la carátula, la indicación de la dirección del Norte Geográfico (dirección marcada por los meridianos del mapa), la dirección del Norte magnético (la dirección marcada por las brújulas) y la dirección del Norte de la Cuadrícula (la dirección de las ordenadas de la cuadrícula cartesiana de la proyección). Estas tres direcciones de norte formarán entre sí unos ángulos que variarán de unas zonas a otras. La dirección del Norte Geográfico y la del Norte Magnético, cuando no coinciden, forman un ángulo llamado **declinación magnética**, que deberá figurar en el mapa como dato para el centro de la hoja, en una determinada fecha. Y como la declinación magnética varía con el tiempo, habrá que incluir también el valor de su **variación anual**. El ángulo que forman la dirección del Norte Geográfico y el Norte de la Cuadrícula recibe el nombre de **convergencia de la cuadrícula**, y será otro dato que los mapas deberán llevar al margen.

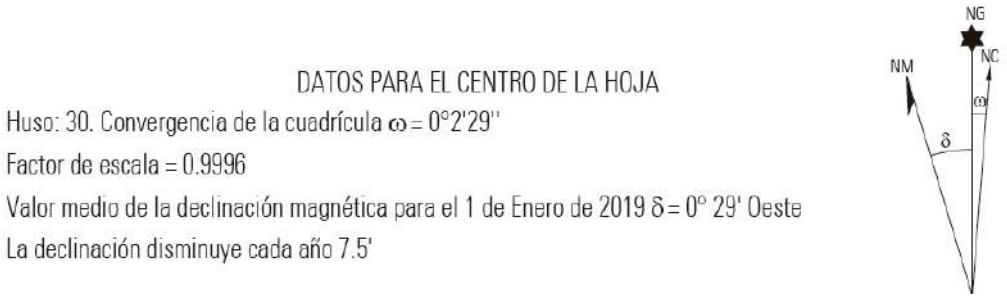


Ilustración 5 - 9: Valores de declinación y convergencia para una hoja del MTN25

5.1.2. Cartografía Topográfica

Un **mapa topográfico** puede ser definido como "*la cartografía que representa a escala, y mediante algún sistema de proyección, la superficie terrestre en sus tres dimensiones, incluyendo elementos tanto físicos como humanos que se encuentren sobre ella. El relieve queda representado mediante isolíneas o isopletas, que unen puntos con la misma altitud (curvas de nivel), o con distintos colores (tintas hipsométricas). El resto de elementos utiliza símbolos y colores convencionales que van descritos en la leyenda. También incluyen anotaciones toponímicas*"[1].

5.1.2.1. Contenido de un Mapa Topográfico 1:25,000

Una vez finalizadas las fases de captura, integración, almacenamiento y evaluación de la calidad de la información geográfica que se precisa para realizar la representación topográfica del territorio a escala 1:25,000 (fases explicadas en los capítulos 2, 3 y 4 de la guía), se está en disposición de generar la cartografía correspondiente.

Según la tipología de la representación del objeto geográfico.

En base a las primitivas geométricas que puede adoptar la representación de un objeto geográfico, el contenido del mapa se puede clasificar según la naturaleza de su simbolización en elementos:

- Puntuales: ej. pozo, depósito de agua, faro, piscina, etc.
- Lineales: ej. puente, río, vías de comunicación, línea eléctrica, gasoducto, etc.
- Superficiales: ej. áreas ocupadas por viñas, bosques, embalses, edificios, etc.

A la hora de representarlos será necesario especificar en cada caso los criterios generales de prioridad, distancias y orientación y establecer las normas de edición aplicables a la geometría o simbología del elemento (ej. en simbologías puntuales o signos convencionales especificar si son orientables y escalables o no) y las correspondientes al texto que pudieran tener asociado (ej. si el topónimo a representar admite que se puedan separar las palabras o las letras que lo definen o no es posible, etc.).

En el caso del MTN25 de España, las normas concretas de aplicación de la información que se recoge en esta serie se encuentran documentadas en las "Normas de edición del MTN25".

[1] Definición especificada en el **Glosario de términos geográficos**, elaborado por miembros del Grupo de Didáctica de la Geografía de la Asociación Española de Geografía (AGE) en colaboración con el Instituto Interuniversitario de Geografía de la Universidad de Alicante, el Instituto Geográfico Nacional y el Centro Nacional de Información Geográfica de España.

Según las temáticas que componen al mapa

Conforme a la propia definición de un mapa topográfico, se pueden deducir contenidos temáticos básicos o esenciales que todo mapa de estas características debe contener para la correcta representación del territorio, como son: la orografía, hidrografía, coberturas y usos del suelo, construcciones, transportes o las delimitaciones administrativas del territorio. A partir de este contenido estructural o de base, según el ámbito específico del planeta que se tenga que representar, se concretan los objetos que por cada una de estas temáticas es necesario dibujar e, incluso, se ajustan las temáticas añadiendo otras que, aunque a nivel global puedan resultar secundarias, en la zona de estudio en cuestión son especialmente relevantes. Por último, una vez definido el contenido topográfico del interior del mapa, es necesario especificar los componentes de los exteriores del mapa: marco, portada, leyenda, etc.

5.1.2.2. Contenidos del MTN25

A título de ejemplo, a continuación se ilustra esta sección con los contenidos temáticos, y otros auxiliares, que componen las hojas del MTN25 de España.

Orografía

La morfología del relieve y los valores de altitud se definen principalmente mediante las **curvas de nivel**, que consisten en la intersección del terreno con superficies paralelas al nivel medio del mar. Estas líneas son de color siena y se representan con una equidistancia de 10 m. Para facilitar la interpretación del relieve, se representa cada quinta curva de nivel con una línea más gruesa y además se le acota su valor de altura (curvas de nivel maestras). Cuando el terreno es muy plano y no quedan bien definidas mediante la equidistancia elegida, se recurre a las llamadas curvas de nivel auxiliares, que se trazan a la mitad del intervalo. En estas zonas también se incorporan los puntos acotados, al igual que en las cimas montañosas, collados, depresiones y en otras zonas destacables donde se precisa especificar el valor de altitud. Para facilitar la interpretación de las ondulaciones del terreno, se incorpora un **sombreado** digital que ilumina el mapa desde un foco teórico situado al noroeste del mismo.

Ilustración 5 - 10: Orografía de MTN25

Hidrografía

Compuesta por los elementos tanto naturales como artificiales de la red hidrográfica (incluidos componentes de drenaje y de almacenamiento de agua), es decir, ríos, arroyos, lagos, canales, embalses, etc. Se representan con tonos azules, de distintos grosores según su importancia y cuando existen láminas de agua que se extienden superficialmente, se rellenan con fondo azul.

Ilustración 5 - 11: Hidrografía de MTN25

Comunicaciones

Es la red de transporte constituida por carreteras, ferrocarriles, tendidos aéreos, tuberías, caminos, etc. Constituye el eje vertebrador del territorio. Para poder mostrar las diferentes categorías en las que se subdivide este tipo de información, se usan múltiples colores que facilitan la clasificación visual de estos elementos, que pueden abarcar desde autopistas hasta carriles-bici.

Ilustración 5 - 12: Comunicaciones de MTN25

Construcciones

Recoge las edificaciones ya sean urbanas, industriales, históricas o religiosas, así como el resto de obras humanas, puentes, túneles, acueductos, murallas, pasarelas, etc. En este tema predominan los tonos rojizos, a diferencia de la tendencia europea que lo suele representar con colores grises virando a negros.

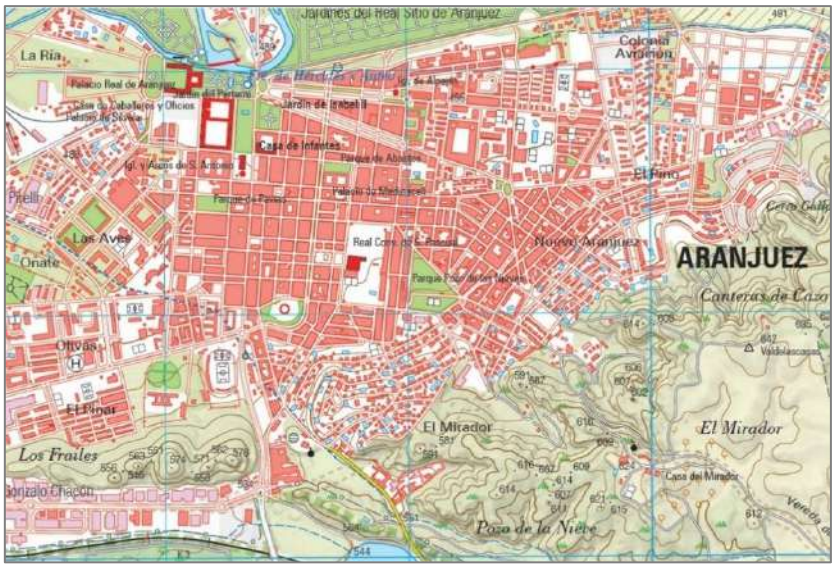


Ilustración 5 - 13: Construcciones de MTN25

Coberturas y Usos del suelo

Se trata de la categorización de la superficie terrestre en distintas unidades según sus propiedades biofísicas.

En los últimos años, gracias al desarrollo de la teledetección y al tratamiento digital de las imágenes, la captura de este tipo de información geográfica es hoy en día mucho más sencilla, rápida y eficaz, obteniéndose bases de datos de ocupación del suelo de prácticamente toda la superficie terrestre. En el caso de España, la información se recoge de **SIOSE,** el **Sistema de Información sobre Ocupación del Suelo de España**, cuyo objetivo es generar una base de datos de Ocupación del Suelo para toda España a escala 1:25,000, integrando la información disponible de las Comunidades Autónomas y la Administración General del Estado. El resultado es una exhaustiva clasificación donde se distinguen áreas de bosques, matorrales, cultivos agrícolas, pastizales, zonas sin vegetación, etc. En este apartado los colores predominantes son verdes y amarillos.

Ilustración 5 - 14: Usos del suelo de MTN25

Las temáticas descritas a continuación no son elementos naturales ni obras hechas por el hombre, pero deben ser consideradas en los mapas topográficos por ser muy relevantes:

Líneas límite administrativas

Son las líneas que delimitan municipios, provincias, comunidades y naciones limítrofes. Obedecen a geometrías acordadas entre municipios y provincias que son registradas de manera oficial en la base de datos del Sistema de Información Geográfica de Líneas Límite (SIGLIM). En el mapa se representan por líneas jerarquizadas de color negro o gris.

Ilustración 5 - 15: Representación de las líneas límite jurisdiccionales

Lugares o áreas de especial protección

Incluye la delimitación de las zonas protegidas o espacios naturales protegidos, es decir, los espacios del territorio que cumplan al menos uno de los requisitos siguientes y sean declarados como tales:

- contener sistemas o elementos naturales representativos, singulares, frágiles, amenazados o de especial interés ecológico, científico, paisajístico, geológico o educativo.
- estar dedicados especialmente a la protección y el mantenimiento de la diversidad biológica, de la geodiversidad y de los recursos naturales y culturales asociados.

En la serie del MTN25 se representan como líneas continuas y discontinuas de color verde.

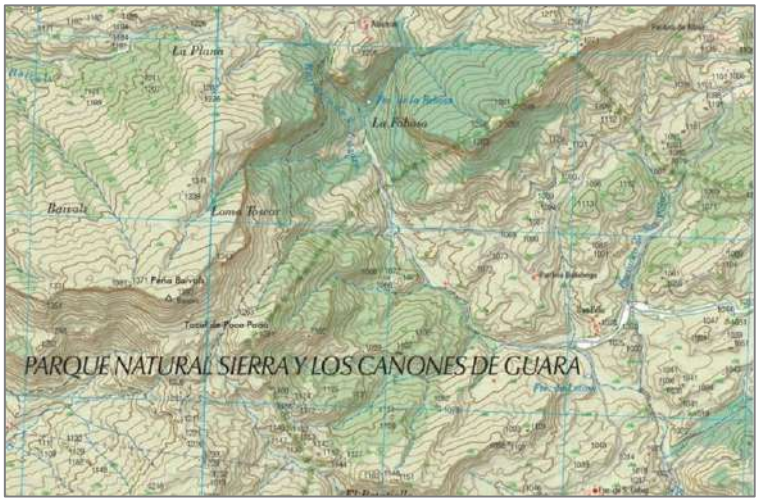


Ilustración 5 - 16: Representación de los límites de las zonas de especial protección

Vértices geodésicos

Incluye la información de las redes nacionales geodésicas y de nivelación mediante la representación de los vértices geodésicos, con su nombre y altitud asociada.

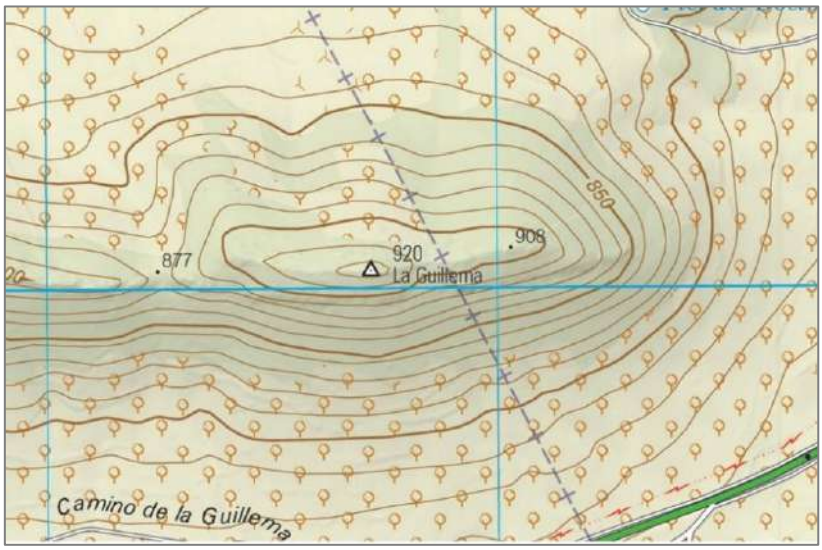


Ilustración 5 - 17: Representación de un vértice geodésico

Toponimia

Casi todos los accidentes geográficos tienen nombre propio, es una información recogida en encuestas de campo a lo largo del tiempo y que ha sido volcada en el Nomenclátor Geográfico Básico de España (NGBE). Cada topónimo está asociado a la posición que ocupa sobre la superficie terrestre y goza de atributos tales como la tipografía o el color de la letra, que dependerá del tema al que pertenece. El cuerpo del texto será tanto más grande, cuanto más destacado sea el elemento cartográfico al que nombra.

Ilustración 5 - 18: Toponimia y Líneas límite administrativas de MTN25

En definitiva, los **rótulos** que figuran en un mapa con su variabilidad de tipos de letra, colores y tamaños, se comportan como signos convencionales que dan una idea muy aproximada de la importancia del elemento geográfico al que nombran, ya sean ciudades, ríos, sitios, embalses, sierras, picos, etc.

Exteriores

Cualquier mapa topográfico necesita incorporar, además del contenido, una cantidad enorme de **metadatos** que deben ponerse en los exteriores del propio mapa. Estos son:

- Marco con coordenadas geográficas y cuadrícula de la proyección cartográfica.
- Portada con el logotipo de la serie cartográfica, el nombre y número de la hoja.
- Contraportada con una lista y un gráfico de los municipios incluidos y otro croquis con la ubicación de la hoja en España y la distribución de sus hojas colindantes.
- Sistema de referencia geodésico y proyección cartográfica, además del origen de altitudes.
- Tabla de signos convencionales de los elementos susceptibles de ser representados.
- Escala numérica y gráfica.
- Año de edición.

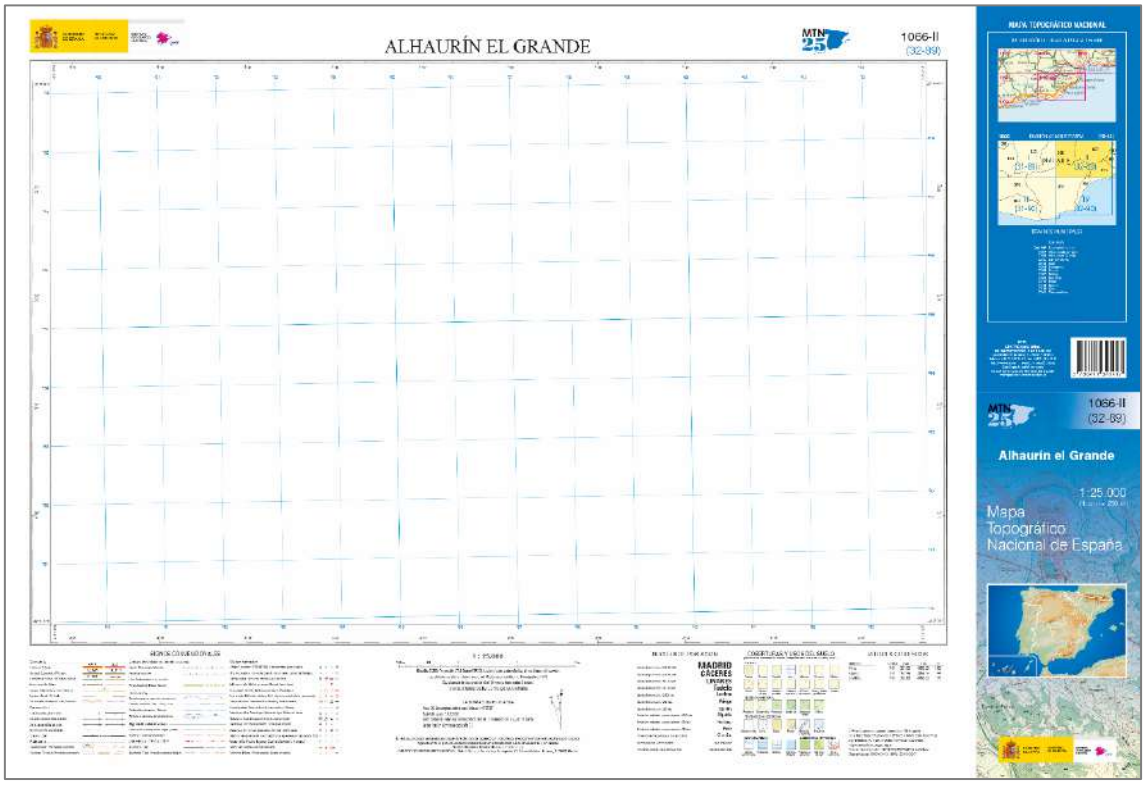


Ilustración 5 - 19: Exteriores de una hoja de MTN25

5.1.2.3. La simbología y la jerarquía de capas en la estructuración de contenidos

El mapa es un **medio de comunicación** para explicar cómo es el territorio. Debe combinar la calidad métrica en cuanto a la posición y tamaño de los elementos con la capacidad de

transmitir visualmente la información. La posición espacial que tienen los objetos geográficos es inamovible, porque responde a su ubicación en la superficie terrestre. Así, el cartógrafo dispone de cierto grado de libertad para diseñar la simbología empleada, que cuando es adecuada, contribuye sobremanera a la óptima interpretación de la información representada. Antes de simbolizar, la casuística de los elementos geométricos se reduce a cuatro clases básicas (puntos, líneas, superficies y rótulos), como ya se ha mencionado anteriormente. La habilidad del cartógrafo para combinar colores, grosores, rellenos y tamaños de los símbolos empleados en estas cuatro clases será trascendental para que el resultado final sea fácilmente legible por el usuario.

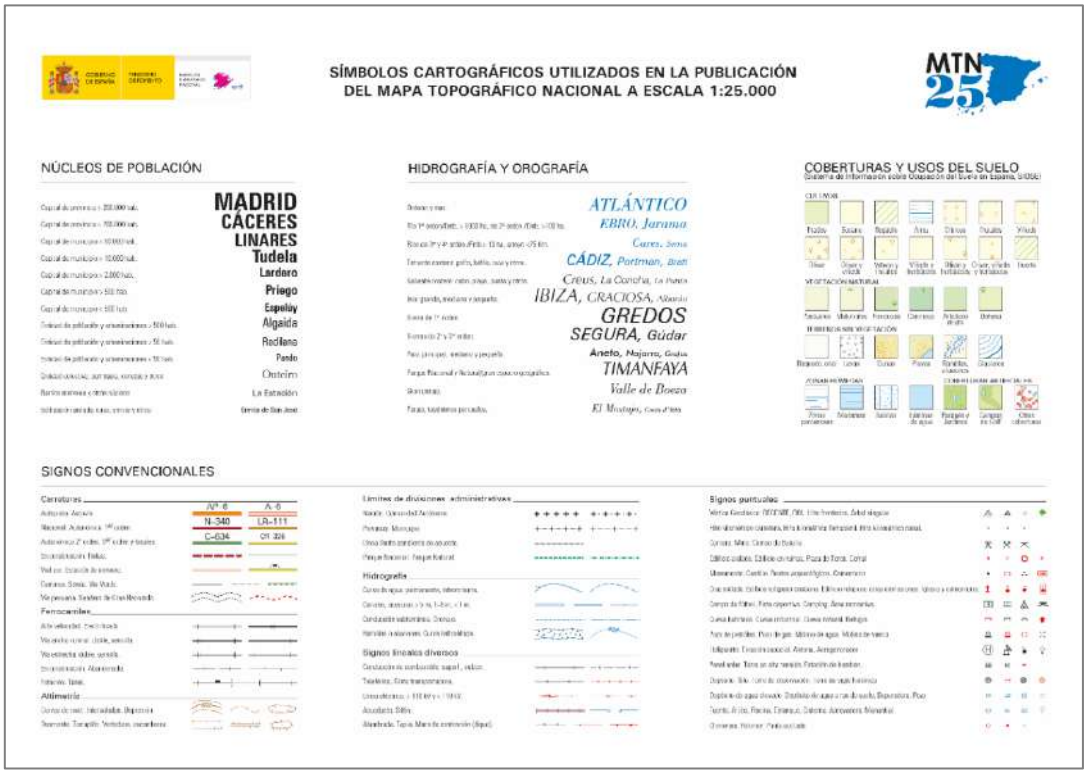


Ilustración 5 - 20: Cuadro de símbolos cartográficos de MTN25

Como los signos empleados son una convención del cartógrafo, los símbolos puntuales deben ser pictogramas evocadores de su significado, los lineales, deben combinar colores y grosores para expresar cualidades y jerarquías, y los símbolos que afectan a superficies se pueden rellenar con un fondo de color al que, en ocasiones, acompaña un patrón repetitivo que define el tipo concreto de uso de suelo.

GEOMETRIA	ATRIBUCIÓN SEMÁNTICO	SIMBOLIZACIÓN
Puntual	Campo de futbol	
Líneal	Autopista	
Superficial	Olivar y viñedo	
Texto	Nación limítrofe	FRANCIA

Ilustración 5 - 21: Fragmento de tabla de simbolización por atribución semántica

En definitiva, el cartógrafo necesita deconstruir el paisaje seleccionando solamente los temas que quiere representar, para después volverlo a reconstruir, empleando un dibujo que sintetice el territorio de una forma comprensible, sencilla y agradable a la vista. La armonía en la combinación de los símbolos convencionales es fundamental para que el lector del mapa recree en su imaginación el paisaje geográfico cartografiado de una manera efectiva.

Los elementos que aparecen en el mapa se distribuyen por capas. Estas capas tienen un orden de prioridad visual establecido para que los objetos geográficos no se tapen entre sí. Es muy razonable que todos los rótulos tengan mayor prioridad visual que el resto de los elementos, al objeto de poder ser leídos sin dificultad.

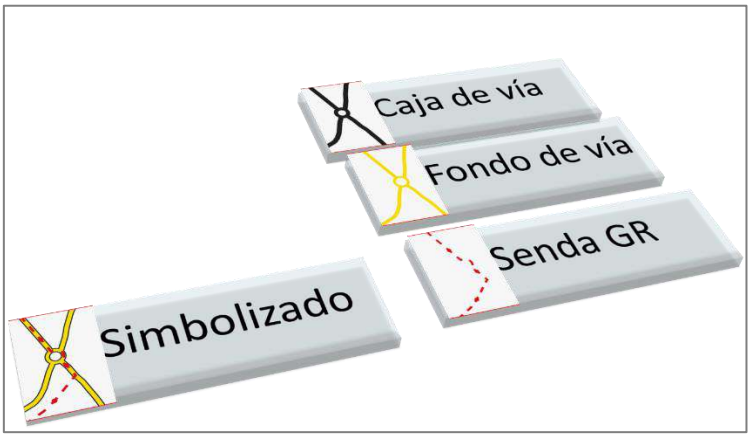


Ilustración 5 - 22: Ejemplo de estructura ordenada de capas para su correcta visualización

En el ejemplo de la figura, hay varios elementos geográficos superpuestos, que deben ser dibujados de manera que no ofrezcan dudas de interpretación al usuario.

El caso de una carretera local se compone de dos capas de idéntica geometría, pero distinta simbolización. La capa de la caja es negra, más ancha y con prioridad más baja mientras que la capa del fondo es amarilla, más estrecha y con prioridad más alta.

Si además una senda de gran recorrido coincide con la carretera en algún tramo, su color rojo y con mayor prioridad visual se superpondrá a la vía, resultando finalmente simbolizada en el mapa de la forma que se muestra en la figura.

Por lo tanto, es imprescindible definir una estructura ordenada de capas, cuyas prioridades visuales estén predeterminadas, en las cuales se depositarán los distintos elementos del mapa.

5.1.2.4. El mapa topográfico comparado con la mera fotografía aérea

Al final, la agregación de todos los temas reseñados con anterioridad da lugar al mapa topográfico publicado. Mientras que la fotografía aérea es una instantánea objetiva del medio, que plasma la realidad tal como es, el mapa resulta de un proceso de comunicación entre el autor y el usuario. El cartógrafo interpreta lo que ve en la foto y añade otra información no fotointerpretable, seleccionando el subconjunto de la realidad que constituye el universo que quiere cartografiar. Luego lo representa con símbolos que facilitan la comprensión del paisaje.

Además, aporta rótulos con topónimos que identifican unívocamente lugares y que, al ubicarse de forma discrecional sobre el mapa, contribuyen a la orientación del lector. Cuando

un usuario que busca desorientado en un mapa, lee un topónimo conocido, automáticamente se ubica y puede empezar a reconocer la zona representada.

En conclusión, el mapa –interpretado, simbolizado y rotulado– aporta un valor añadido que no tiene la foto aérea y, por tanto, comunica mejor el territorio, como se puede apreciar en la siguiente imagen.

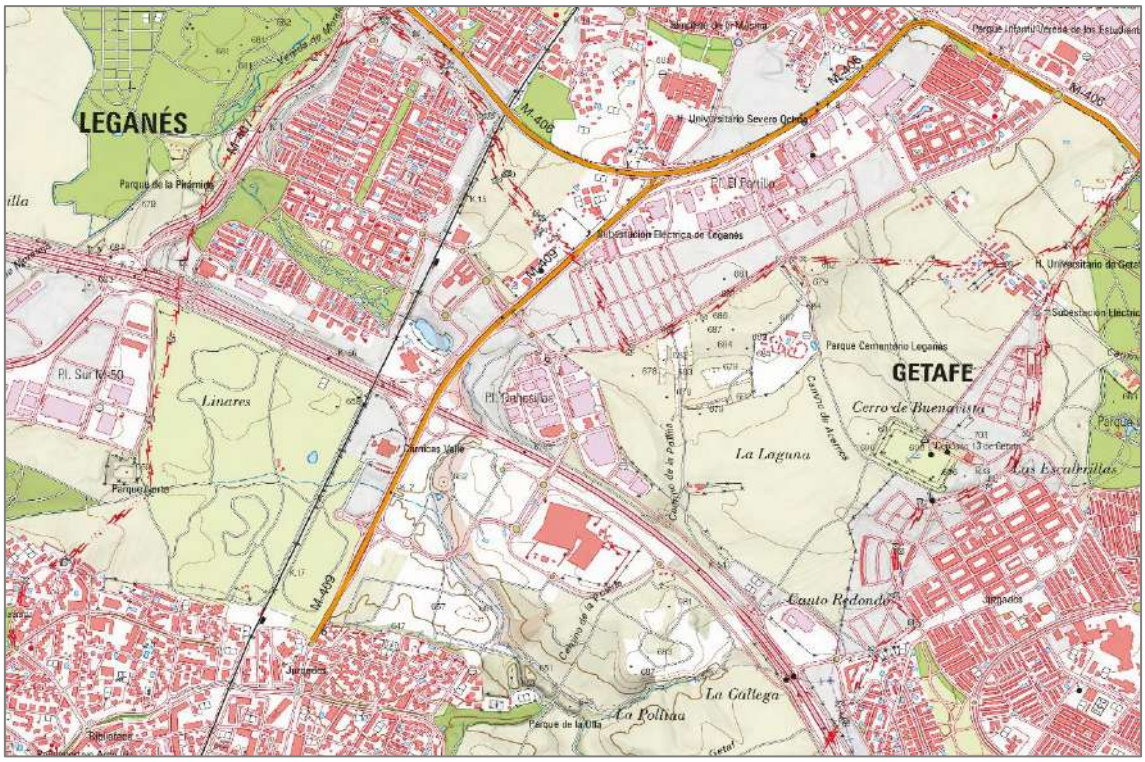


Ilustración 5 - 23: Detalle del Mapa Topográfico Nacional a escala 1:25,000 (MTN25)

El poeta Paul Valéry decía que las tres grandes creaciones de la inteligencia humana son la poesía, las matemáticas y el dibujo. En esos ámbitos se desenvuelve la cartografía, porque aspira a recrear de una manera poética el territorio, empleando las matemáticas con absoluta precisión geométrica y trazando un dibujo con la finalidad de expresar, de forma sencilla, la complejidad del mundo real.

5.2. Producción automática de cartografía: Mapa Automático

La producción de cartografía ha experimentado una fuerte evolución gracias al desarrollo de nuevas tecnologías, tanto en la captura directa e indirecta de los objetos geográficos como en la edición y redacción cartográficas y en la difusión de los productos. Esta evolución precisa, entre otros factores, de la trasmisión de conocimientos de unas generaciones a otras a través de la correcta documentación que será la base sobre la que se optimicen e implementen los procesos automáticos de producción.

Esta sección recopila la experiencia del IGN de España en cuanto a la evolución del proceso de producción cartográfico del MTN25, desarrollado a partir de la documentación correspondiente a la metodología de generación tradicional del mapa y la implementación de los automatismos que simulan dichas pautas.

5.2.1. Antecedentes: documentación del proceso tradicional

En la evolución de la producción del MTN, de la toma directa de datos en campo con instrumentación y métodos de topografía clásica se pasó a la toma de fotogramas y posterior restitución, primero analógica, después digital, terminando con la rectificación y generación de ortofotogramas, modelos digitales del terreno, y datos LiDAR, todo ello cada vez con mayor resolución espacial y frecuencia temporal.

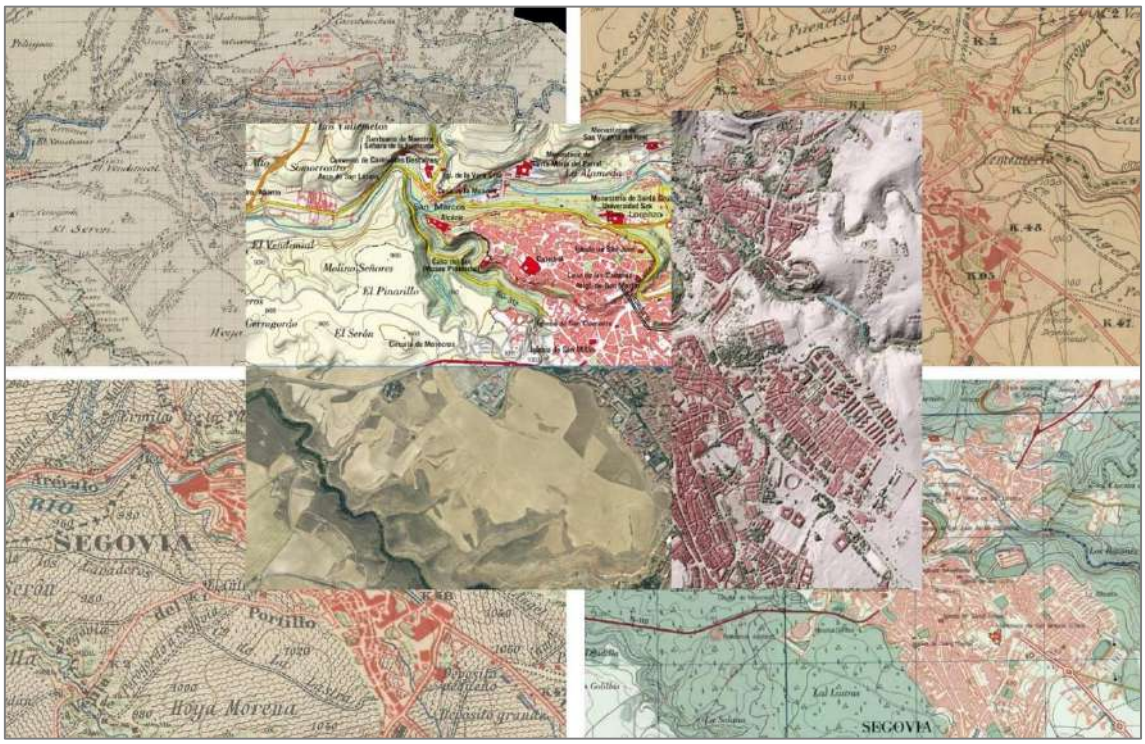


Ilustración 5 - 24: Composición de diferentes ediciones de MTN, Ortofoto y LiDAR

La captura, formación y edición de las hojas se realizaba directamente en el mapa. La edición y redacción cartográfica es un proceso muy costoso que requiere de personal altamente especializado porque tiene una componente estética destacada. En el IGN de España este meticuloso trabajo era realizado, primero con métodos analógicos y más tarde con procesos digitales, por operadores cartográficos con un alto conocimiento en edición cartográfica.

Adquirir estos conocimientos requería un gran proceso de aprendizaje, que solo se lograba a través de muchos años de experiencia en este trabajo, de modo que conforme la plantilla se iba reduciendo se corría el riesgo de perder el conocimiento. Con el objeto de afianzar la trasmisión sistemática de conocimientos entre operadores expertos y aprendices, se redactaron las **Normas de edición del MTN25**, resultado de la recopilación y formalización de los conocimientos del operador sobre la revisión de mapas.

Con este documento se pretendía alcanzar dos objetivos principales:

- recoger en un único documento la experiencia y conocimiento de los expertos en la edición cartográfica, así como la información contenida en las diferentes fuentes bibliográficas que se utilizan para la edición de esta cartografía.
- establecer una serie de normas para la edición de cada uno de los elementos del mapa y para la edición del mapa en su conjunto, de forma clara y homogénea, que facilitasen el aprendizaje y que sirvieran de guía para realizar el proceso de edición. La normalización iba a permitir la mejora de los procesos de automatización de la edición y del control de calidad posterior.

Este documento está dividido en tres bloques; en el primero se describen los criterios de edición de las simbologías contenidas en el mapa, en el segundo los criterios de edición de textos y en el tercero se incluyen las normas de edición de la parte comprendida entre el marco interior y el exterior, y de los exteriores del mapa: cabecera, contraportada, portada, leyenda y vocabulario.

Gracias a la documentación de estas normas el producto queda perfectamente definido, tanto que a partir de los criterios establecidos en ellas ha sido posible diseñar los procesos para la generación automática del mapa, que es el procedimiento vigente en la producción cartográfica de esta serie.

Otro factor de vital importancia a destacar es la toponimia. Teniendo en cuenta que es parte indispensable del mapa, para facilitar su comprensión y para la localización de los lugares, también se estimó necesario el redactar unas **Normas para la toponimia en el MTN25**.

El objeto de esta publicación era, por un lado, establecer unas normas para el tratamiento de la toponimia en el MTN25, que proporcionen igualdad de criterios en toda la serie cartográfica y, por otro lado, dar unas nociones básicas sobre los nombres geográficos y la toponimia (tipos de topónimos, la importancia de la toponimia desde una perspectiva geográfica, cultural, etc.), así como presentar una selección de nombres comunes que forman parte de la toponimia en lengua española, con indicación de su etimología (términos relacionados con actividades agrícolas, elevaciones de terreno, cursos de agua, etc.).

En España, la normalización y el uso homogéneo de los topónimos resultan especialmente complejos por su carácter plurilingüe, ya que, además del castellano como lengua oficial para todo su conjunto, hay otras lenguas oficiales en determinadas Comunidades Autónomas, así como dialectos oficialmente reconocidos. Pero eso no significa que el plurilingüismo sea el único problema de la toponimia, ya que la normalización también es esencial en las zonas monolingües.

Los criterios recogidos en este documento están inspirados en la filosofía de las recomendaciones de las Naciones Unidas sobre la Normalización de los Nombres Geográficos,

en las que se define la normalización internacional como «la actividad encaminada a lograr una forma escrita única de cada nombre geográfico de la Tierra» y se subraya «que los nombres locales normalizados se utilicen en los mapas» (Resolución 31 de la 2.a Conferencia).

Se divide en dos capítulos, tratando el primero sobre los criterios generales de la toponimia y en el segundo, se dan instrucciones para la rotulación de la toponimia en el MTN25, se clasifican las entidades o elementos geográficos rotulados en el MTN25, definiendo los que se incluyen en cada clase y subclase. En todos ellos se indica el tratamiento de los topónimos de acuerdo con los criterios generales y las fuentes de información para obtener las denominaciones correctas de los topónimos, así como los criterios objetivos (número de habitantes, longitud, extensión) o relativos (principales, grandes, menores, pequeños) para la clasificación jerárquica de los elementos, añadiendo ejemplos de topónimos para una mejor comprensión.

De forma equivalente a como ocurría con las Normas de Edición, la documentación de estas Normas de Toponimia han sido la base para la implementación de las tareas de rotulación en el vigente proceso automático de generación del mapa.

Hoy en día, tras 143 años de tradición cartográfica, los rápidos canales de difusión de la información geográfica junto con la alta demanda actual por parte de los usuarios, en cuanto a precisión de la información y frecuencia de su actualización, han obligado a reformular toda la metodología de trabajo, orientándola hacia métodos mucho más eficientes, basados en la automatización de los procesos de la cadena de producción. Estos procesos, implementados en base a las directrices recogidas los dos documentos de normas explicados en este apartado, comenzaron a desarrollarse en 2013 y actualmente ya son una solución sólida que permite la generación completa de la serie completa MTN25, compuesta por más de 4,000 hojas, en el periodo de un año, frente a los 25 años que se tardó en generar la primera versión completa por métodos tradicionales.

En los siguientes apartados del capítulo se explica con más detalle la motivación de la evolución de la metodología de producción, además de la descripción de los principales componentes y operaciones del proceso automático.

5.2.2. Del mapa tradicional al mapa automático. Motivación

Debido a la evolución de los procesos de producción de cartografía, el MTN25 ha pasado de tener un papel protagonista en la producción de datos, sobre el que se desarrollaban las diferentes fases de captura, formación y edición de la información geográfica, a convertirse en un producto derivado de las Bases de Datos Geográficas. El principal motivo es porque éstas permiten integrar más fácilmente las fuentes datos, a la vez que obtener productos derivados de mayor calidad geométrica y carga semántica. Como consecuencia, ahora la producción del MTN25 ha de ajustarse lo máximo posible a la planificación de las actualizaciones de estas Bases de Datos Geográficas.

Como muestra la siguiente figura, el proceso de producción ha pasado de generar Bases Cartográficas a partir de MTN25, a generar esta serie a partir de Bases Topográficas producidas de manera independiente.

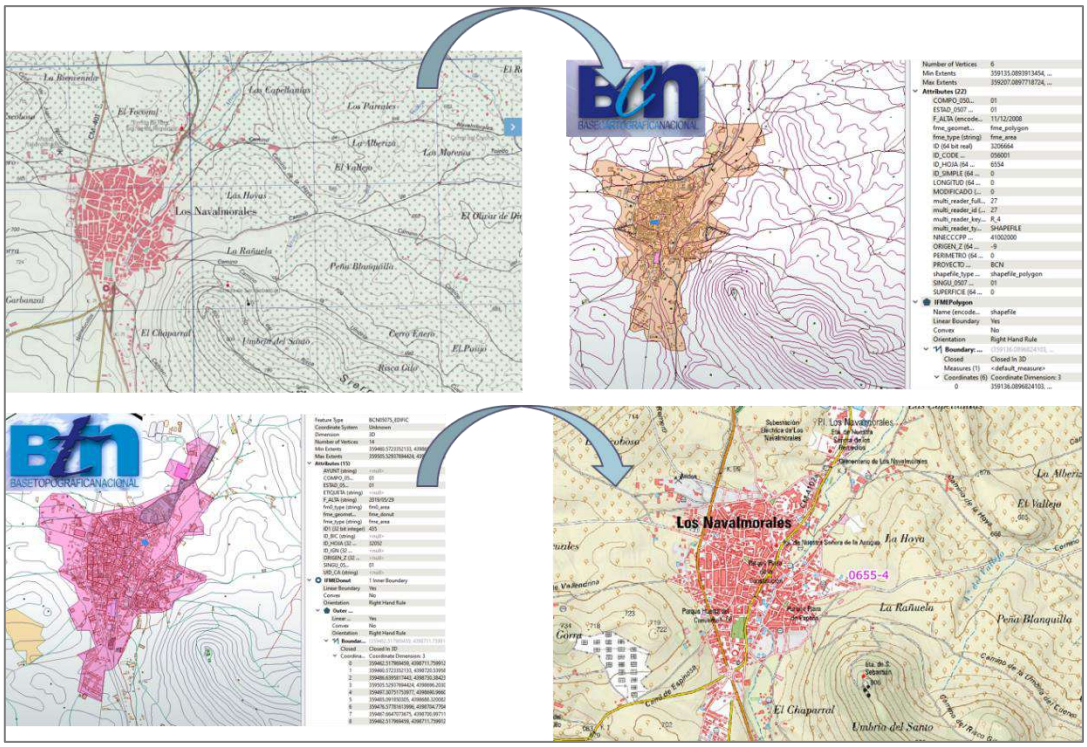


Ilustración 5 - 25: Cambio de paradigma en la producción de cartografía.

El ámbito de actuación del MTN25 tradicionalmente ha sido la hoja. En los nuevos procesos de producción, éstos son diversos: desde la producción por unidades territoriales (provincia, nación, etc.) a capas completas de información (comunicaciones, construcciones). Estas actuaciones obligan a actualizar un gran número de hojas que se ven afectadas.

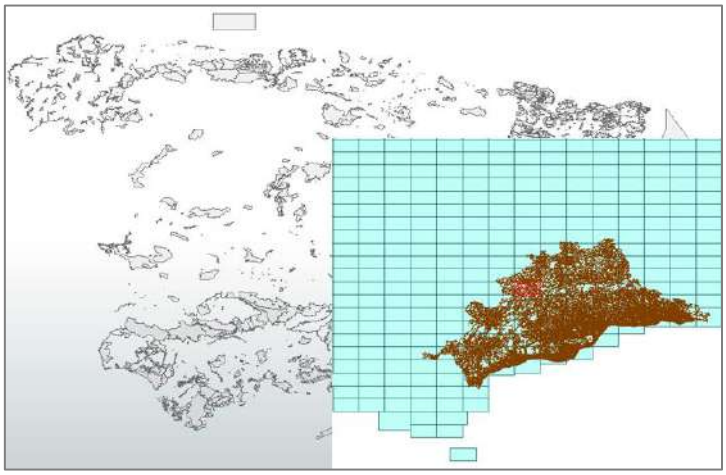


Ilustración 5 - 26: Diferentes ámbitos de actuación: por hojas, por provincias, por capas.

La difusión de los productos cartográficos también ha experimentado una gran evolución. La demanda de publicaciones impresas disminuye en favor de la demanda de acceso a esta información a través de servicios de visualización (WMS/WMTS) y descargas de archivos digitales. Utilizando estos servicios, la información cartográfica convive con la información geográfica procedente de las bases de datos, imágenes aéreas y espaciales, por lo que es importante que el tiempo de actualización sea lo más sincronizado posible.

Ilustración 5 - 27: Visualizador de cartografía e imágenes

Por otro lado, la edición y la redacción cartográfica son procesos muy costosos que requieren de personal altamente especializado porque tienen una componente estética muy fuerte, y los recursos, tanto humanos como materiales, se reducen y, en consecuencia, el volumen de hojas que pueden publicarse es inferior en comparación con la difusión de otros conjuntos de datos.

Como consecuencia de todos estos cambios, se concluyó en la necesidad de realizar un profundo cambio en el modelo de producción.

Por otro lado, los avances tecnológicos presentan una oportunidad de llevar a cabo esta transformación, orientada en gran medida a la automatización de los procesos de edición, redacción y publicación de la cartografía.

La automatización de los procesos de edición intenta emular el trabajo del especialista cartográfico, aplicando una serie de reglas de edición a través de una herramienta informática. De esta primera afirmación se deduce que la componente artística de la redacción cartográfica queda eliminada, con la consiguiente pérdida de calidad con respecto a la edición por métodos clásicos.

Como ya se indicó al inicio de esta sección, el proceso de automatización debe sustentarse sobre unas **normas de edición claras, objetivas y parametrizadas**, correctamente documentadas. De este conjunto de normas, se seleccionarán aquellas susceptibles de ser automatizadas, es decir, que cuenten con unos parámetros medibles y que se puedan traducir a un algoritmo que sea capaz de aplicar dicha norma durante el proceso de edición. Esto nos lleva a pensar que parte de esas normas no cumplirán ese requerimiento, o que el algoritmo no funcionará de manera eficaz, por lo que en este aspecto el producto cartográfico sufrirá también pérdida de su calidad. Para minimizar este aspecto, es importante establecer protocolos de revisión y de gestión de incidencias que proporcionen retroalimentación al desarrollador para conseguir así una mejora continua del proceso y, en consecuencia, un producto con suficiente calidad como para ser publicado.

En conclusión, difícilmente un proceso automático puede llegar a generar un producto cartográfico con la misma calidad con la que se obtiene mediante la edición tradicional (manual, fundamentalmente), debido a la fuerte componente estética que requiere un mapa. Sin embargo, también se puede afirmar que, la ventaja que supone la posibilidad de obtener

de manera casi inmediata la información actualizada procedente de las bases de datos geográficas, puede llegar a compensar dicha pérdida de calidad estética.

Ilustración 5 - 28: Aplicación de reglas de edición para generar el Mapa

A continuación, se explican los puntos principales de la generación del contenido del mapa, sin hacer referencia a la información marginal o exteriores de la hoja, que se tratará en secciones posteriores.

5.2.3. Fuentes origen

Para la generación del mapa por procesos automáticos necesitaremos partir de datos geográficos que hayan sido capturados previamente y almacenados en un formato compatible con el proceso que se vaya a utilizar. Generalmente, estos datos de partida estarán almacenados en una o varias bases de datos geográficas.

5.2.3.1. Selección de las fuentes de datos

Para seleccionar las bases de datos adecuadas se han de tener en cuenta diferentes factores:

- **La escala del mapa**: determina la resolución, tanto geométrica como semántica, de los datos de partida. La resolución ha de ser al menos la correspondiente a la escala del producto cartográfico. Si ésta es mayor (más detalle), será necesaria una generalización. Por tanto, cuanto más se acerque la resolución de los datos a la escala del mapa, se requerirá menos trabajo de generalización y, por tanto, menos carga en el proceso.
- **El ámbito de actuación**: la fuente de datos a escoger tendrá que ocupar todo el ámbito de actuación de la unidad de producción. Puede ocurrir que una hoja de mapa ocupe varias unidades territoriales, cuyos datos estén almacenados en distintas bases de datos.
- **El modelo de datos**: debe ser compatible con el modelo de mapa. Se necesitarán atributos para clasificar los objetos geográficos de acuerdo con la leyenda definida en el mapa. También es importante que los datos tengan la mayor jerarquización posible de manera que se faciliten al máximo los procesos de generalización en función de dicha jerarquía.
- **El Sistema Geodésico de Referencia**: hay que tenerlo en cuenta para reproyectar los datos en caso de que se encuentren en un sistema de referencia diferente al del mapa que se quiere generar.

- **La calidad de los datos**: para que los procesos automáticos funcionen correctamente, es imprescindible que los datos tengan una calidad suficiente. Es muy importante garantizar la consistencia de las fuentes de datos de partida. La automatización requiere partir de fuentes de datos armonizadas, con relaciones semánticas y topológicas conocidas y aseguradas. Para ello, es necesario un control de calidad exhaustivo de esas fuentes.

Un proyecto cartográfico de envergadura necesitará coordinar todas las fases de producción de manera que estén relacionadas entre sí, y tengan en cuenta las necesidades de cada una de ellas a la hora de modelizar y diseñar los diferentes productos.

En el caso del IGN de España, la serie del MTN25 se obtiene de diferentes bases de datos producidas por este organismo. La principal fuente de datos es la Base Topográfica Nacional (BTN), que a su vez integra la información de la Red de Transportes, que se produce en otra base de datos. Ambas bases de datos han adaptado sus modelos de datos a las necesidades de los productos que se derivan de ellas, entre otras, las del MTN25. Otras fuentes de datos proceden de la base de datos de líneas límite jurisdiccionales (SIGLIM), de la red geodésica y del Sistema de Información de Ocupación del Suelo en España (SIOSE).

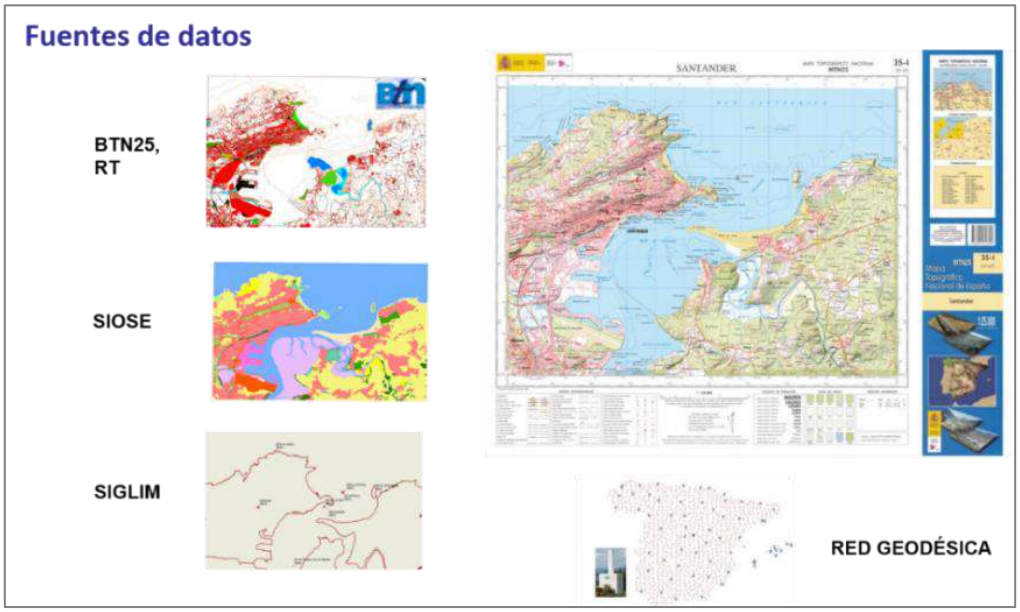


Ilustración 5 - 29: Fuentes de referencia para el MTN25

5.2.3.2. Extracción por unidades de producción

Una vez seleccionadas las fuentes de partida, hemos de extraer la información necesaria para generar el mapa. Hay que tener en cuenta que se va a realizar una serie de tratamientos automáticos, tanto de atributos como de geometrías, y si el conjunto de datos es muy grande, los procesos automáticos pueden ralentizarse y consumir demasiados recursos. Por otro lado, si se considera que la unidad de producción es la hoja de Mapa, lo más conveniente sería extraer la información de cada hoja y procesar cada una de manera independiente.

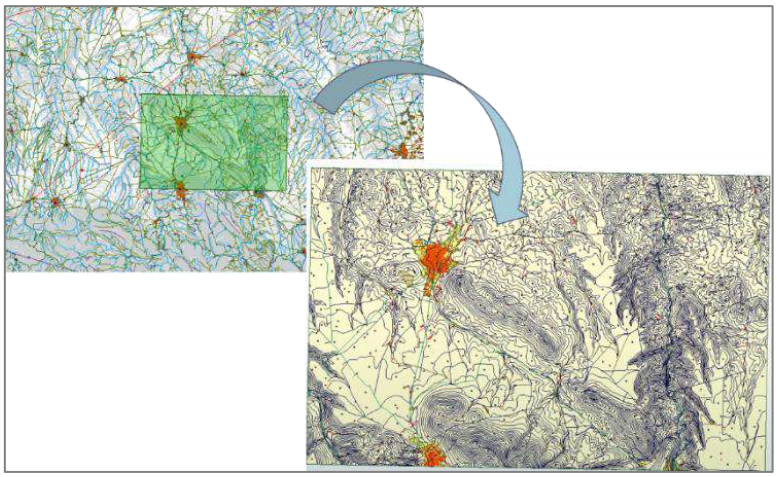


Ilustración 5 - 30: Extracción desde una base de datos continua, por hoja de MTN25

Esta extracción, realizada sobre una base de datos que se supone continua, puede efectuarse de dos formas: 1) cortando directamente las geometrías por el marco de cada hoja. 2) seleccionando los elementos que intersecten con el marco, pero sin cortar las geometrías. Este último modo de extracción conserva las geometrías originales, aunque excedan del ámbito de la hoja. Una vez se hayan efectuado los procesos necesarios para la generación del mapa, se podrían cortar definitivamente esas geometrías. La razón de no cortarlas al principio es para conservar las características de la geometría original, como longitud, superficie, etc., que pueden ser útiles a la hora de realizar procesos de clasificación, asignación de jerarquías, etc.

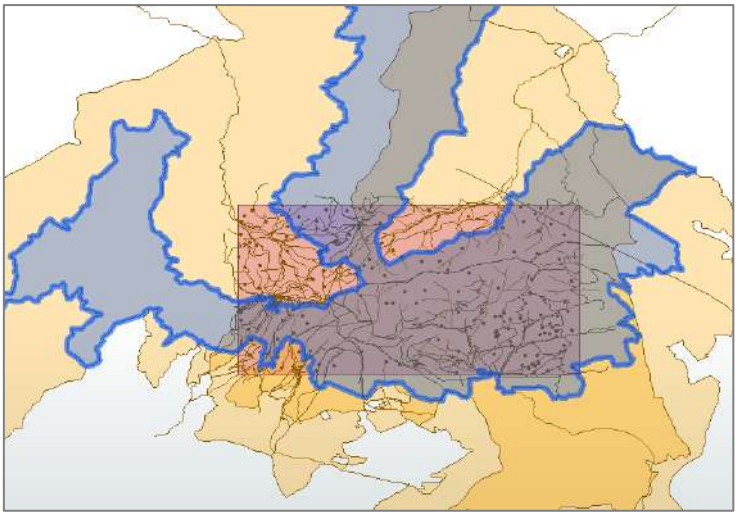


Ilustración 5 - 31: Extracción de elementos superficiales por intersección con hoja de MTN25

5.2.3.3. Traducción al modelo cartográfico

La traducción del modelo, o mapeo de los datos, consiste en establecer una correspondencia entre dos modelos de datos, en este caso serán los pertenecientes a la base de datos origen y al mapa. Para ello debemos conocer la fuente de origen, identificar los distintos objetos geográficos en base a su modelo de datos, y hacerlos corresponder con las entidades del mapa, clasificadas por la leyenda.

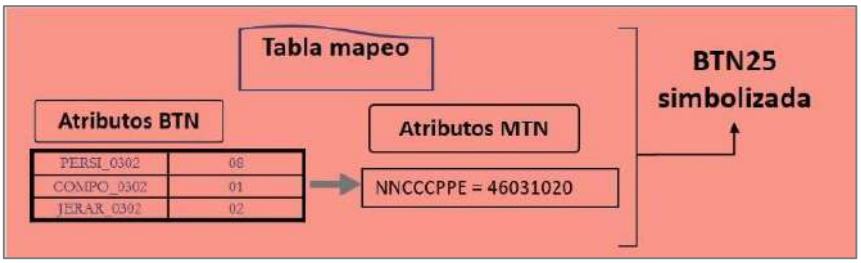


Ilustración 5 - 32: Esquema del proceso de traducción del modeloBTN-MTN25

Hay modelos más sencillos que tienen una correspondencia más directa con el mapa y otros más complejos que requieren una serie de operaciones con los atributos como, por ejemplo, la combinación de varios atributos para identificar las entidades geográficas con una relación unívoca con los elementos del mapa.

En la siguiente ilustración se representa el caso de la base de datos de líneas límite jurisdiccionales del IGN, donde el atributo 'Tipo' establece una correspondencia directa con la leyenda del MTN25.

Ilustración 5 - 23: Correspondencia directa entre modelo de líneas límite jurisdiccionales y leyenda de Mapa.

En el caso de la tabla de carreteras de la BTN del IGN, es necesaria la combinación de varios atributos para identificar el objeto geográfico y darle correspondencia con la leyenda del MTN25. En la tabla utilizada para el mapeo se concatenan: el nombre de la tabla, la clave de cada atributo y el valor; el resultado se almacena en el campo 'COMBINACIONES'. Cada una de esas combinaciones tendrá una correspondencia con el código utilizado en el mapa, en este caso almacenado en el campo 'VALIDO_CTR'.

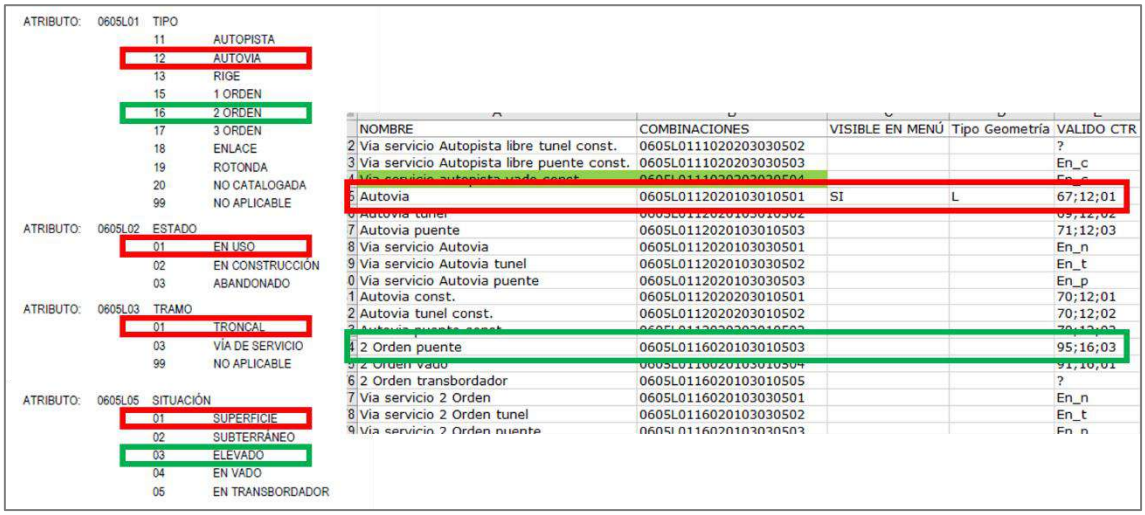


Ilustración 5 - 33: Correspondencia a través de combinaciones de atributos.

En la tabla de mapeo se puede incluir otro tipo de atributos que ayuden a los procesos automáticos como, por ejemplo, si el elemento puede colapsar a símbolo puntual, si es desplazable, si se le asigna jerarquía, etc.

NOMBRE	COMBINACIONES	VISIBLE EN MENÚ	Tipo Geometría	VALIDO CTR	Colapso CTR	Sups Desplazables	Simbs POSTERIORI	Jerarquia	Clave_ETIQUETA
Depósito agua en superficie	0331S01020501			357	357	sl		1	140
Depósito agua en superficie ruinas	0331S01020503			216	340	sl		1	140
Depósito agua elevado	0331S01030501			356	356	sl		1	140

Ilustración 5 - 34: Traducción del modelo. Atributos asignados a cada combinación

En la propia traducción del modelo, ya se pueden aplicar generalizaciones de tipo semántico, pues determinadas combinaciones de atributos identifican objetos que no tendrán representación en el mapa y, por tanto, se descartan en el proceso de mapeo.

También se pueden asignar directamente códigos de mapa a objetos de la fuente de datos mediante análisis espacial, según las necesidades de representación. Por ejemplo, como ilustra la figura 5-35, si dentro de un elemento de la tabla 'fortaleza' hay un edificio, se le asignará a ese edificio el código de castillo en mapa.

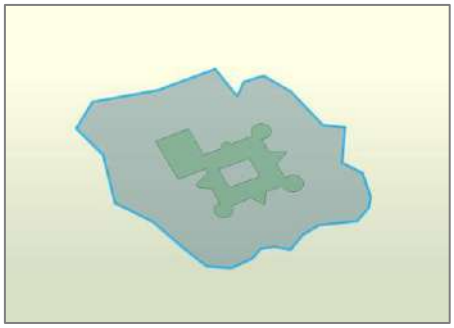


Ilustración 5 - 35: Traducción del modelo por análisis espacial.

5.2.4. Procesos automáticos de redacción cartográfica

Este apartado, tras introducir el concepto de redacción cartográfica, describe las bases en las que se fundamenta el proceso de su automatización, explicando posteriormente algunos de los operadores utilizados en dicho proceso. Como se explica más adelante, estos conceptos forman el soporte de partida sobre el que diseñar la implementación del flujo de este proceso automático.

5.2.4.1. Redacción cartográfica: generalización y edición cartográfica

La **redacción cartográfica** comprende la secuencia de operaciones que meticulosamente aplica el cartógrafo sobre el conjunto de geometrías que componen el contenido del mapa, con la finalidad de dotarle de legibilidad, representatividad y estética.

Estas operaciones, en términos generales, las podemos clasificar en dos grandes grupos en función de su finalidad:

- las **operaciones de generalización,** que adecúan el contenido del mapa a la escala final de visualización. Cuando la resolución de los datos de partida, en este caso de la base de datos topográfica, es mayor que a la que va a ser percibida según la escala de representación (determinada por el límite de percepción visual de 0,2 mm), se genera una visualización borrosa y empastada del contenido del mapa, además de un

consumo de memoria ineficiente. Un ejemplo de este tipo de generalización es la eliminación de elementos no representativos, o simplificación de formas, que se explican con detalle más adelante.

- las **operaciones de edición cartográfica**, que se encargan de proporcionar estética y legibilidad al mapa, favoreciendo su lectura e interpretación. Para ello, se aplican una serie de operaciones, como rotaciones de elementos, desplazamientos para evitar empastes, exageraciones de distancias, entre otras muchas, que también son explicadas más adelante.

5.2.4.2. Automatización del proceso

Los procesos implicados en la automatización del proceso de redacción cartográfica pretenden reproducir todas las operaciones que se suceden en la elaboración del mapa por técnicas clásicas. Para ello, es necesario diseñar el flujo automático que consiga aplicar las técnicas de redacción cartográfica a los elementos provenientes de las fuentes de origen, incluyendo por tanto los siguientes objetivos:

- generalización a la escala de visualización
- clasificación según una leyenda
- edición de elementos para asegurar la legibilidad (evitar solapamientos o empastes, acentuar distancias para facilitar la visualización)
- rotulación de entidades (dada su relevancia, se explica con detalle más adelante)
- refinamiento estético

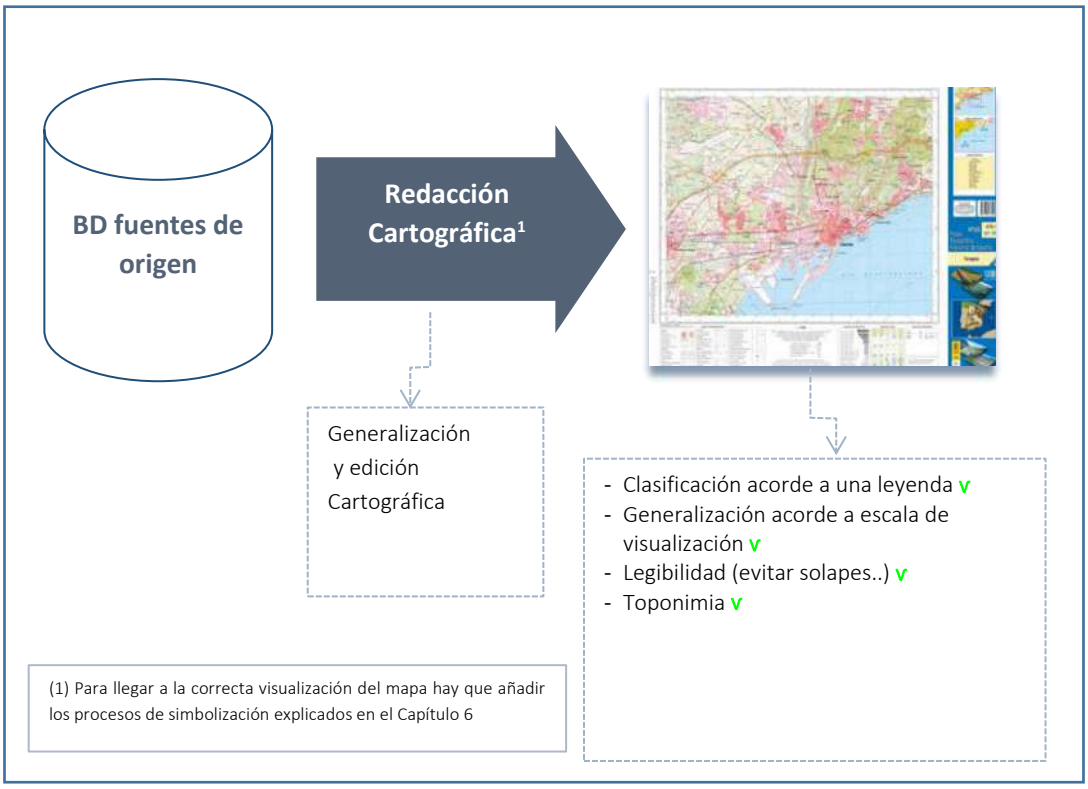


Ilustración 5 - 36: Redacción Cartográfica, contexto y objetivo

Llegados a este punto, conviene aclarar que este apartado se centra en la automatización de la redacción cartográfica, referida a la obtención de geometrías adecuadas para la escala del mapa. Para la correcta visualización final del mapa, faltaría añadir las operaciones de simbolización (colores, estilos, fuentes, formatos, etc.), cuya explicación detallada requiere un capítulo entero aparte (Capítulo 6).

El flujo automático es una sucesión de procesos sincronizados que actúan sobre fuentes de origen, o salidas parciales internas del proceso, hasta generar la salida cartográfica. En su diseño hay que tener en cuenta las características de los objetos geográficos, así como los diferentes entornos en los que se encuentran, y la alta interrelación entre los mismos. Como consecuencia, al actuar sobre un objeto se pueden alterar otros objetos que ya han sido tratados en procesos previos. Por ello se requiere un diseño que permita realizar el análisis y determinar las actuaciones a llevar a cabo, tanto a nivel individual, de cada fenómeno, como a nivel global, considerando el contexto e interrelaciones de estos fenómenos.

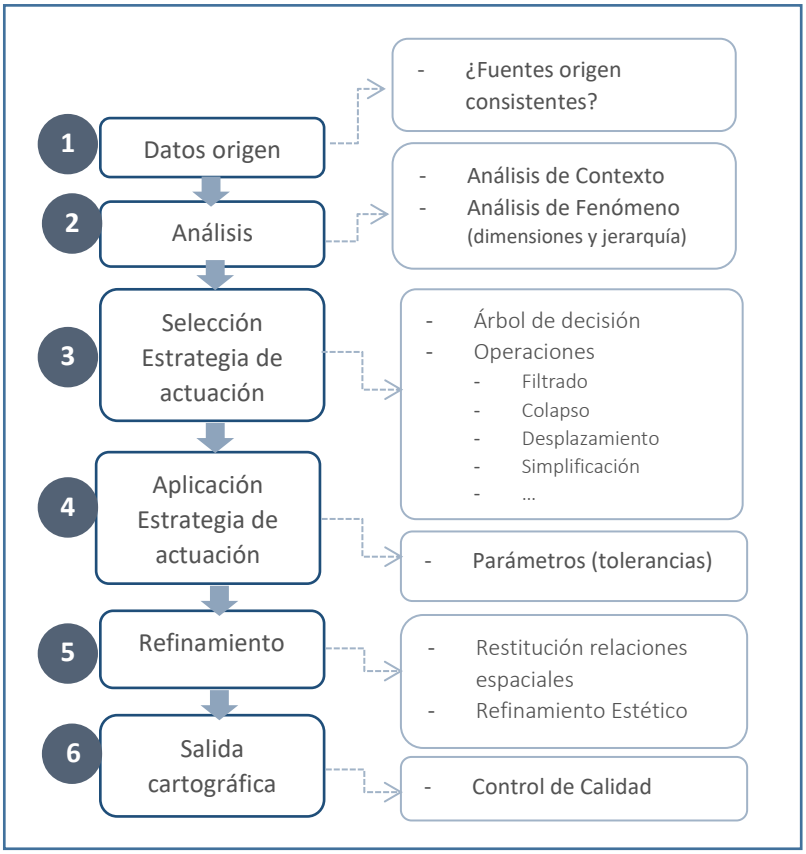


Ilustración 5 - 37: Esquema conceptual del flujo de automatización

A continuación, se detallan cada uno de los elementos que componen el esquema general de automatización reflejado en la figura anterior:

(1) Datos origen

Una de las dificultades que atañen a la automatización de procesos en cartografía es, además de la gran cantidad y variabilidad de los conjuntos de datos que componen las fuentes de origen, la elevada casuística y elevado número de relaciones entre los fenómenos que representan estos datos.

Por este motivo, es fundamental partir de fuentes de datos armonizadas, con relaciones semánticas y topológicas conocidas y aseguradas según una definición de modelo de datos. De lo contrario, el proceso de automatización está sometido a una serie de riesgos tan indeterminados como lo sea el conocimiento de las fuentes de origen, y a una elevada dificultad en la convergencia de los resultados.

Esto pone de manifiesto la importancia de la consistencia de las fuentes origen. Partir de un modelo estructurado y definido mediante sus correspondientes especificaciones, posibilita, o al menos facilita, la implementación de un proceso automático que actúe sobre ellos.

(2) Análisis

De igual manera que durante el proceso de redacción cartográfica manual un cartógrafo analiza cada objeto geográfico y su entorno antes de actuar sobre él, durante el flujo automático intervienen algoritmos que examinan el contexto de cada elemento a la hora de decidir las posibles estrategias de tratamiento que se pueden adoptar. Se analizan tanto las características intrínsecas a cada objeto (dimensiones, jerarquía) como su contexto (densidad, relaciones).

Así, por ejemplo, si las dimensiones de un edificio por sí solas no son determinantes en la elección de su tratamiento (eliminación, colapso, escalado, etc.) habrá que tener en cuenta otros factores influyentes de su entorno, como puede ser la densidad de edificación en la que se encuentre.

(3) Selección de estrategia de actuación

Una vez analizado cada objeto y su entorno, el flujo selecciona la estrategia de actuación acorde a dicho análisis. Estas estrategias o reglas de actuación se llevan a cabo mediante operaciones específicas de generalización y edición cartográfica, o más frecuentemente mediante una combinación de varias (operaciones que se explican detalladamente más adelante).

Dada la elevada casuística que puede existir, para determinar la estrategia más adecuada se requiere de una clasificación lo más pormenorizada posible de dicha casuística, a partir de la que asignar una regla de actuación a cada escenario. Como ejemplo de selección de estrategia de actuación, en el apartado 5.2.8.2 se muestra un árbol de decisión sobre la capa de edificaciones, asignando una regla a cada escenario de edificación.

(4) Aplicación de estrategia de actuación

Una vez decidida la estrategia, ésta se aplica siguiendo los parámetros definidos en cada operación (flechas, tolerancias, superficies mínimas, lados mínimos, etc.).

Especialmente en este punto hay que tener en cuenta la alta dependencia entre los elementos que componen el mapa. Después de cada actuación pueden provocarse efectos secundarios que afectan a elementos ya tratados en pasos anteriores. Por ello, es necesario definir un orden de actuación en función de unas prioridades establecidas en cuanto a su probabilidad de modificación: prioridades de desplazamiento, de selección, de etiquetado, etc.

En primer lugar, se llevarán a cabo las actuaciones sobre los elementos **con mayor jerarquía y mayor restricción de modificación**, a partir de las cuales se irán sucediendo las siguientes manipulaciones sobre elementos con menor restricción.

(5) Refinamiento

Tras las actuaciones descritas que modifican y adecúan las geometrías a la escala y propósito del mapa, aparecen las que tienen por objeto conseguir una correcta estética y legibilidad, como son:

- devolver las relaciones espaciales que se hayan perdido durante los anteriores procesos. Un ejemplo es el de devolver las relaciones topológicas a los enlaces de una carretera tras la unificación de calzada.
- resolver los conflictos visuales y eliminar los solapes o empastes mediante desplazamientos o colapsos de elementos.
- aplicar procesos estéticos como son: orientar edificios paralelos a las carreteras, generar de símbolos representativos para el relleno del patrón de cementerios o las minas, etc.

(6) Salida Cartográfica

Una vez generada la salida cartográfica, es fundamental realizar el control de calidad, normalmente por métodos visuales de revisión, de forma previa a su publicación y, en su caso, someterla a un procedimiento de edición simplificada o reducida a los ítems más importantes, tal y como se explica en el correspondiente apartado dedicado a este tipo de edición.

5.2.4.3. Operadores utilizados

Como explica el epígrafe anterior, durante la redacción cartográfica automática se aplica una serie de actuaciones a los fenómenos geográficos basadas en procesos de generalización y edición automática. Estos procesos se basan en los operadores que se describen a continuación.

Filtrado de elementos (Eliminación/Selección)

Este operador detecta qué elementos de la base de datos de partida son adecuados para el propósito y escala final del mapa, con el fin de seleccionar los que se quieren conservar y los que se precisa descartar. Por ejemplo, la capa de edificios de la base de datos de partida sufrirá un complejo proceso de selección: se eliminarán, por ejemplo, edificios aislados en urbanizaciones sin alterar la forma general de la urbanización y se seleccionarán los edificios representativos por su importancia.

Esta operación se apoya en las dimensiones de los elementos, tanto lineal como superficial, analizando si la dimensión es adecuada para la escala final, y en el grado de representatividad por su categorización o unicidad en el entorno (Ejemplo: el número de vecinos próximos de la misma categoría).

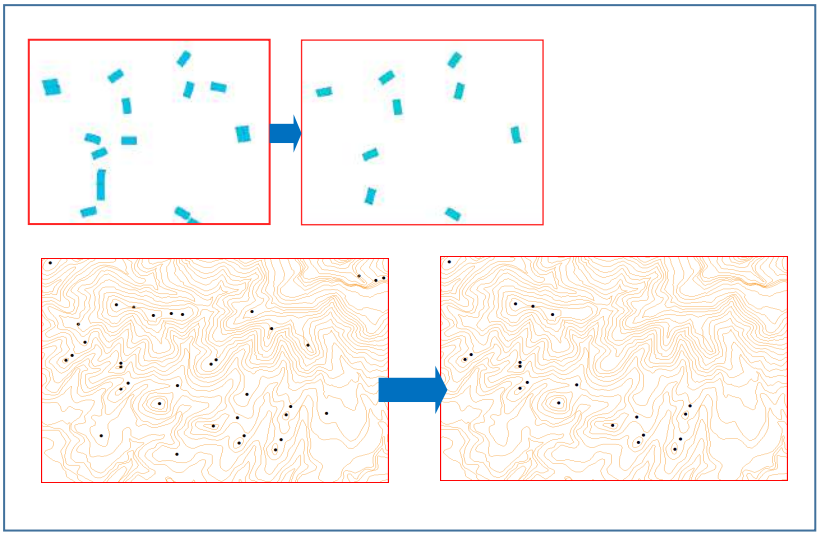


Ilustración 5 - 38: Proceso de filtrado aplicado a piscinas y puntos de cota

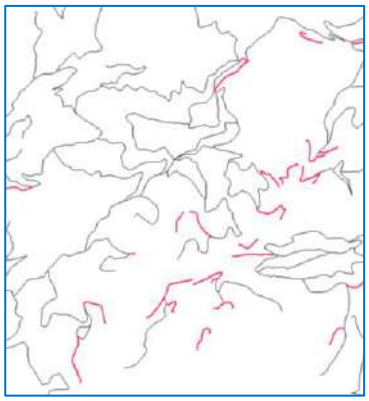


Ilustración 5 - 39: Selección de caminos a eliminar en función de su longitud y ubicación

Simplificación de formas

Mediante este operador se eliminan detalles que por su dimensión no son perceptibles en la escala final del mapa. Se aplica tanto a elementos lineales como superficiales, eliminando vértices que generan flechas menores a las apreciables por el ojo humano. Su aplicación es más fácil sobre elementos de formas irregulares, por ejemplo, formas de caminos o bordes de lagos. Sin embargo, tiene una aplicación más compleja cuando se trata de elementos de formas regulares, por ejemplo, en edificios, ya que la eliminación indiscriminada de vértices que solo atiende a flechas y tolerancias haría perder la forma representativa del edificio de lados ortogonales, dando lugar a superficies deformadas.

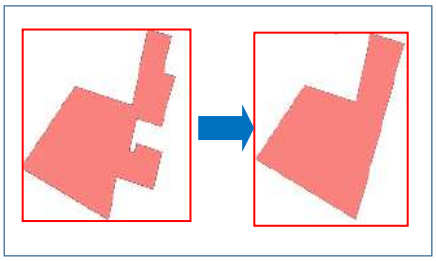


Ilustración 5 - 40: Simplificación de formas aplicada al contorno de un edificio

Suavizado

Los operadores de suavizado están más orientadas a la estética final de visualización que a la generalización. Se suelen aplicar en combinación con operadores de simplificación de formas para evitar que la eliminación de vértices genere formas picudas o quiebros bruscos en elementos con geometrías de formas suaves (carreteras, caminos), y nunca en elementos de formas regulares. Por ejemplo, no lo usaríamos en contornos de edificios u otras figuras ortogonales.

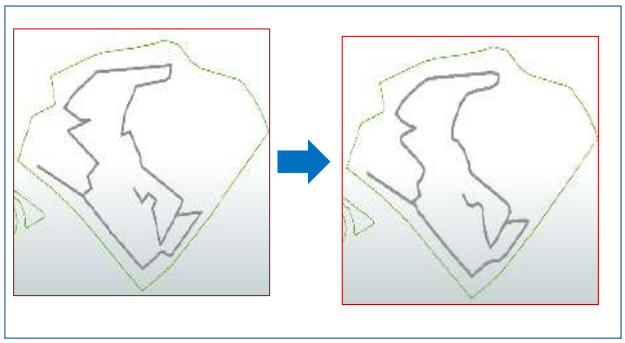


Ilustración 5 - 41: Suavizado de caminos dentro del perímetro de un parque

Amalgamado

Este operador pretende fusionar superficies de ciertas dimensiones que se encuentren a una distancia menor a la perceptible a la escala final. Por debajo de este límite, equivalente al de percepción visual, esa separación no se percibe con claridad, generando una visión borrosa y poco estética.

Un ejemplo muy característico de amalgamado es el que sufren los edificios dentro de casco de población, fusionándose aquellos cuyos bordes se encuentran a poca distancia, tal y como se muestra en la siguiente figura.

Ilustración 5 - 42: Amalgamado de edificaciones que se encuentran próximas

Colapso

Este operador sustituye a una o varias geometrías por otra más sencilla. Puede aplicarse a elementos de tipo superficial, sustituyendo el objeto por símbolos, o a elementos de tipo lineal, como en el típico ejemplo de la sustitución de los dos ejes de una carretera de doble calzada (uno por sentido) por el eje por eje central de dicha carretera.

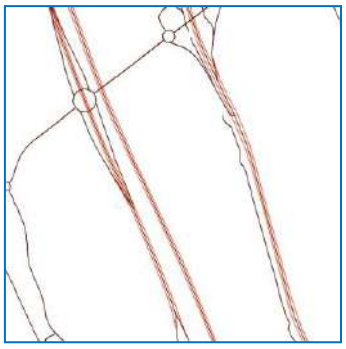


Ilustración 5 - 43: Colapso de las dos calzadas de una carretera en una sola

Exageración/Ensalzamiento

Al aplicar operaciones de simplificado se puede perder la particularidad de los objetos, por ello se deben ensalzar determinados aspectos que los caractericen. Por ejemplo, en los cascos de población, se ensalza el entramado urbano recortando las edificaciones con las vías.

Ilustración 5 - 44: Exageración de distancias entre edificios para acentuar esqueleto urbano

Desplazamientos

Los desplazamientos están destinados a evitar solapes y empastes entre las distintas geometrías. Además, nos permiten exagerar distancias entre elementos aumentando su legibilidad.

Encontramos su aplicación en ejemplos muy característicos como en los desplazamientos entre elementos lineales, entre edificaciones o de edificaciones respecto vías de comunicación.

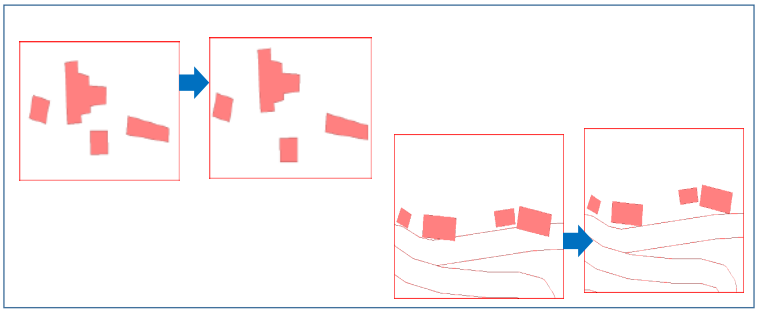


Ilustración 5 - 45: Desplazamientos aplicados a edificaciones y vías de comunicación

Tipificación

La tipificación consiste en la extracción de un patrón y su posterior aplicación de forma simplificada de manera que se mantenga la percepción general de una zona. En la siguiente figura se muestra un ejemplo de aplicación de colapso y filtrado de edificaciones en una zona de urbanización, conservando su estructura visual.

Ilustración 5 - 46: Tipificación aplicada a la capa de edificaciones de una urbanización

Ajuste de contornos

Esta operación supone un ajuste de los perímetros de áreas colindantes, bien porque durante la redacción cartográfica se hayan generado alteraciones geométricas que hagan perder esa continuidad, o bien para aumentar la legibilidad de ciertos elementos del mapa.

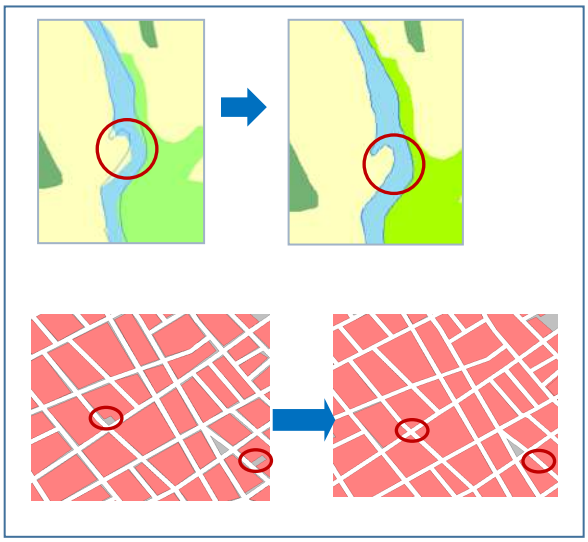


Ilustración 5 - 47: Ajuste de contornos en bordes de cultivos y edificios

Refinamiento estético

Durante el refinamiento estético se persigue, además de dotar al mapa de una percepción atractiva, aumentar la representatividad de algunos elementos. Para ello se pueden aplicar recursos tan variados como son orientar edificios paralelamente a las carreteras o generar símbolos representativos para rellenar contornos de cementerios, minas o centrales solares.

Ilustración 5 - 48: Aplicación de símbolo con un patrón en un área de central solar

5.2.5. Etiquetado automático

El etiquetado automático es, sin duda, uno de los desafíos más complejos de todo el proceso: **la selección y correcta ubicación de la toponimia en forma de rótulos**. En este proceso se generan etiquetas para los elementos, con un estilo definido en función de la clasificación del elemento rotulado, en una posición que identifique sin ambigüedad el objeto etiquetado, y que no genere conflictos con otros textos o elementos del mapa. Gran parte de esta dificultad dependerá de la cantidad de rótulos provenientes de las fuentes de datos origen, y de la existencia de una jerarquía adecuada para una misma clase de elementos a etiquetar. Hay que prever que un proceso automático seguramente no sea capaz de rotular todas las entidades que intervienen en el proceso, y por ello, será necesaria una selección de esos elementos y una asignación de jerarquías para dar prioridad a unos elementos sobre otros.

5.2.5.1. Clasificación de elementos. Asignación de propiedades de texto

Previamente al etiquetado automático, se deben realizar las correspondencias entre los objetos geográficos con representación en el mapa y el tipo de texto con el que se rotulará. Cada tipo de texto tendrá asociados una fuente de texto y un tamaño determinados, que deberán ser tenidos en cuenta a la hora de calcular el espacio real que va a ocupar cada texto, y así poder relacionarlo con los demás rótulos y elementos, para eliminar conflictos entre ellos.

Los tipos de texto deben estar especificados en el modelo o leyenda del Mapa.

Ilustración 5 - 49: : Categorización de los textos de entidades de población en la leyenda del MTN25

La correspondencia entre objetos geográficos y tipos de texto puede realizarse directamente en el proceso de traducción del modelo, asignando en cada elemento el tipo de texto junto con el código asociado a la salida en el mapa.

En otros casos, cuando el modelo, tanto de la fuente de datos origen como del mapa, no permite una asignación directa, será necesario realizar otra clasificación dentro de una misma clase de entidad, bien por análisis de determinados atributos, o bien de las propias geometrías. Por ejemplo, la clasificación de un elemento según su superficie requerirá analizar el valor de un atributo (superficie) si existiera, o analizar la geometría para obtener ese valor y efectuar la clasificación.

HIDROGRAFÍA CONTINENTAL

RÍOS, RÍAS Y ARROYOS

Proyecto	CLAVE	MENÚ	DESIGNACIÓN	RÓTULO	FUENTE – TX NI CO PE ES
MTN 25	430	RÍOS Y RÍAS DE 1º	Ríos de 1er orden y rías principales	RÍO EBRO	25 Mayúsculas – 75 4 121 03 0
MTN 25	440	RÍOS Y RÍAS DE 2º LAGUNAS>100 Ha	Ríos de 2º orden (más de 90 km) rías medianas y lagunas mayores de 100 ha	Río Jarama	25 Minúsculas – 80 4 121 02 0
MTN 25	450	RÍOS, AYOS. 3er y 4º RAMBLAS GRANDES LAGUNAS MEDIANAS	Ríos o arroyos de 3er y 4º orden, (de 25 a 90 km),ramblas grandes y lagunas de 10 a 100 ha	Río Negro	25 Minúsculas – 65 4 121 00 0
MTN 25	460	RÍOS Y ARROY.5º y 6º LAGUNAS PEQUEÑAS FUENTES, MANANT.	Ríos o arroyos de 5º y 6º orden (menos de 25 km), lagunas pequeñas menos de 10 ha, charcas, fuentes, pozos, manantiales, salinas, depuradoras, etc.	Arroyo Claro	25 Minúsculas – 50 4 181 00 0
MTN 25	490	HIDRÓNIMO PUNTUAL CASCAD. RAPID. VADOS	Hidrónimos puntuales: cascadas rápidos, pozas, vados, cuevas marítimas, etc.	Cola de Caballo	25 Minúsculas – 50 4 081 00 0
EMBALSES					
MTN 25	500	EMBALSES >1.000 Ha	Embalses mayores de 1000 ha	EMBALSE DEL ZÚJAR	25 Mayúsculas – 74 a 98 Hor. 4 151 02 0
MTN 25	510	EMBALSES 100 a 1.000 Ha	Embalses de 100 a 1000 ha	Embalse de Valdeinfierno	25 Minúsculas – 80Hor. 4 061 02 0
MTN 25	520	EMBALSES 10 a 100 Ha	Embalses de 10 a 100 ha	Embalse de Revenga	25 Minúsculas – 65Hor. 4 061 01 0
MTN 25	530	EMBALSES <10 Ha	Embalses de menos de 10 ha	Embalse de Valdeósgua	25 Minúsculas – 50Hor. 4 061 00 0

Ilustración 5 - 50: Categorización para los rótulos de los elementos río y embalse en la leyenda del MTN25

En la siguiente figura puede observarse que, partiendo del modelo de datos de BTN, hay una relación directa entre los valores del atributo Jerarquía de la tabla 0302L_RIO y los de la leyenda de mapa. Sin embargo, en la tabla 0325S_EMBALSE no existe ese atributo y, por tanto, la clasificación se hará por análisis de la superficie.

JERARQUÍA - JERAR_0302 **CodeList**

Define la clasificación que se establece para el río en su conjunto, según su longitud e importancia respecto al conjunto de ríos de España

01	PRIMERA	Duero, Ebro, Guadalquivir, Guadiana, Júcar, Miño, Segura y Tajo
02	SEGUNDA	Ríos excluidos de la lista de PRIMERA y longitud > 90 Km
03	TERCERA Y CUARTA	Ríos de tercera categoría (40-90 Km) y cuarta categoría (25-40 Km)
04	QUINTA Y SEXTA	Ríos de quinta categoría (10-25 Km) y sexta categoría (< 10 Km)

Ilustración 5 - 51: Atributo Jerarquía de la tabla 0302L_RIO en el modelo de datos de BTN

5.2.5.2. Selección de elementos a etiquetar

Al igual que en el tratamiento de las geometrías, a la hora de rotular elementos hay que tener en cuenta la escala del mapa y la de la fuente de datos origen. En este contexto de rotulación, la escala determina la resolución, no geométrica, sino temática, de los datos. Una base de datos geográfica, aunque tenga una escala equivalente a la del mapa, puede contener una

mayor carga temática. Esto ocurre especialmente en los nombres de las entidades, que se encuentran almacenados en un atributo.

Para seleccionar los elementos a etiquetar, hay que tener claro los requerimientos del mapa que se va a publicar, así como qué objetos pueden ser etiquetados. Por ejemplo, aunque se representen en el mapa los edificios de un casco de población, no siempre se rotulará el nombre de todos ellos, en muchos casos será necesario realizar una selección de los nombres a rotular.

0590P_SERV_INS	ATRIBUTO_TI	ETIQUETAR
01 CENTRO DOCENTE NO UNIVERSITARIO	01	NO
02 CENTRO UNIVERSITARIO	02	NO
03 OTROS CENTROS DOCENTES	03	NO
04 HOSPITAL	04	SI
05 OTROS CENTROS SANITARIOS	05	NO
06 RESIDENCIA DE MAYORES	06	NO
07 RESIDENCIA DE ESTUDIANTES	07	NO
08 ALBERGUE	08	SI
09 BALNEARIO	09	SI
10 PARADOR	10	NO
11 MUSEO	11	SI
12 CENTRO DE INTERPRETACIÓN-VISITANTES	12	SI
13 CENTRO CULTURAL	13	NO
14 ARCHIVO-BIBLIOTECA	14	NO

Ilustración 5 - 52: Tabla de decisión sobre rotulación según el atributo 'Tipo'

Otro aspecto a tener en cuenta es el de la categorización de los datos. Si los objetos se encuentran suficientemente clasificados por medio de sus atributos, la selección será más sencilla. Por ejemplo, si disponemos de una tabla de viales urbanos y estos se encuentran clasificados en varias categorías, el proceso podría seleccionar las vías más importantes para rotularlas y descartar el resto.

Por último, se pueden efectuar selecciones de elementos a rotular mediante análisis espacial, por ejemplo, selección de determinados nombres geográficos dentro de cascos de población, de un máximo de elementos en una partición espacial, etc. Cuando sea viable, se puede combinar la selección mediante operaciones espaciales con el análisis de la jerarquía de los elementos.

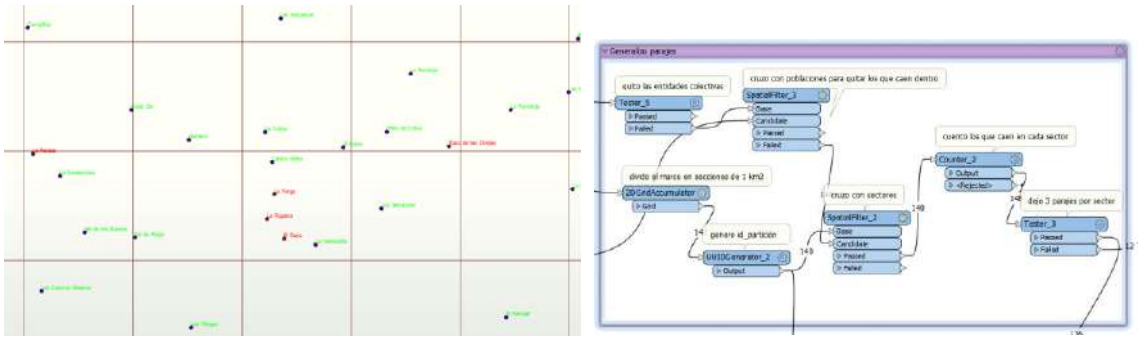


Ilustración 5 - 53: Generalización de textos por selección dentro de una partición

5.2.5.3. Asignación de jerarquías

Como se ha comentado, la jerarquía de los elementos es muy importante para la rotulación. Nos ayuda a realizar una primera selección de los elementos a etiquetar y, posteriormente, establece las prioridades de unos elementos sobre otros para el proceso de etiquetado automático. Se pueden establecer jerarquías a diferentes niveles:

- **A nivel de tabla**: se pueden asignar directamente jerarquías a todos los elementos de una tabla. Por ejemplo, establecer una prioridad alta para rótulos de una tabla de entidades de población, de carreteras, o de puntos acotados, y otra más baja para las tablas de curvas de nivel, cauces artificiales, etc.
- **A nivel de clase**: dentro de una misma tabla, se pueden asignar jerarquías según el valor de un determinado atributo, o de una combinación de varios. Por ejemplo, dentro de una tabla de ríos, si existe el atributo adecuado, podemos dar prioridad a los de mayor categoría. Esta categorización puede establecerse en la misma tabla de traducción del modelo, o más adelante, durante el proceso automático.
- **Por análisis espacial**: cuando no se puede categorizar por medio de los atributos de los elementos, se pueden ejecutar operaciones de análisis espacial para establecer esas categorías. Por ejemplo, clasificar por superficie los elementos de una tabla de embalses y dar prioridad a los de mayor categoría.
- **Por comparación con otras bases de datos**: otro tipo de análisis sería comparar los elementos con versiones anteriores del mapa, de manera que tengan mayor jerarquía los textos existentes en dichas ediciones históricas, o en cartografía existente a otras escalas, etc.

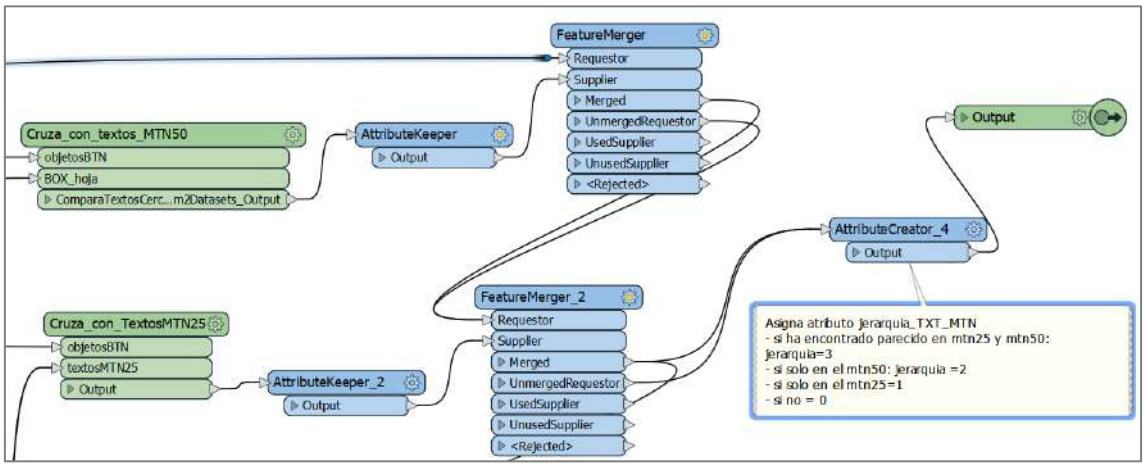


Ilustración 5 - 54: Asignación de jerarquías a textos mediante comparación con ediciones históricas de MTN25 y MTN50 en FME

Este procedimiento de asignación de jerarquías es análogo al de clasificación, visto en el punto *5.2.5.1. Clasificación de elementos. Asignación de propiedades de texto*, pero se puede llevar más allá, de manera que un mismo tipo de texto puede tener distintas prioridades, si hemos podido establecer diferencias entre los elementos. Por ejemplo, en una tabla de edificaciones, podemos configurar el proceso para que dé menos prioridad a textos de edificaciones dentro de los núcleos de población, donde puede existir abundancia de nombres, en favor de otros textos, y mantenga una prioridad más alta fuera de ellos.

5.2.5.4. Reglas de etiquetado. Elementos puntuales, lineales y superficiales

Una vez tenemos preparados todos los elementos para realizar el etiquetado, es importante establecer unas reglas para el desarrollo o implementación de los algoritmos o herramientas que se van a emplear. Estas reglas determinan la forma en que deberán situarse los textos

respecto de los objetos a etiquetar. Podemos establecer varios grupos de reglas, atendiendo al tipo de geometría que se va a etiquetar:

- **Elementos puntuales:** las reglas para estos objetos establecen fundamentalmente la distancia a la que debe estar el texto respecto del punto, que dependerá del tamaño y la forma de la representación del objeto puntual, y la posición respecto del punto. En algunos casos también puede establecerse si el texto puede ir orientado según una posible orientación del objeto puntual.
- **Elementos lineales:** las reglas en este grupo establecen si el texto debe ir orientado o no sobre la línea, la distancia a la misma, dependiendo del ancho de su representación y de la escala, y la posición sobre la línea. Dependiendo de la longitud del elemento, también se puede indicar si el texto debe ir repetido a lo largo del elemento.
- **Elementos superficiales:** en este caso puede establecerse la posibilidad de situar el texto en el interior o en el exterior del polígono, si puede o no sobresalir de la superficie a etiquetar, y si se puede orientar o debe ser siempre horizontal.
- **Reglas comunes:** se trata de un grupo de reglas generales aplicables a todos los elementos, como la posibilidad de que solapen con otros elementos, si pueden partirse en más de una línea, posibilidad de reducir el tamaño del texto, la utilización de líneas auxiliares, etc.

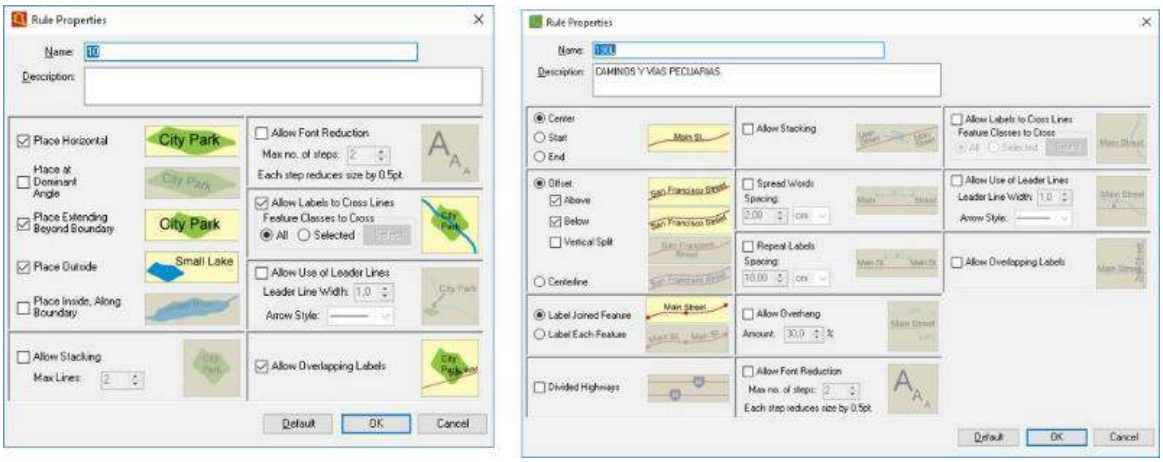


Ilustración 5 - 55: Reglas de edición de elementos superficiales y lineales. Módulo MapTextLabeller en FME

A continuación, el proceso de etiquetado comenzaría a rotular, siguiendo las reglas establecidas: primero, los elementos de mayor prioridad y después, reservando el espacio de los textos ya colocados, los elementos pertenecientes a los siguientes niveles de prioridad.

5.2.5.5. Detección de solapes. Resolución de conflictos

Finalizado el proceso de etiquetado automático es importante tener previsto un mecanismo de control de posibles solapes entre textos. Una vez detectados estos solapes, se necesitará un proceso de decisión para eliminar aquellos textos en los que existan conflictos. Aquí podría intervenir de nuevo la jerarquización, de manera que de dos textos que se solapen, elimine el de menor jerarquía.

Es importante almacenar de forma independiente todos los textos que no hayan sido finalmente etiquetados, con el fin de poder recuperarlos en una posible nueva edición de la hoja.

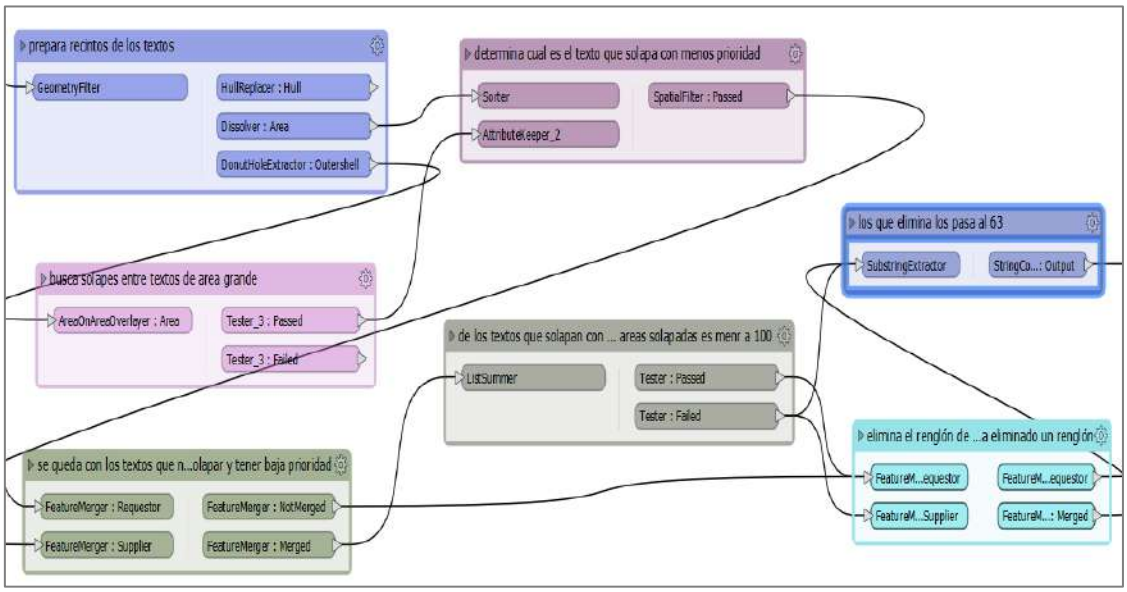


Ilustración 5 - 56: Esquema del proceso de eliminación de solapes de textos en FME

5.2.6. Salidas Intermedias

En este apartado se explican los diferentes tipos de salida que se puede dar a la información cartográfica, generada mediante los procesos automáticos de generalización y edición.

Los diferentes formatos de salida dependerán del uso y de los procesos que se vayan a aplicar a estos datos.

La salida más inmediata es su almacenamiento en una base de datos cartográfica, donde la información se organiza según: un modelo de datos definido en conformidad con los requisitos de uso de esta información, la simbolización de los elementos y su publicación (que puede ser por diferentes canales (edición impresa, servicios de visualización, ediciones digitales en archivos descargables). Este almacenamiento en base de datos se explica más adelante.

En relación a la calidad de las salidas, y tal como se ha apuntado en la introducción, aunque los procesos automáticos tratan de emular las tareas del especialista cartográfico, en ocasiones los resultados no son los esperados y requieren de una edición complementaria. Los problemas más frecuentes que se pueden encontrar en la salida cartográfica son:

- **Omisiones o errores en las fuentes de datos**: pérdidas de información, desactualización, entradas no válidas, falta de consistencia entre capas.
- **Errores del proceso**: diferente casuística en los datos, fallos inesperados, cases de hoja.
- **Rotulación**: solapes de textos, omisiones por el proceso de etiquetado, colocación errónea de textos.

Cuando se precisa de esta edición a posteriori, es recomendable exportar los datos a un formato adecuado para trabajar en un entorno de edición de cartografía que muestre la

información con su simbolización, y permita así revisar y editar los conflictos entre elementos. El formato de salida dependerá del entorno en el que se vaya a trabajar.

En el caso del MTN25, el entorno de edición está desarrollado sobre software de CAD MICROSTATION. Por esa razón se utiliza un formato de salida intermedio, previo al almacenamiento en base de datos, que es el DGN, formato propio de este software CAD. En el entorno de edición se trabaja con dos salidas distintas, una para planimetría y otra para coberturas y usos del suelo, tal y como se explica a continuación.

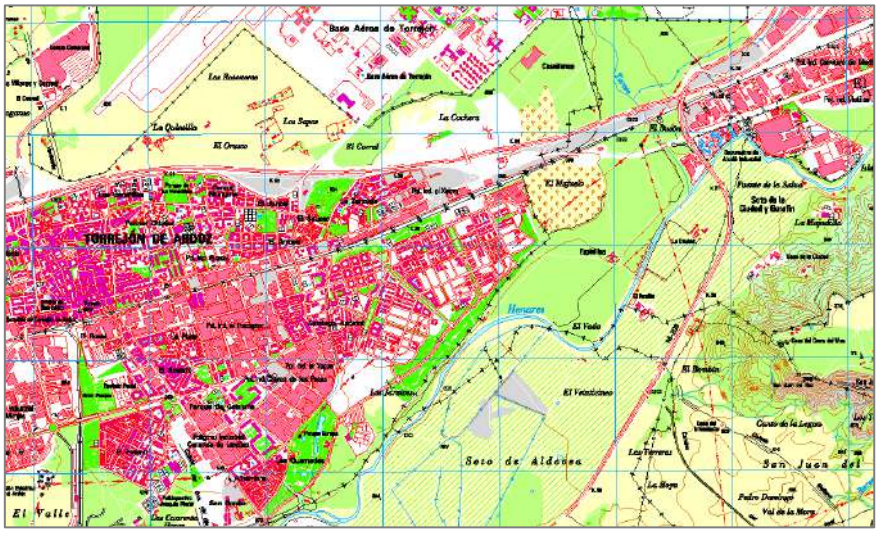


Ilustración 5 - 57: Salida en formato DGN simbolizada

5.2.6.1. Planimetría

En esta salida se almacena toda la información proveniente de las bases de datos BTN, Líneas límite administrativas y red geodésica. La información está estructurada en niveles, agrupados por temas de información, y atributos propios de MICROSTATION con los que quedan identificados según el modelo de datos del MTN25.

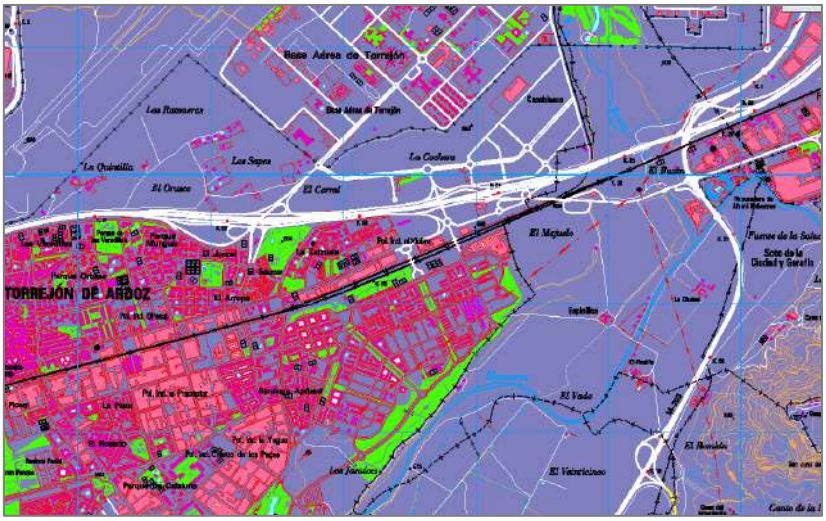


Ilustración 5 - 58: Salida de planimetría en formato DGN

5.2.6.2. Coberturas y usos del suelo

En esta salida se almacena la información proveniente de la base de datos del SIOSE. La información consta en su mayoría de elementos superficiales que representan las diferentes coberturas del suelo.

Ilustración 5 - 59: Salida de coberturas del suelo en formato DGN

El resultado de la edición se puede exportar de nuevo a una base de datos cartográfica, o generar directamente salidas para publicación en los formatos mencionados anteriormente.

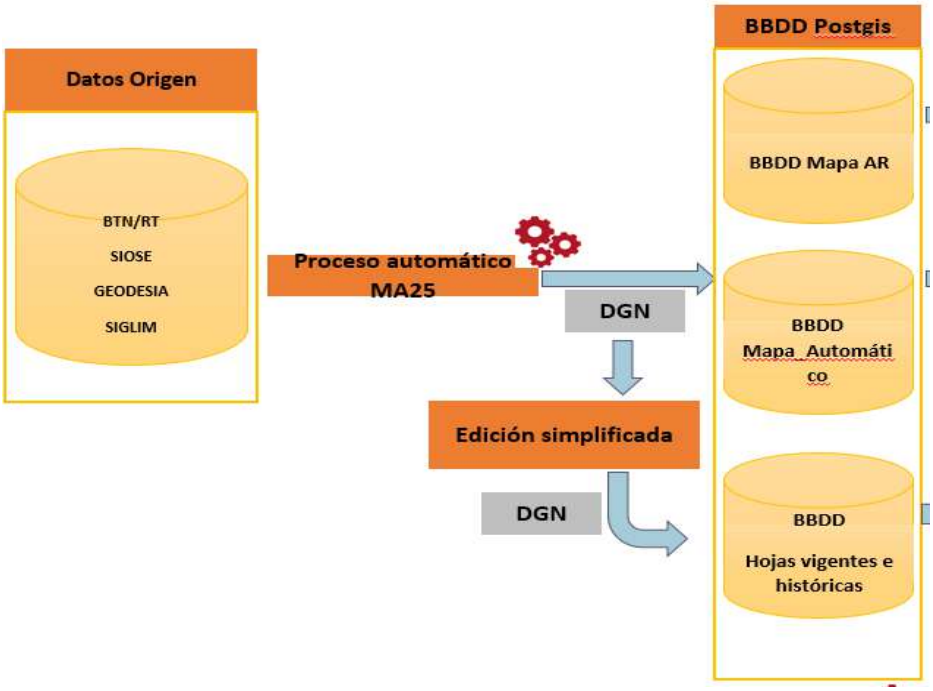


Ilustración 5 - 60: Salidas intermedias en el flujo de producción

5.2.7. Almacenamiento en Base de Datos Cartográfica

Una vez ejecutados los procesos automáticos de redacción cartográfica y generalización, las geometrías se almacenan en una base de datos cartográfica que será, a su vez, el punto de partida para el siguiente proceso automático de simbolización y rasterización, que se detalla en el Capítulo 6.

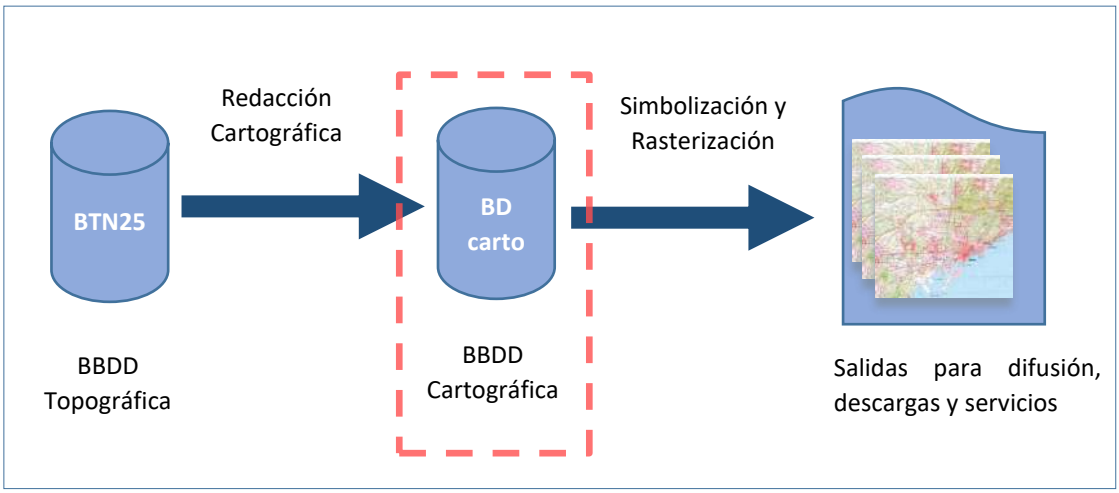


Ilustración 5 - 61: Ubicación de la BBDD Cartográfica dentro del flujo de automatización

Las geometrías almacenadas en esta base de datos están preparadas para su visualización, es decir, están generalizadas, desplazadas y modificadas para su correcta visualización a escala 1:25,000, por lo que hablamos de una **base de datos cartográfica**, a diferencia de la base de datos topográfica de partida, que almacena las geometrías en su verdadera posición en función de la precisión que tengan los datos.

Cada una de estas geometrías se almacena acompañada de un identificador de estilo de visualización que, posteriormente, el proceso de simbolización utilizará para aplicar la plantilla de estilos correspondiente. Es decir, cada uno de los identificadores de visualización almacenados estarán relacionados con una entrada en la plantilla de estilos del proceso de simbolización.

Además, también se almacenan los rótulos resultantes del proceso automático de rotulación, que van acompañados de su estilo de visualización y todos los parámetros necesarios para su correcta reproducción (orientación, tamaño, fuente, etc.).

En el caso del entorno productivo del MTN25, esta base de datos está implementada en Postgres, con la extensión espacial Postgis, y en el flujo automático de producción se generan dinámicamente las consultas de inserción y lectura de las geometrías y rótulos.

La base de datos debe considerar toda la información necesaria para que el proceso de simbolización sea capaz de reproducir y simbolizar cada uno de los objetos geográficos. Para ello, su diseño debe contemplar los requisitos que se muestran en la siguiente tabla.

Objeto geográfico	Requisito
Geometrías lineales y poligonales de carácter general	Debe almacenar su geometría y su identificador del estilo de visualización.
Geometrías puntuales de carácter general	Debe almacenar su geometría y estilo de visualización, y también un atributo que defina en su caso la escala y orientación del símbolo.
Geometrías lineales y puntuales de elementos orográficos (puntos de cota + curvas de nivel)	Aunque no se necesita para el proceso de simbolización, es recomendable almacenar, además de los atributos especificados, el atributo cota.
Rótulos	Una base de datos Postgis no permite almacenar geometrías tipo texto, por lo tanto, se almacenan como geometría puntual, añadiendo la información mínima necesaria para posteriormente reproducir el rótulo: contenido del texto, punto de inserción, orientación, escala, estilo de visualización.
Marcos de las unidades de producción (marcos de las hojas)	Debe almacenar las geometrías de cada uno de los marcos, más su numeración dentro del sistema cartográfico.
General	Debe vincular cada una de las geometrías almacenadas a la unidad de producción (marco de hoja), de manera que podamos hacer explotaciones de la información almacenada filtrando por hoja (inserciones o lecturas de las geometrías de una determinada hoja).
General	Debe permitir la actualización de la información de cada hoja. Cada vez que se almacena una nueva hoja se comprueba si existe una inserción de hoja con la misma numeración, en cuyo caso la eliminaría, eliminando además en cascada todas las geometrías vinculadas a ellas. Estas operaciones están automatizadas mediante consultas dentro del flujo de redacción cartográfica. Otra opción sería permitir almacenar las distintas versiones de la misma hoja.
General	Debe referir el sistema de coordenadas en el que se almacenan las hojas.
General	(Requisito aconsejable pero no necesario) Debe permitir la vinculación de las geometrías con sus rótulos, generando un atributo de vinculación.
General	(Requisito aconsejable pero no necesario) Debe almacenar el atributo 'ETIQUETA', que almacena el nombre de las entidades, acompañando a cada geometría almacenada, ya que en función del tipo de explotación puede ser vital (por ejemplo, si se quisiese hacer un etiquetado dinámico).

Tabla 5-2: Listado de requisitos para el diseño de la Base de Datos Cartográfica

5.2.8. Diseño e implantación del flujo de automatización

Este apartado se centra en la **implementación del flujo de automatización del proceso de redacción cartográfica** explicado anteriormente, siguiendo las directrices que ahí se plantean. Como refleja la siguiente figura, esta implementación corresponde con la parte de flujo que transcurre desde las fuentes de origen hasta la carga en la Base de Datos Cartográfica.

Este epígrafe se ilustra con el caso de buenas prácticas del flujo de producción automática implementado en el IGN de España.

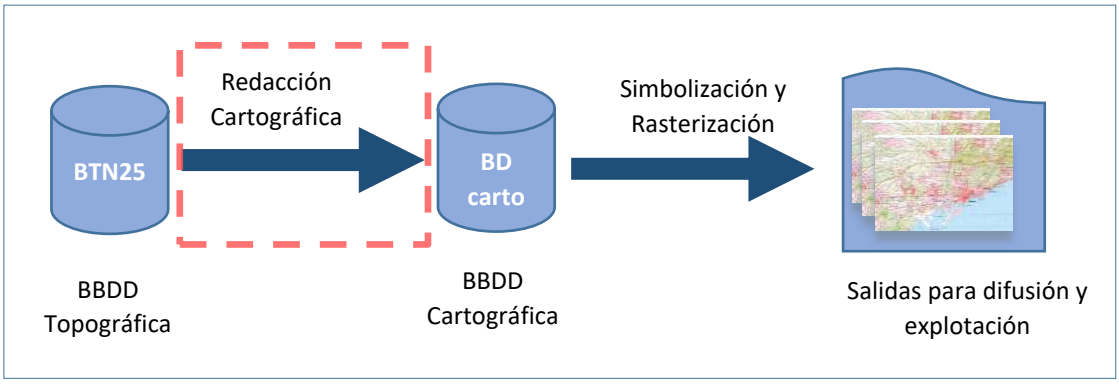


Figura 5 - 62: Procesos de redacción cartográfica, ubicación dentro del flujo general

Esta parte del flujo está implementada dentro de una única herramienta que integra y sincroniza todos los procesos que lo componen, transfiriendo salidas parciales y parámetros entre ellos. Su ejecución se realiza por lotes de hojas del MTN25, al seleccionar la numeración de hojas que se quieren procesar, y de forma automática sin interacción alguna por parte del operador, se generan las correspondientes salidas cartográficas.

5.2.8.1. Componentes generales

Entrada del proceso

La entrada al proceso la configuran las conexiones a las Bases de datos de las fuentes de referencia: Base Topográfica, líneas límite, SIOSE (usos del suelo) y vértices geodésicos.

Por otro lado, también recibe como input el listado de las hojas a procesar, es decir, las hojas que componen cada lote de ejecución.

Salida del proceso

El proceso genera dos tipos de salidas para cada una de las hojas del lote de ejecución:

- Para edición impresa, se generan salidas en formato DGN, a las que se aplicará un proceso de edición simplificada, según se detalla en la última sección de este capítulo.
- Para salidas de archivos raster a partir de los que se generan servicios de visualización (WMS/WMTS) y son accesibles por descarga, se genera el almacenamiento de las geometrías resultantes en la base de datos cartográfica. Esta salida será el punto de partida para el proceso de simbolización y rasterización, que se detalla en el capítulo 6.

Software utilizado

La implementación de esta parte del proceso se realiza sobre tres entornos diferentes:

- El flujo está implementado principalmente mediante el **software FME**. Se trata de una herramienta del grupo ETL (*Extract Transform Load*) que, entre sus amplias capacidades, como son las transformaciones entre múltiples formatos, permite el análisis y manipulación de geometrías espaciales mediante la secuencia de operaciones dispuestas de una forma gráfica. Es, por lo tanto, una herramienta que permite programar rutinas de una manera muy intuitiva, permitiendo trasladar fácilmente el flujo conceptual de diseño en una secuencia de transformadores.

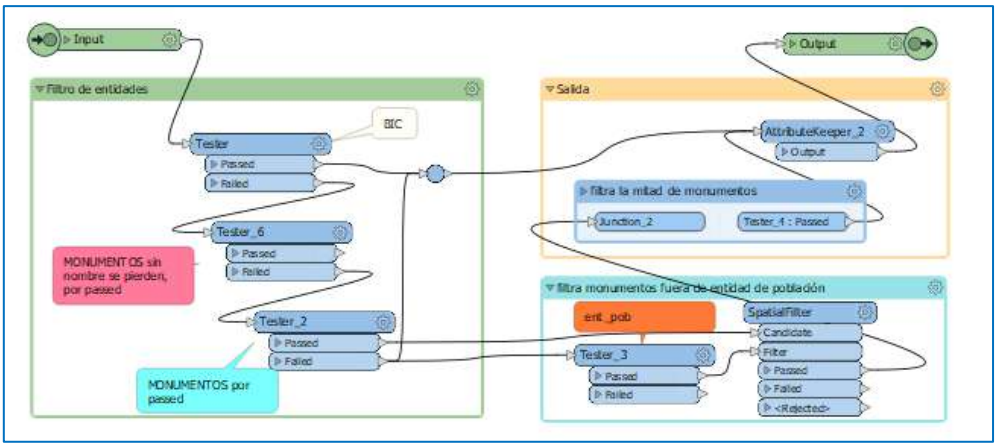


Ilustración 5 - 63: Ilustración de implementación mediante FME

Mediante una herramienta principal implementada en este software, se integran y sincronizan cada uno de los procesos que conforman el flujo. La mayoría de estos procesos se encuentran implementados a su vez con el mismo software, como, por ejemplo, la mayoría de algoritmos de generalización y árboles de decisión.

- Se utilizan **Librerías Arcpy**, un paquete de ArcGis que permite realizar análisis y automatizaciones de datos en el campo de la cartografía con Python. Estas librerías son utilizadas para procesos de edición y generalización preliminares, como la simplificación de formas de contornos, el colapso de doble vía a un eje central en comunicaciones o la generalización de entramado urbano.
- Por último, en la implementación de esta parte del flujo se utiliza el sistema gestor de **Postgres** con su extensión **Postgis** para el almacenamiento en la base de datos cartográfica. Mediante rutinas SQL, generadas y ejecutadas por FME, se gestiona el almacenamiento de las geometrías cartográficas en la BBDD.

Parámetros

Los parámetros que rigen las operaciones de generalización están orientados a satisfacer el límite de percepción visual que impone la salida en papel de la escala 1:25,000. Considerando un límite de percepción visual de 0,2 mm para el ojo humano, obtenemos los siguientes parámetros:

- Dimensión lineal: 0,2 mm a escala 25,000 = 5 metros. Este parámetro determinará longitudes y separaciones mínimas percibidas (luz entre elementos), por ejemplo, lados mínimos en contornos de edificaciones, separaciones entre elementos lineales como carreteras o entre superficies de edificaciones.
- Dimensión superficial: 0,2 mm x 0,2 mm a escala 25,000 = 25 metros cuadrados. Este parámetro corresponde a las superficies mínimas percibidas, determinando las superficies mínimas de un elemento para ser colapsado o seleccionado.

5.2.8.2. Flujo de operaciones

Ilustración 5 - 64: Flujo implementado para automatizar la redacción cartográfica

La implementación de los procesos automáticos de redacción cartográfica está compuesta por los siguientes subprocesos que se ejecutan secuencialmente uno tras otro:

1. *Proceso preliminar*

Consiste en un proceso preparatorio de la información a procesar, que engloba los siguientes hitos:

- La extracción de la información de las BBDD origen según los marcos de las hojas seleccionadas.
- Una generalización preliminar de las geometrías mediante conexión a librerías de Arcpy que actúa sobre:
    - Vías de comunicación: (1) colapso de carreteras de doble calzada a un solo eje, (2) desplazamiento de vías de comunicación entre sí para no producir solapes y (3) generalización del entramado urbano.
    - Edificaciones: simplificación de los contornos de los edificios según el parámetro de longitud mínima.

2. *Traducción del modelo*

A continuación, se realiza el proceso de traducción del modelo, descrito anteriormente y que consiste en un mapeo de las entidades que provienen de la base topográfica a las entidades simbolizadas en el mapa topográfico. Es decir, se asigna un identificador vinculado a las propiedades de visualización de cada entidad, que será almacenado junto con las geometrías en la base de datos cartográfica para su posterior explotación con el proceso de simbolizado, que se explica en el capítulo 6.

3. *Tratamiento de elementos lineales*

Se trata de un proceso que actúa sobre elementos lineales para adaptar su geometría a una correcta visualización a escala 1:25,000, mediante una serie de operaciones como las que se indican a continuación:

- Simplificación de vértices: esta operación tiene como finalidad eliminar el exceso de vértices en una línea poligonal en función del tamaño de la flecha determinado para la escala de visualización.
- Selección y eliminación de determinados elementos lineales: esta operación se aplica a elementos lineales concretos como cercas, alambradas, etc. Incluso a la red de caminos o a la de ríos, en caso de que la densidad de elementos capturados fuese superior a la admisible para la escala destino. En este recurso se realiza un filtrado de los elementos en función de su densidad, su conectividad con otros elementos y sus dimensiones.
- Generación de símbolos de puente y túnel sobre vías de comunicación: en este punto, a diferencia de los dos anteriores, no se trata de un recurso de generalización sino de mejora de la representatividad, con el que se destacan estéticamente los tramos elevados y subterráneos de las vías de comunicación.

4. Tratamiento de edificaciones

El proceso de edificios es el que conlleva más carga de ejecución, ya que contiene todas las operaciones necesarias para editar la capa de edificios con el objetivo de proporcionar una visualización representativa de los cascos de población y del resto de edificaciones a la escala 1:25,000, evitando conflictos visuales como solapes entre edificios o entre éstos y las vías de comunicación. Recordemos la necesidad de analizar el contexto de cada elemento, en este caso de cada edificación, para decidir las estrategias de actuación sobre él más adecuadas. Por tanto, habrá que analizar tanto sus características intrínsecas como su contexto, siendo la densidad de edificación la característica de contexto más influyente. Por ello, las actuaciones sobre la capa de edificaciones están estructuradas precisamente en función de la clasificación de densidad, generando tres tipos de particiones:

Zona de baja densidad

En baja densidad de edificaciones, el objetivo es intentar mantener los elementos existentes ya que éstos pueden resultar representativos para esa área. Como no hay problemas de espacio y sí hay más posibilidades de movilidad de edificios, se realizan desplazamientos entre los edificios para asegurar la luz entre ellos (distancia mínima perceptible entre elementos).

Zona de media densidad

Esta partición es característica de las urbanizaciones y otras zonas con edificaciones aisladas agrupadas en una distancia de vecindad. Aquí, la mayoría de edificaciones aisladas se colapsan a símbolo, y como consecuencia pasan a ocupar más espacio del que ocupaban en su dimensión real, por lo que se genera apelotonamiento de símbolos de edificaciones. Para evitar este empaste visual, se realiza un filtrado en el que se eliminan algunas de las edificaciones, manteniendo la impresión visual de la zona. Este filtrado se aplica seleccionando símbolos mediante una rejilla que conserve el patrón de la zona edificada.

Ilustración 5 - 65: Salida cartográfica de edificaciones en zona de densidad media

Zona de alta densidad

Esta partición es característica de casos urbanos, donde encontramos edificaciones en forma de manzanas y menos cantidad de edificaciones aisladas. El objetivo es proporcionar una imagen nítida del casco, simplificando elementos, pero a la vez ensalzando la estructura del esqueleto urbano (el esquema esencial representativo de cada casco). En este caso, al haber poco espacio, hay poca libertad de movimiento, por lo que se emplean otro tipo de operadores, como la agrupación en edificaciones cercanas, o de recursos, como acentuar las separaciones para marcar el límite de la manzana mediante recortes con la vía urbana.

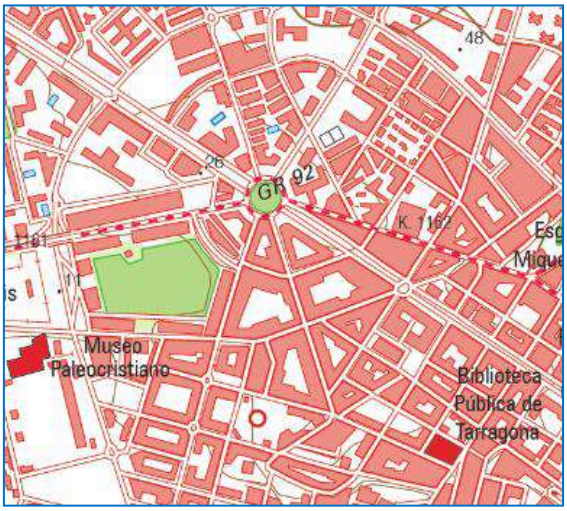


Ilustración 5 - 66: salida cartográfica de edificaciones en casco urbano

Árbol de decisión de actuaciones sobre edificaciones en función de su partición de densidad

La selección de estrategias a aplicar a cada escenario obedece a un árbol de decisión que permite utilizar unos operadores en función de las características y las circunstancias de entorno de cada edificio ya que, como se ha mencionado en varias ocasiones la redacción cartográfica debe actuar teniendo en cuenta el contexto y relaciones con otros elementos. La siguiente figura muestra el esquema del árbol de decisión implementado con FME.

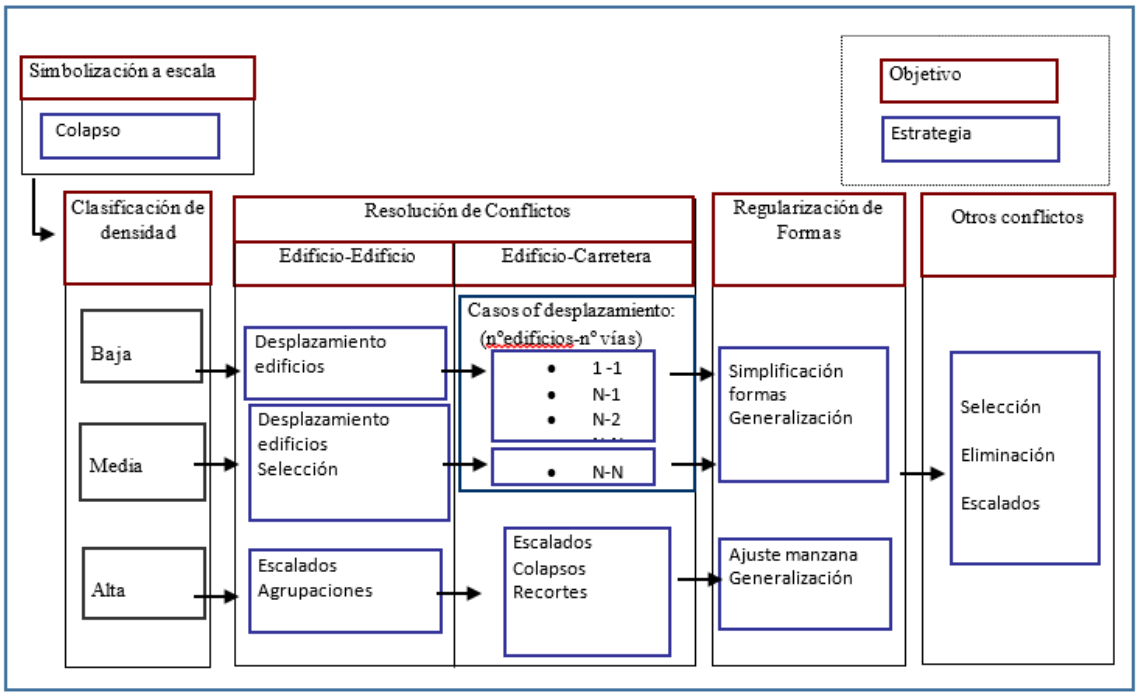


Ilustración 5 - 67: Árbol de decisión implementado para gestionar las actuaciones sobre edificaciones

Generación de particiones

Para poder tratar a los edificios según su contexto de densidad de edificación (alta, media o baja), se generan particiones del espacio de edificaciones mediante agregaciones de zonas de influencia, generando una serie de nubes que se clasifican en función del número de construcciones o tamaño de edificios. En función de la pertenencia de los edificios a estos tres tipos de clases se les someterá a distintas operaciones.

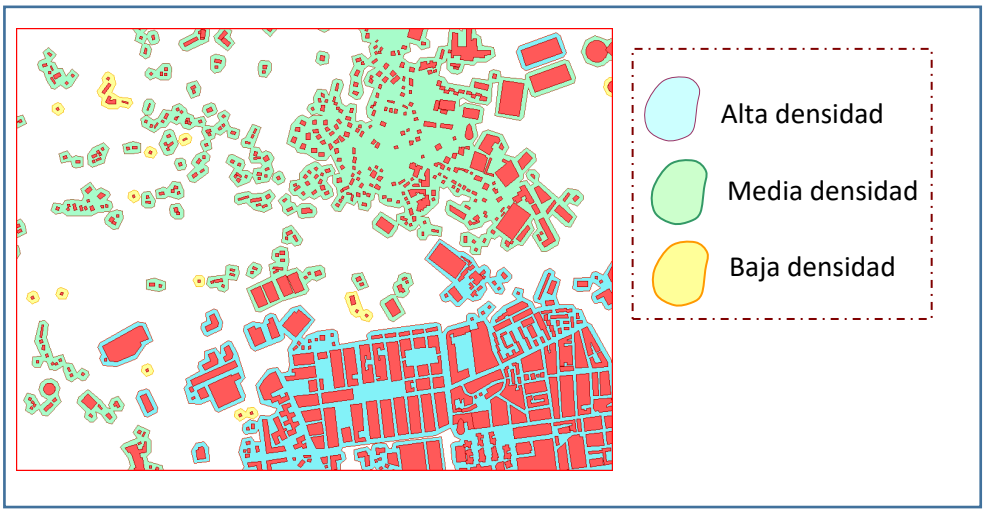


Ilustración 5 - 68: Clasificación de edificios según particiones de edificación

Una vez clasificados los edificios en particiones, se les aplica, según el árbol de decisión, las operaciones de redacción cartográfica explicadas en apartados anteriores y que podemos concretar en las siguientes:

- Desplazamientos de edificios de menor tamaño para evitar solapes entre edificios o entre edificios y vías.
- Colapsos de los edificios que no tienen una superficie representativa a la escala de visualización.
- Rotación de edificios de hasta una cierta magnitud que se encuentran en las inmediaciones de una vía de comunicación para que su dirección principal quede paralela a dicha vía.
- Selección de edificios que pueden ser eliminados por su extensión o por su categorización, y que entran en conflicto con otras entidades. Por ejemplo, en zonas de urbanización se realiza un filtrado de edificios, aplicando un patrón de selección mediante coincidencia con una rejilla definida, obteniéndose así una representación simplificada de la urbanización.
- Amalgamado más recortes a los edificios que se encuentran en la partición de alta densidad, es decir, edificios en cascos de población. En este caso se aplican operaciones de amalgamado y recorte con el ancho de vía principal. Para eliminar la luz, se realiza una agrupación o amalgamado de los edificios que se encuentren a menos de esa distancia. Y para recuperar y ensalzar vías urbanas principales que puedan quedar ocultas con los edificios, se realizan operaciones de recorte de los edificios con el ancho de dichas vías (clipper).
- Expansión a manzana que se aplica a los edificios en alta densidad para eliminar las pequeñas diferencias que pudiera haber entre el borde de una agrupación de edificios y el borde de la manzana.

5. *Altimetría*

En este subproceso se efectúa una selección de puntos de cota que se consideren representativos para el terreno.

6. *Tratamiento de símbolos puntuales*

El proceso de tratamiento de simbología puntual se encarga de las siguientes operaciones:

- colapsos de las superficies que no tienen representación superficial, sustituyéndolas por su simbología puntual. A partir de estos símbolos se realizan selecciones en función de su jerarquía para evitar conflictos visuales o solapamientos de símbolos puntuales.
- recolocaciones de algunos de estos símbolos, por ejemplo, se colocan los hitos de carreteras en el centro de las vías de comunicación, y las estaciones de servicio paralelas a sus correspondientes vías.
- aplicación de patrones de símbolos que aportan representatividad a ciertas áreas, por ejemplo, el relleno del área minera con un patrón de símbolo de mina, o el relleno de superficie de cementerio con un símbolo de cruz.

7. *Proceso automático de rotulación*

Este proceso permite generar rótulos a partir del atributo Etiqueta proveniente de BTN, con fuentes y tamaños acordes a la leyenda de estilos de MTN25, y atendiendo a las reglas establecidas en las directrices del apartado de etiquetado automático.

Se obtiene enteramente bajo el motor de FME, mediante el módulo *MapTextLabeller*, que como se muestra en la siguiente figura, permite definir, para cada tipo de elemento, su fuente, tamaño y regla de etiquetado.

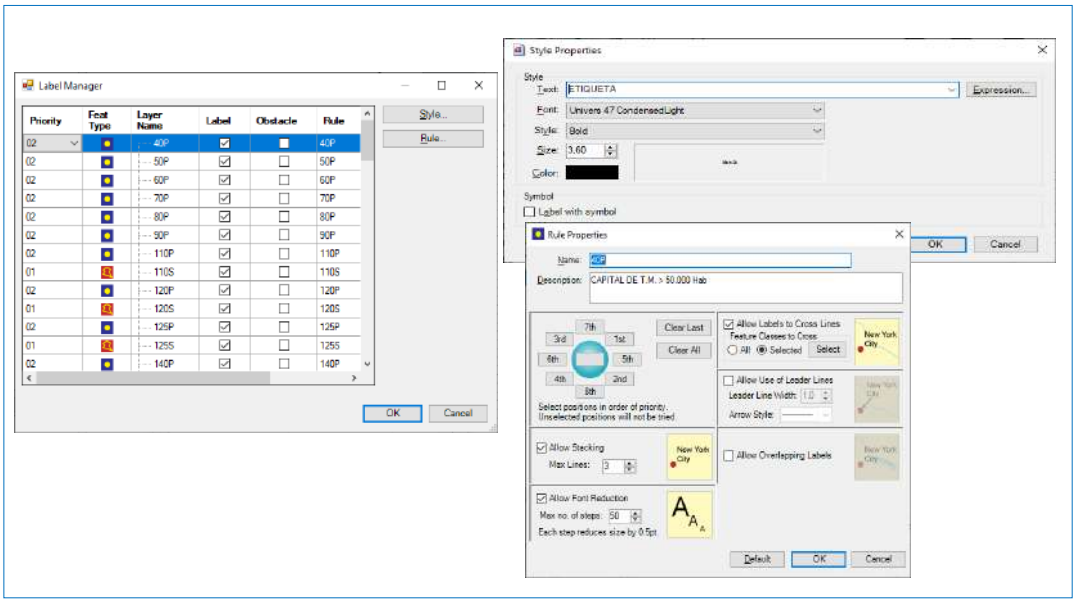


Ilustración 5 - 69: Parametrización de rotulado mediante el módulo MaptextLabeller de FME

Previamente al proceso de etiquetado, se aplica una serie de **subprocesos** que preparan los elementos para el etiquetado automático:

- Clasificación de las etiquetas según su leyenda. Por ejemplo, las entidades de población son analizadas para asignarles un estilo en función de su capitalidad y número de habitantes, o en el caso de la hidrografía, se asigna un estilo en función de su jerarquía.
- Selección de elementos a etiquetar, ya que no todos los elementos son etiquetables, por ejemplo, rótulos genéricos o cotas de curvas no maestras.
- Detección de etiquetas duplicadas, para evitar que aparezcan rótulos iguales referidos al mismo fenómeno.
- Sustitución de una selección de textos por sus abreviaturas.
- Generación de límites y elementos que no deben ser pisados (marco de la hoja o símbolos).

8. *Incorporación de otras fuentes de origen*

Este paso se realiza justo antes de almacenar la salida en la base de datos cartográfica. El proceso se conecta a la base de datos que almacena SIOSE, la base de datos de ocupación del suelo en España a escala 1:25,000, para incorporarlos a la salida cartográfica. En este punto también se aplica una operación de conflación geométrica para hacer coincidir las fronteras de bordes de las áreas debido a la falta de coincidencia total de las geometrías de SIOSE y las cartográficas.

En el caso de límites administrativos, el proceso extrae la información procedente de la base de datos de líneas límite, y tras una traducción del modelo, la incorpora directamente en la salida, sin realizar ningún tipo de operación de generalización, ya que este tipo de geometrías debe representarse en su verdadera posición. Los posibles conflictos con otros elementos se resuelven en la parte de tratamiento de simbolización, aplicando transparencias.

9. Almacenamiento en la base de datos cartográfica

El punto final de este flujo es el almacenamiento de las geometrías transformadas en la BBDD cartográfica, cuyo diseño se describe a continuación.

Definición de tablas

Para cumplir los requisitos anteriormente mencionados se elaboró el diseño de una BBDD mediante la siguiente estructura de tablas y atributos:

Tablas que contienen geometrías

Las geometrías se estructuran en tablas clasificadas **por tema** y **por tipo de geometría**. La clasificación por temas está basada en los dígitos de la clasificación de los 10 temas de BTN: ej. 01 para límites administrativos, 02 altimetría, 03 hidrografía, etc., más el 11 para SIOSE. El tipo de geometría contempla los tipos P para puntual, L para líneas y S para superficial. Como se explicó anteriormente, la base de datos Postgis no almacena textos, por lo que serán almacenados en tablas de tipo puntual con sus correspondientes atributos.

La denominación de las tablas se obtiene del cruce de Tema y Tipo de geometría, dando lugar a los nombres de tablas 01L, 01P, 01S, 02P, 02L, etc., como muestra la siguiente tabla:

TEMA/ TipoGeom	01	02	03	04	05	06	07	08	09	10	11
L	01L	02L	03L	04L	05	06L	07L	08L	09L		11L
P	01P	02S	03P	04P	05P	06P	07P	08P	09P	10P	
S	01S		03S	04S	05S	06S	07S		09S		11S

Tabla 5-3: Denominación de las tablas en función del Tema y Tipo de geometría

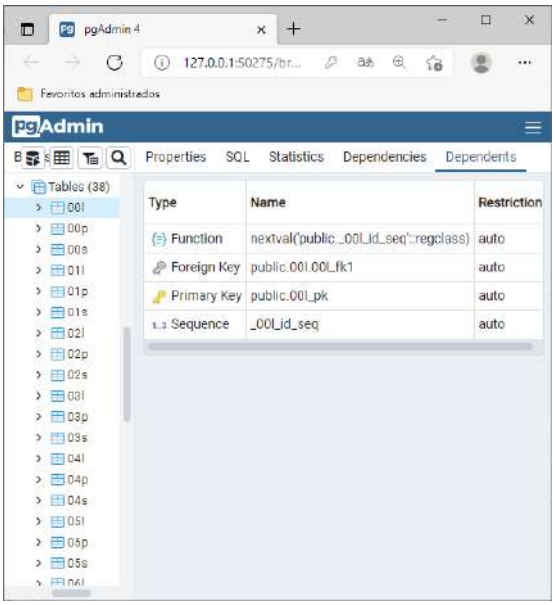


Ilustración 5 - 70: Detalle de la estructura de tablas almacenada en Postgis

Tabla con el registro de las hojas almacenadas

La tabla "*hoja_c*" se crea para registrar las hojas almacenadas en la base de datos y permitir la vinculación de cada una de ellas con las geometrías contenidas.

Esta tabla almacena para cada registro la información de cada una de las hojas: el bounding box de la geometría de su marco y una serie de atributos que la detallan (número de hoja, nombre, fecha de alta, etc.). Para vincular cada hoja a las geometrías se genera además un atributo identificador (atributo id) que está referenciado en cada una de las geometrías almacenadas mediante el atributo id_hoja_c (Ej. de relación: hoja_c.id=01L.id_hoja).

Definición de atributos

Una vez definidas las tablas, procedemos a definir sus atributos, también teniendo en cuenta los requisitos anteriormente descritos.

Tablas geomP	Tablas geomL	Tablas geomS	Descripción
id	id	Id	
Id_hoja_c	Id_hoja_c	Id_hoja_c	Referencia al id de la tabla hoja_c
nncccppe	nncccppe	nncccppe	Identificador para proceso posterior de simbolización
cota	cota	Cota	Sólo se rellena para para altimetría tema 02
geom	geom	geom	
escala			
orient			
valor			Sólo para toponimia tema 08. Contenido del texto

Tabla 5-4: Atributos almacenados para cada geometría

Atributos tabla id_hoja_c	Descripción
id	Identificador autonumérico único. Se asigna a cada elemento almacenado mediante id_hoja_c
num_hoja	Se saca del nombre del archivo
fecha alta	Fecha en la que fue subida la hoja a Postgis (cuando se ejecutó programa para esa hoja)
procesador	Nombre del equipo desde el que se procesó la hoja
tipo_hoja	CULTIVOS/MAPA
huso_marco	Proviene del marco que efectuó el corte
nombre_marco	Proviene del marco que efectuó el corte
geom_marco	Geometría del marco que se usó para el recorte en geográficas

Tabla 5-5: Atributos almacenados para la información de cada hoja

5.3. Mapa automático con edición adicional

La generación automática de la cartografía satisface los requisitos de los usuarios que prefieren un producto más actualizado, aunque tiene la desventaja de presentar una pérdida de calidad en la estética final, si lo comparamos con el resultado del mapa tradicional elaborado de forma manual.

El hecho de que la generación sea automática implica que puede conllevar determinados errores. Haciendo referencia a la importancia de la toponimia en la cartografía, uno de los procesos más complicados de realizar es el etiquetado automático. Este problema se hace más visible en el caso de que haya un elevado número de textos en una zona, debido a que el algoritmo intenta colocar todos los textos y en el resultado, a veces, se generan solapes entre ellos o ausencias de su etiquetado, dando lugar a un resultado no satisfactorio conforme a las especificaciones del producto cartográfico.

Este aspecto es especialmente relevante en las ediciones impresas del mapa, cuya adquisición generalmente supone una aportación económica del usuario. Por este motivo, es razonable considerar **un protocolo adicional al automatismo** con objeto de incrementar la calidad del resultado final mediante la edición de los conflictos que el proceso no resolvió de forma satisfactoria. Además, para la edición impresa, es necesario generar la hoja completa, añadiendo unos exteriores, que también son generados automáticamente.

En esta sección se describe la solución adoptada por el IGN de España para abordar este caso de uso, y que es la metodología denominada "**MTN25 Edición Simplificada**".

5.3.1. Objetivo

La denominación "Edición Simplificada" se debe a que en este proceso no se hace una corrección completa y exhaustiva conforme a todas las indicaciones recogidas en el documento de las Normas de Edición del MTN25, sino que se limita a aquéllas que, por su transcendencia, se considera de vital importancia que aparezcan corregidas, además de la subsanación de los errores producidos por el automatismo.

El objetivo de esta metodología es obtener de forma eficaz y eficiente ediciones para la publicación impresa de la serie del MTN25, mediante un proceso optimizado en tiempo y recursos.

El alcance del proyecto contempla tres ámbitos:

- la edición del contenido de cada hoja del mapa generado automáticamente, siguiendo los ítems establecidos en un protocolo (check-list).
- la composición de la hoja completa con los exteriores (generados también automáticamente).
- y la edición de dichos exteriores, siguiendo igualmente un check-list.

Se trata, por un lado, de mejorar la edición resultante del proceso automático (del contenido y de los exteriores) según la revisión de determinados ítems, empleando como referencia los datos de partida, la BTN, no datos adicionales capturados del mundo real ni la ortofotografía. Es decir, se realizan únicamente trabajos de edición y no se lleva a cabo tareas de

actualización, ya que se parte de la premisa de que la información está actualizada o al menos lo está en el mayor grado en que ha sido viable. El flujo de trabajo se focaliza en:

- solucionar los errores de edición que hayan podido producirse en el proceso de generación automática del mapa, como pueden ser las pérdidas de geometrías (ej. tramos de carreteras), las pérdidas de textos (ej. sitios) o la alteración de elementos.
- editar aquellos elementos en los que el proceso todavía no ha alcanzado el grado de legibilidad y de calidad deseado, como pueden ser los solapes de textos y de simbologías, la orientación de simbologías, etc.

Por otro lado, de componer la hoja completa con los exteriores, generando de esta forma una nueva edición del MTN25, ya sea para su impresión en papel o para cualquier otro tipo de servicio WEB.

5.3.2. Proceso de producción

5.3.2.1. Archivos de partida y finales

Las hojas del MTN25 a producir se distribuyen en paquetes o bloques de hojas, llamados **entregas**. Es un conjunto de hojas que se ejecutan bajo las mismas condiciones, es decir, con el mismo operador y revisor. De cara a trabajar con asistencia técnica externa, se obtienen resultados satisfactorios al validar los resultados con inspecciones por muestreo. Éstas suelen contener alrededor de 15 hojas.

Los **archivos de partida** para cada hoja son:

- El resultante de la generación del contenido de mapa automático correspondiente a planimetría y altimetría, que es el que debe editarse siguiendo el protocolo.
- El correspondiente a la cobertura del suelo, que no se debe editar.
- Archivo BTN correspondiente a la hoja.
- El archivo de exteriores (FFS) que también se genera automáticamente, y que tras una mínima edición de fusionará con el contenido para generar la hoja completa.
- Sombreado generado para cada hoja desde el MDT.
- Información adicional para colocar en la portada y contraportada (código de barras).
- Ortofotografía. Necesaria para ayudar a la captura de elementos perdidos en el proceso y para incorporar aquellos que por su importancia deban ser recogidos, aunque no aparezcan en BTN.

Los **archivos finales** obtenidos:

- El correspondiente al resultado de editar la salida de mapa automático con el check-list (planimetría y altimetría).
- Archivo con la cobertura del suelo, que sería el mismo que el archivo de partida si sobre él no se ha hecho ningún tipo de modificación.
- Archivo resultante de la generación automática de un FFS a partir del dgn editado y el dgn de SIOSE.

- PDF final, resultante de la generación automática de un PDF a partir del FFS de contenido, FFS de exteriores, sombreado y código de barras.
- Protocolo (check-list o Registro de Calidad). Se trata de la hoja de cálculo del protocolo que se va rellenando a lo largo del proceso por un operador, un revisor y un revisor de control de calidad.

5.3.2.2. Protocolo (check-list)

La implementación de un protocolo de edición o checklist ha sido elaborada después de muchas revisiones para detectar conflictos no resueltos automáticamente de forma satisfactoria, y aquellos que por su importancia en un mapa tienen que aparecer corregidos para cumplir con las especificaciones del producto. Aunque la premisa es que el proceso parte de información actualizada, a veces puede que no sea así, y por ello el checklist debe contemplar ciertos ítems enfocados a evitar que aparezcan errores graves debidos a dicha falta de actualización de los datos. Es una hoja de cálculo en la que se encuentran todos los ítmes que deben acometerse para llevar a cabo el proceso de edición y revisión del contenido y de los exteriores, del control de calidad y el envío al departamento donde se imprimirán.

Este protocolo está pensado para la cartografía a escala 1:25,000. En función de la resolución geométrica y semántica de la cartografía que se quiera generar, así como de la zona a cartografiar, habrá que adaptarlo y modificar lo que se estime necesario.

El personal cualificado que interviene en la cumplimentación de este registro de calidad es:

- Un operador (O), que realiza las fases de edición manual: edición de contenido y edición de exteriores siguiendo el protocolo y rellenándolo (check-list: datos generales y la columna del “operador”). Además, ejecuta los procesos de obtención de los archivos necesarios para la obtención del archivo resultante final. Si fuera necesario, también realizaría las tareas de corrección en el caso de rechazo de la entrega.
- Un revisor que realiza las fases de revisión del pdf final, que son los ítems con sombreado de color naranja, siguiendo el protocolo y rellenándolo (check-list: datos generales y la columna del “revisor”). Además, realiza las tareas de 2ª revisión en el caso de rechazo de la entrega.
- Revisor de control de calidad, que supervisa que la ejecución de los ítems ha sido correcta y marca para cada uno de ellos su estado de validación. Tras ello, compila los ítems no aptos, para que, en el caso de que se rechace la entrega, se efectúe la corrección correspondiente.
- Revisor en talleres, que realiza una última revisión de la hoja siguiendo el check-list de la pestaña de talleres. Por lo tanto, modifica tanto el pdf como el check-list finales, dando por finalizado el proceso.

Los ítems son tareas concretas que deben realizarse o revisarse y marcarse como hechas, como se muestra en la siguiente figura:

		O	R	CC
22	Rotular Sierras siguiendo la lín. de BTN (hasta ambos extremos) con espacios entre letras **y si hay más de un rótulo eliminarlo.**			
23	Etiq. de Edif. Aisladas recuperadas del N63 si no dificulta legibilidad**			

MARCA OPERADOR: **X**: Revisado y corregido, o ejecutado **/**: Revisado pero No se ha encontrado ningún caso
MARCA REVISOR Y CC: C: Correcto **E:** NO LO CUMPLE

Ilustración 5 - 71: Ejemplo de ítems y roles que participan en la Edición Simplificada

Si el operador marca "X" en la columna O, indica que ha revisado el ítem y ha tenido que llevar a cabo alguna acción: ejecutar un proceso, corregir algún error, o editar algún elemento. Si por el contrario el operador marca "/" en la columna O, estará indicando que ha revisado el ítem pero que no ha tenido que hacer ninguna acción. El revisor marca con "C" o "E" al revisar los items de color naranja en el pdf final, y el revisor de cc anota si está todo correcto "C" o por el contrario no lo cumple, "E".

En función de la gravedad de los ítems incumplidos, la hoja será apta o no apta.

GRAVE
MEDIO
LEVE

Ilustración 5 - 72: Calificación de los ítems incumplidos

A la hora de trabajar con servicios externos, es importante que se rellenen los apartados del cabecero del check-list, que tiene este aspecto:

PROTOCOLO EDICIÓN Y REVISIÓN MAPABTN25 V2.6	OPERADOR CONTENIDO EMP:_____FECHA: _____Nº ACTUACIÓN:___
HOJA: _____EMPRESA:_____CTO:_____Nº ENTREGA: _____TIPO PROCESO:_____	OPERADOR EXT EMPRESA:_____FECHA: _____Nº ACTUACIÓN:___
REVISOR CC IGN:_____FECHA: _____Nº REVISIÓN:___RESULTADO:_____	REVISOR FINAL EMPRESA:_____FECHA: _____Nº REVISIÓN:___

Cada excel check-list tiene 3 pestañas:

- **Contenido_exteriores.** Es la pestaña que contiene todos los ítems a realizar y revisar del contenido y exteriores, por el operador, el revisor y el revisor de control de calidad.

13	Etiquetas de capital munic <10mil: eliminado efecto diana	X		C	38	Geometrías de FF.CC. Inte correspondientes reparad
14	Etiq. y texto de kms de vías hasta Carr. 2º, recuperados del N63	/		C	39	Geom. de Vías paralelas a
15	Texto de kms de FFCC, recuperados del N63	/		C	40	Símbolos de estación de F
16	Etiq. de Estaciones de FFCC recuperadas del N63*	/		C	41	Cambiada la clasificación Pks o nomenclatura
17	Etiq. de Parques Nacionales y Naturales recuperadas del N63	/		C	7	HIDROGRAFÍA (BTN como

CONTENIDO_EXTERIORES | IMPRENTA | RESULTADO_CC

Ilustración 5 - 73: Extracto de la pestaña de contenido_exteriores

- **Imprenta.** Es la pestaña que contiene los ítems a revisar en talleres antes de proceder a la impresión en papel.

A	B	C
	OPERADOR:_____FECHA:_____	
	C: Correcto E: NO LO CUMPLE	
	PROTOCOLO REVISIÓN PDF FINAL TALLERES	R
1	Existen todos los elementos fijos del mapa: (contenido, portada, contraportada, leyenda) y no están cortados, ni con rayas, ni con un fondo	
2	El nombre del cabecero es igual que el de la portada (este último en minúsculas)	
3	El número del cabecero es igual que el de la portada	
4	si existe vocabulario no está cortado	
5	El año de generación es el vigente	
6	Tiene depo, ISBN y NIPO	
7	Observaciones:	

CONTENIDO_EXTERIORES | IMPRENTA | RESUL

Ilustración 5 - 74: Extracto de la pestaña imprenta

- **Resultado_cc.** Se copian de cada una de las hojas que forman la entrega, los ítems que no están corregidos o revisados. También se anota el número de hojas de cada entrega, así como las fechas y todas las revisiones que ha tenido hasta que ha sido aceptada y publicada.

CC	NUM HOJAS	HOJAS CC	FECHA CC	RESULTADO	FECHA CC 2º	RESULTADO FINA
	17	17	28/05/2020			
	ID_HOJA	ITEM	DESCRIPCIÓN	CC POST REPAROS C/E		
APTO	0001-4	6	Detec.. casual (No requiere marcarlo) de elemento RELEVANTE omitido en el fichero "b" de referencia: Embalse, Río de 1º y 2º, hasta carr. de 2º incluidas.	Falta texto de edición anterior "O Cotelo", "A Xunqueira",		
NO APTO		44	Control de fugas (marcas verdes)	Aparece una fuga verde.		
		NOTA: Continuar curva batimétrica 100 interrumpida.				
		NOTA: La portada, al ser hoja especial, se hará en IGN.				
	0002-1	7	Detec.. casual (No requiere marcarlo) de Elem. en BTN y omitidos en mapaBTN, recuperados	Falta cota 100 en curva batimétrica y no se interrumpe curv		
		27	Reparo solap. text (inclu. cotas) con simb. puntual (excep edif. Aislado y piscinas)	Interrumpir curvas batimétrica en cotas 50 y 100.		
		52	Coordenadas UTM no faltan y no se solapan con coord Geog. en esquinas	Borrar coord UTM "48 54".		
		NOTA: La portada, al ser hoja especial, se hará en IGN.				
	0002-3	5	Detec.. casual (No requiere marcarlo) de elemento RELEVANTE omitido en el fichero "b" de referencia: Embalse, Río de	Inconsistencia de capas, caminos cruzan río de doble marge		

CONTENIDO_EXTERIORES | IMPRENTA | RESULTADO_CC

Ilustración 5 - 75: Extracto de la pestaña de resultados

5.3.2.3. Desarrollo, actividades y mapa del proceso

En esta sección se describen las diferentes actividades del proceso, desde la planificación hasta la revisión final previa a la impresión. Todas ellas se describen con detalle en documentación de trabajo interna del IGN, cuya consulta sería recomendable realizar en caso de necesitar profundizar más allá de la descripción que en esta sección se aporta de cada una de ellas. En la siguiente figura 5-76 se puede visualizar el mapa del proceso con todas ellas.

Dado el volumen de hojas, poco más de 4,000, a escala 1:25,000 que cubren el territorio español, es necesaria la contratación de servicios técnicos que ayuden a la producción anual, de tal forma que en periodos mínimos de tiempo se pueda tener España actualizada. Por ello se habla en el documento antes mencionado, de empresa, pudiendo sustituir los roles de descritos pertenecientes a empresas por personal propio.

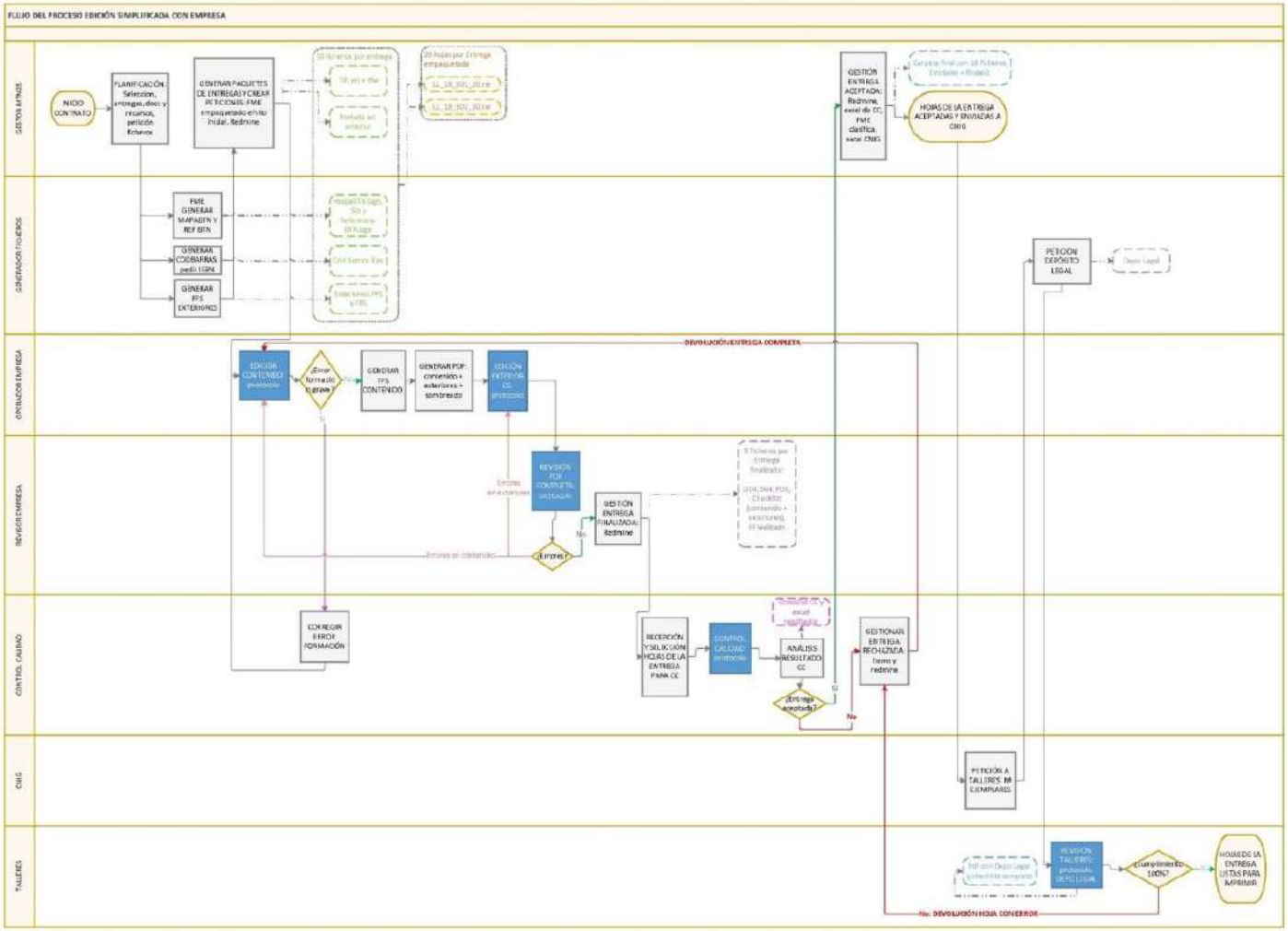


Ilustración 5 - 76: Mapa de proceso de la Edición Simplificada

Planificación y gestión del trabajo en el IGN

Semiautomática. El gestor y generador de archivos llevan a cabo las tareas de planificación y gestión de archivos y de BBDD. La comunicación con la empresa se efectúa por medio de un sistema de gestión de proyectos online, por medio de pestañas de documentación, noticias y foro, y lleva a cabo el seguimiento de las peticiones a lo largo del proceso. Los archivos de partida se cargan en este sistema, al igual que los archivos resultantes. Se crean las peticiones (tareas de "Ejecución bloque") con los archivos iniciales de cada entrega, así como el resto de peticiones que son necesarias a lo largo del proceso (validación de la entrega, corrección de la entrega, con los archivos finales y soporte para cualquier duda).

Edición del contenido de la hoja

Manual, siguiendo el protocolo. Un operador cartográfico, partiendo del archivo obtenido por el proceso de mapa automático, edita el contenido de la hoja siguiendo lo establecido en el protocolo check-list. Para ello cuenta además con un archivo con a la información de BTN (principal fuente de datos), para utilizarlo como referencia en determinadas tareas. También dispone de la ortofotografía, para cuando sea imprescindible incorporar información geográfica de especial relevancia, ausente en la fuente de datos. Este proceso está formado por 10 tareas principales:

1. Errores de formación graves

Antes de comenzar con la edición se deben revisar los siguientes ítems (1 al 4). Si en alguno de ellos existe error, se debe revisar el archivo auxiliar de referencia BTN (b). Si en este archivo el error persiste, se trataría de un error considerado como grave y se debe parar inmediatamente la hoja; es decir, se comunica al gestor para que lo valore y en caso necesario se pararía el proceso. Normalmente se actualiza la información directamente, aunque otra opción es incorporarlo en BTN, y volver a generar una salida de mapa automático.

1	**ERRORES GRAVES EN LA INFO. ORIGEN (en el archivo "b" de referencia):** SE PARA
1	Hasta Carr. de 2º incluidas: sin texto asociado, en el archivo "b" de referencia
2	Hasta Carr. de 2º incluidas: omisión tramos (>1km), en el archivo "b" de referencia
3	Ríos de doble margen y Embalses: sin texto asociado, en el archivo "b" de referencia
4	Ferrocarril: omisión tramos (>1km) en el archivo "b" de referencia
5	Detec. casual (No requiere marcarlo) de elemento RELEVANTE omitido en el archivo "b" de referencia: Embalse, Río de 1º y 2º, hasta carr. de 2º incluidas, Núcleos de población, textos de Mar, País e Islas grandes, Parque Nac o Nat. Formato de etiqueta de Caminos con artículo al final (CAMINO (DE**), inconsistencia de vías de comunicación con relieve, en general inconsistencia de capas.**

Ilustración 5 - 77: Extracto de ítems relativos a errores en la información origen

2. Composición

A continuación, se debe poner de referencia las líneas límite administrativas para ver si en el proceso de mapa automático se han incluido correctamente. Por otro lado, se encuentra el ítem en el caso de que casualmente se detecte algún elemento en BTN de referencia que se

haya omitido o haya sido mal resuelto durante el proceso automático de generación del contenido, que en el caso de detectarse y capturarse se macaría con “X”, y en caso contrario se deja vacío.

2	**COMPOSICIÓN**
6	Faltan todas o **algunas** líneas límite
7	Detec. casual (No requiere marcarlo) de Elem. en BTN y omitidos **o mal resueltos** en mapaBTN

Ilustración 5 - 78: Extracto de ítems relativos a composición de la hoja

3. Proceso previo

El objeto de los siguientes ítems es darle el aspecto que va a tener la información geográfica en la cartografía. Con una tabla de características, simbolizará determinados elementos, pondrá un ancho determinado a las vías de comunicación, rellenará edificios, dibujará con estilos de línea personalizados la hidrografía y algunos elementos lineales, etc. Debemos tener en cuenta que la salida de mapa automático es un archivo con líneas y contornos, y para poder editar correctamente tenemos que ver cómo va a quedar realmente. Se trata de ítems que deberán de estar marcados siempre con “X”, ya que son tareas en las que siempre se debe de actuar.

La salida del proceso de mapa automático para MTN25 es un archivo con extensión dgn, de forma que el entorno de edición está desarrollado sobre software de CAD MICROSTATION.

3	**PROCESO PREVIO**
8	BTN25: Simbol. de Elem. Lineales pequeños
9	BTN25.1: Proceso de Hidrografía
10	BTN25.1: Proceso de Construcciones
11	BTN25.1: Simbolizar los elementos lineales

Ilustración 5 - 79: Extracto de ítems relativos al proceso previo

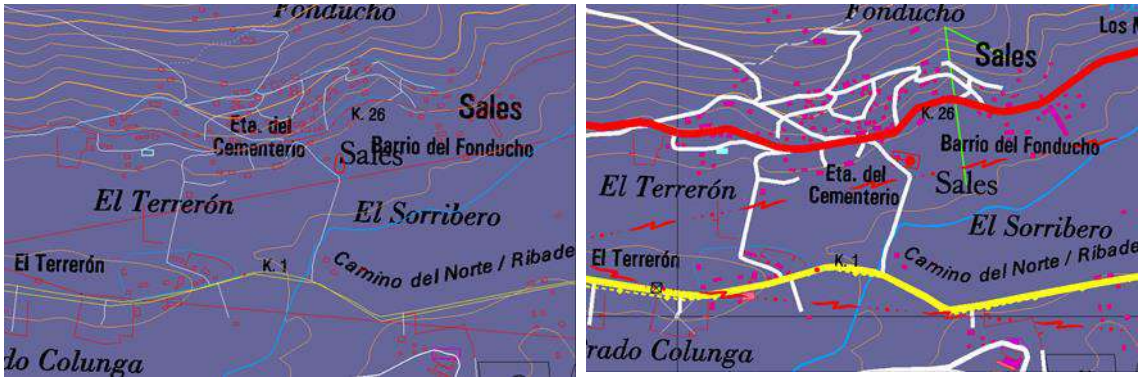


Ilustración 5 - 80: Antes y después de ejecutar los ítems del proceso previo

4. Edición de textos

En relación a los textos se deben de llevar a cabo los siguientes grupos de tareas:

- Un trabajo de recuperación de los textos que se han perdido durante el proceso de generación del mapa automático. El etiquetado automático es uno de los algoritmos más complicados y cuando el espacio disponible es insuficiente no rotula algunos textos. Estos textos son seleccionados y se copian en una capa o nivel determinado (63). Toda la información que aparezca en este nivel no se mostrará ni será visualizada en el archivo final. Al realizar esta tarea, hay que tener una serie de consideraciones:
    - hay que asegurarse de que no existe duplicidad en la hoja, es decir, que el texto aparentemente perdido y almacenado en el N63 no se encuentre también en la hoja a editar.
    - una vez seleccionado el texto a recuperar, se le debe asignar la fuente correspondiente.
    - finalmente, editar según las normas.
- Rotulación de determinados elementos según las normas:
    - eliminar el “efecto diana” en capitales de municipio menores de 10,000 habitantes, es decir, tratar de evitar que el caso de población quede tapado por el rótulo, desplazando fuera del casco, los rótulos que ocupen la mayor parte del éste.
    - editar las sierras (no solo las recuperadas) siguiendo la línea de rotulación de referencia (línea base recta para cada palabra, no curva), e introduciendo espacios entre letras en vez del espaciado entre caracteres, y si hay más de un rótulo de sierra, eliminarlo dejando sólo uno.
- Reparo de solapes entre textos y con simbologías. Para ello se debe desplazar los textos, evitando mover textos de cotas de curvas.
- Eliminación de textos de determinados genéricos y eliminar determinados textos dentro de un casco de población: los casos más habituales son ES (estación de servicio), EF (estación de ferrocarril) y Helipuerto, pero pueden existir casos de Camping, Central solar, Área recreativa, etc.
- Recuperación de ediciones anteriores. Se trata de poner como referencia un archivo de textos de la edición anterior, que incluye orografía, hidrografía y sitios, y otro archivo que contiene barrios en grandes ciudades y barrios menores, y recuperar lo que falte. Este ítem solo se realizará si existiese una edición anterior de la hoja.

4	**TEXTOS**
12	Etiquetas de EEPP recuperadas del N63
13	Etiquetas de capital munic <10mil: eliminado efecto diana
14	Etiq. y texto de kms de vías hasta Carr. 2º, recuperados del N63
15	Texto de kms de FFCC, recuperados del N63
16	Etiq. de Estaciones de FFCC recuperadas del N63*
17	Etiq. de Parques Nacionales y Naturales recuperadas del N63
18	Etiq. de Embalses recuperadas del N63
19	Etiq. de Ríos de doble margen recuperadas del N63
20	Texto de kms de Canal recuperados del N63
21	Etiq. de Sierras recuperadas del N63
22	Rotular Sierras siguiendo la lín. de BTN (hasta ambos extremos) con espacios entre letras **y si hay más de un rótulo eliminarlo**
23	Etiq. de Edif. Aisladas recuperadas del N63 si no dificulta legibilidad**
24	Etiq. sitios recup. del N63 si no dificulta legib (priorizar orónimos)
25	Resto de etiquetas recuperadas del N63 si no dificulta legibilidad
26	Reparo de solape de textos con textos (inclu. cotas)
27	Reparo solap. text (inclu. cotas) con simb. puntual (excep edif. Aislado y piscinas)
28	Eliminado etiq. de símb. Punt si solo tiene genérico (ejm.ES, EF, Helipuerto)
29	Eliminados textos en cascos que no sean textos de elem. singulares***
30	Textos de pto de cota que daban lugar a confusión, atacando al detalle
31	EEPP recuperadas con el CCEE
32	Textos de elem. hidrog. de costa que dan lugar a confusión colocados
33	**Revisar la falta de toponimia comprobando con la edición anterior y recuperarla**

*Fuera de casco: se recupera del N63 (si hay). En el interior de un casco: se recupera (si hay) si no coincide con el de la población

** Priorizar elementos representativos: Restos Arqueológicos, monumento, castillo, cueva, edif relig, torre vigía histórica

*** bic, monumentos, museos, parques, puentes y otros no bic representativos como edificios religiosos

Ilustración 5 - 81: Extracto de ítems relativos a la toponimia

5. Inconsistencias

Hay que revisar y solucionar las inconsistencias entre capas: no pueden aparecer curvas de nivel ni vías de comunicación dentro de masas de agua.

5	**INCONSISTENCIAS**
34	Inconsistencias de láminas de agua, ríos, embalses y lagunas con vías de comunicación y altimetría.

Ilustración 5 - 82: Extracto de ítems relativos a inconsistencias

Ilustración 5 - 83: Ejemplos de inconsistencias entre la hidrografía y las comunicaciones en la fuente origen

6. Edición de simbologías

En relación a las simbologías se debe llevar a cabo el cambio de orientación de las cruces del cementerio que pueden salir invertidas, y solventar el reparo de los posibles solapes entre símbolos puntuales, que dificultan la legibilidad.

6	**SIMBOLOGÍAS**
35	Cruz cementerio orientado correctamente
36	Reparo de solape de símbolo puntuales entre sí

Ilustración 5 - 84: Extracto de ítems relativos a simbologías

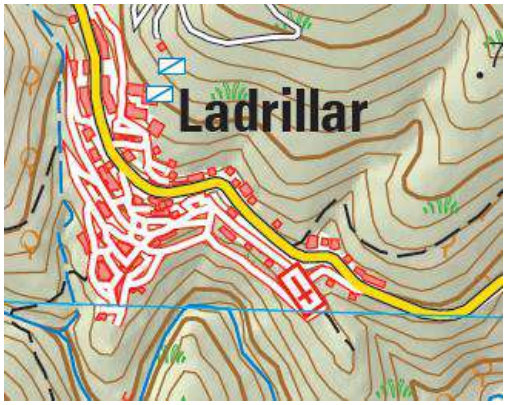


Ilustración 5 - 85: Símbolo de cementerio invertido

7. Edición de vías

Hay que realizar una revisión a lo largo de las vías de comunicación, carreteras y ferrocarriles, en busca de geometrías incompletas (trazado de vías o rotondas, puentes, túneles) o solapes entre vías. Generalmente estos errores son debidos a fallos en el proceso automático. Habrá que comprobar también que haya coherencia entre los puntos kilométricos en todas las vías, es decir, que no se repitan y sean correlativos.

7	**VÍAS DE COMUNICACIÓN (BTN como referencia)**
37	Geom. hasta Carr de 3º interrum. ó falta y sus enlaces ,nudos, rotondas, accesos de cualquier otra vía de comunicación, puentes y **Pks (coherencia de Pks)** reparados.
38	Geometrías de FF.CC. Interrumpidas, enlaces, puentes y **Pks (coherencia de Pks)** reparados.
39	Cambiada la clasificación de Pistas a Carretera de 3º cuando tiene Pks o nomenclatura

Ilustración 5 - 86: Extracto de ítems relativos a vías de comunicación

8. Cortafuegos

Los archivos son proporcionados por el archivo de coberturas del suelo, pero hay veces que no aparecen todos. Se considera que, dada la importancia de este elemento, es necesario recuperarlos, aunque no existan en las fuentes de datos de partida. Para ello será necesaria la ortofotografía como referencia, además de la BTN.

8	**CORTAFUEGOS (BTN como referencia)**
40	**Recuperar cortafuegos que faltan**

Ilustración 5 - 87: Ítem relativo a cortafuegos

9. Edición de la hidrografía

Uno de los errores que puede producir el proceso automático es que los canales aparezcan con geometrías incompletas, es decir, que falten algunos tramos. Habrá que revisar la geometría de todos los canales. Con el resto de la hidrografía no se han observado errores de consideración.

9	**HIDROGRAFÍA (BTN como referencia)**
41	Geometrías de Canal interrumpidas reparadas
42	En hojas con Mar, Isla o Nación limítrofe: Texto incluido (aunque en el archivo "b" de referencia no se encuentre)

Ilustración 5 - 88: Extracto de ítems relativos a hidrografía

10. Control

Las “fugas” suelen ocurrir cuando existe una línea que tiene ángulos muy agudos, y cuando existen 2 puntos de una línea con las mismas coordenadas (puntos dobles), entre otras causas impredecibles. Se genera una línea parecida a un lazo.

En el IGN, hay una utilidad en el menú personalizado de Microstation que nos ayuda a detectar dichas fugas. En el PDF final es dónde realmente vamos a asegurar que el producto final no tenga fugas, porque es donde se visualizan mejor. La experiencia nos dice que es mejor no utilizar en la captura de la información geográfica elementos tipo arco, curvas o B-spline, sino que se debe capturar punto a punto (por ej. vías), y en el caso de capturar una rotonda se debería de hacer con polígono (por ej. con 20 puntos)

Unos de los ítems más importantes es el **control de hojas**, sobre todo para comprobar que no hay textos sin convergencia. Los textos que no van girados (a lo largo de elemento), tienen que ser paralelos al marco de la hoja, es decir, tienen que tener un pequeño ángulo de rotación correspondiente a la convergencia de la hoja.

Dentro de Microstation hay un menú que nos genera un archivo con todos los elementos en los que se ha detectado algo erróneo. Dependiendo del tipo de error, algunos de estos elementos no aparecerán en el PDF final, y la falta de esta información pasará desapercibida, de ahí la importancia de revisar bien este ítem.

10	**CONTROL**
43	Control de fugas (marcas verdes)
44	**Control de hojas**

Ilustración 5 - 89: Extracto de ítems relativos a controles a efectuar en la hoja

Generación del FFS del contenido de la hoja

Una vez concluida la fase de edición del contenido, los archivos se simbolizarán y maquetarán para la impresión final en la aplicación MapPublisher. Para ello es necesario pasar los archivos dgn a un formato compatible con dicha aplicación. Los archivos necesarios serán el dgn del contenido de la hoja y el de coberturas del suelo.

Para este cambio de formato se ha desarrollado un proceso en FME. El formato de salida es un formato de intercambio de FME, compatible con MapPublisher, con extensión .FFS, que contiene los atributos necesarios para la simbolización de los elementos del mapa.

Generación del PDF con exteriores y sombreado

En este apartado se genera el archivo final de la hoja correspondiente, a escala 1:25,000. El software empleado es MapPublisher, y además del FFS del contenido de la hoja serán necesarios también, el FFS de exteriores, el sombreado y el código de barras.

Edición de los exteriores en el pdf

De la misma forma que se han revisado los Ítems en el contenido de la hoja hay que hacerlo también en los exteriores una vez que se ha generado el PDF.

II	EXTERIORES
1	**REVISIÓN CONTENIDO EN EL PDF**
45	Error generalizado capas SIOSE **SE PARA**
46	Falta fondo azul relevante (mar, embalse, ríos 1º) **SE PARA**
47	líneas límite coherente con cuad. De div admin **SE PARA**
48	No hay FUGAS o picos graves (no se detectan todas con el control)
49	Editado el sombreado si excede el tamaño de su máscara
50	Comprobado Horizon. textos (nombre cabecero y texto int =0 grad)
51	Textos que salen del borde reparados
52	**Líneas y coordenadas UTM no faltan y éstas últimas no se solapan con coord Geog. en esquinas**
2	**DATOS DE LA HOJA**
53	Año generación correcto y coincidencia del año de edición con el año de declinación magnética.
54	Número de edición introducida correctamente
55	Datos correctos en cabecero: Nombre y número
56	ISBN y NIPO introducido
3	**PORTADA CONTRAPORTADA y LEYENDA**
57	Nombre, num en Portada, y fila columna = que en Cabecero (minúsc)
58	**Existe pto situación portada y la pos. es correcta**
59	Cuarto de la hoja coloreado en el cuad. de Div. Admin.
60	**Editado el Listado de Térm. Municip y columnas centradas**
61	**Correspondencia entre nº de Ter. Munic y nº en el cuadro de Div. Admin y comprobar líneas de llamada**
62	Reparo de solapes números de cuad. Div. Admin.(ref. portada anterior)
63	Cuadro de Terr muni: Provincias en mayúsculas y centrada en la zona
64	Editado el Vocabulario si no sale dentro de hoja cortada
65	Eliminado o editado Vértices Geod. Si no coincide con interior
66	Existe coherencia entre orientación flecha NM y Declinación
67	**Números de las hojas del 50,000 en posición correcta en el cuadro del millón.**
68	En hojas con batimetría existe texto de "Información batimétrica: Insti…"
69	**Centrados los títulos de los componentes de la portada y contraportada**

Ilustración 5 - 90: Extracto de ítems relativos a revisión del contenido y exteriores

Todos los errores en estos ítems son considerados de vital importancia, sobre todo los de color rojo y naranja.

Observaciones a los ítems:

- Error generalizado capas SIOSE: si falta o aparece sin cobertura de usos del suelo una parte o la hoja en su totalidad, habrá que analizar y solucionar el problema.
- Líneas límite coherente con cuadro de divisiones administrativas: se refiere a que no haya coherencia entre las líneas límite de la contraportada con el interior. El error sería del proceso de generación de exteriores.
- Editado el sombreado si excede el tamaño de su máscara: cuando el sombreado sobresalga del borde exterior del marco. En ese caso se debe extender la máscara blanca que lo cubre para que quede dentro.
- Comprobado horizontalidad textos (nombre cabecero y texto int =0grad): se refiere a que los textos que van paralelos a la base del marco tienen que tener rotación=0º, por eso hay que comprobar uno del exterior (el nombre de la hoja) y otro del contenido (un sitio).
- Se debe comprobar que el año de generación de la hoja, número de edición, nombre y número de hoja son correctos. También que el año de edición coincida con el año de la declinación magnética.
- Existe punto situación portada. Y en costa la posición es correcta: se refiere a que en hojas normales se compruebe simplemente que existe punto, y que en hojas de costa, además, está más o menos bien ubicado (que no esté en el mar, por ejemplo).
- Editado el Listado de Términos Municipales (y si es 1 columna: centrarla): se centra con respecto al título “TÉRMINOS MUNICIPALES”.
- Correspondencia entre nº de Términos Municipales y nº en el cuadro de Divisiones Administrativas: se refiere a que haya correspondencia entre la numeración
- Cuadro de Términos municipales: Provincias en mayúsculas y centrada en la zona: se refiere a que si aparece el nombre de la provincia en el gráfico, debe colocarse centrada con respecto a la zona a la que hace referencia.
- Reparo de solapes de números del cuadro de divisiones Administrativas: se refiere a que la numeración sea legible. Si es necesario se introducen llamadas para sacar el número fuera del recinto.
- Eliminado o editado Vértices Geodésicos si no coincide con interior: hay que comprobar que existe coherencia en el nombre y cota; en el caso de que no haya se debería, en principio, al proceso de generación de exteriores. Si aparecen símbolos de copyright en la leyenda, hay que quitarlo.
- Existe coherencia entre orientación flecha NM y Declinación: el NM (norte magnético) tiene que estar orientado en la dirección que indica el texto.

Revisión del PDF final

Una vez finalizada la hoja y cumplimentados todos los campos del check-list de la columna "operador", se pasa a la fase de revisión final. Se revisarán los ítems que se encuentren sombreados en color naranja.

Se deben seguir igualmente los ítems establecidos en el protocolo, dando por finalizada la hoja en el caso de que esté todo correcto, o enviándola de nuevo al operador para correcciones.

Control de calidad de la entrega por muestreo

Dado el elevado número de hojas de la serie del MTN25 en España, se consideró necesario efectuar un control de calidad por muestreo. También, por la cantidad de trabajo fue necesaria la contratación de asistencia técnica externa, estando los roles de operadores, revisores en servicios externos, separados de los roles de control de calidad en el IGN. En caso de que no se contratara ayuda externa, estos roles tendrían que estar separados en diferentes secciones dentro de la propia organización. Una vez finalizada una entrega, el departamento de control de calidad en el IGN, a efectos de evaluar la aceptación o rechazo de la entrega, revisa todos los check-list por si están incompletos o incorrectos, en cuyo caso se rechaza directamente. Si no, se sigue adelante seleccionando entre un 25 y un 100% de la entrega, dependiendo del operador que haya llevado a cabo las tareas de edición y del tipo de hojas.

El caso real se acerca en muchas ocasiones a la situación en la que hay que revisar la totalidad de las hojas para que no pasen inadvertidos errores, muchas veces, graves.

Se lleva a cabo el control de calidad del pdf final la hoja, revisando todos los ítems, Indicando para cada ítem si cumple "C", o no cumple "E", y finalmente, indicando si es Apta o No Apta la hoja en función de los ítems incumplidos. Resultaría No Apta en los siguientes casos:

- 1 o más ítems de tipo 3º grado (en rojo)
- 3 o más ítems de tipo 2º grado (en naranja)
- 6 o más ítems de tipo 1º grado (en verde)
- 2 o más ítems de tipo 2º grado + 3 o más ítems de tipo 1º grado
- 1 o más ítems de tipo 2º grado + 5 o más ítems de tipo 1º grado

En el caso de que en la entrega haya resultado alguna de las hojas No Apta, se RECHAZA la entrega (es decir, todas las hojas de la entrega), haciendo la devolución de la entrega completa e indicando los ítems (no las hojas) causantes del rechazo, de forma que hasta que no se corrijan por los operadores (en todas las hojas de la entrega) y se compruebe en una 2º validación que se cumplen dichos ítems, no se aceptaría el paquete. En el caso de que todas las hojas sean Aptas, se ACEPTA la entrega.

Revisión del pdf final en talleres

Una vez que el PDF final es enviado al departamento de impresión en papel, se vuelven a revisar algunos ítems que por su transcendencia e importancia es fundamental que aparezcan correctos. Corresponden a la pestaña "imprenta" del excel del chek-list.

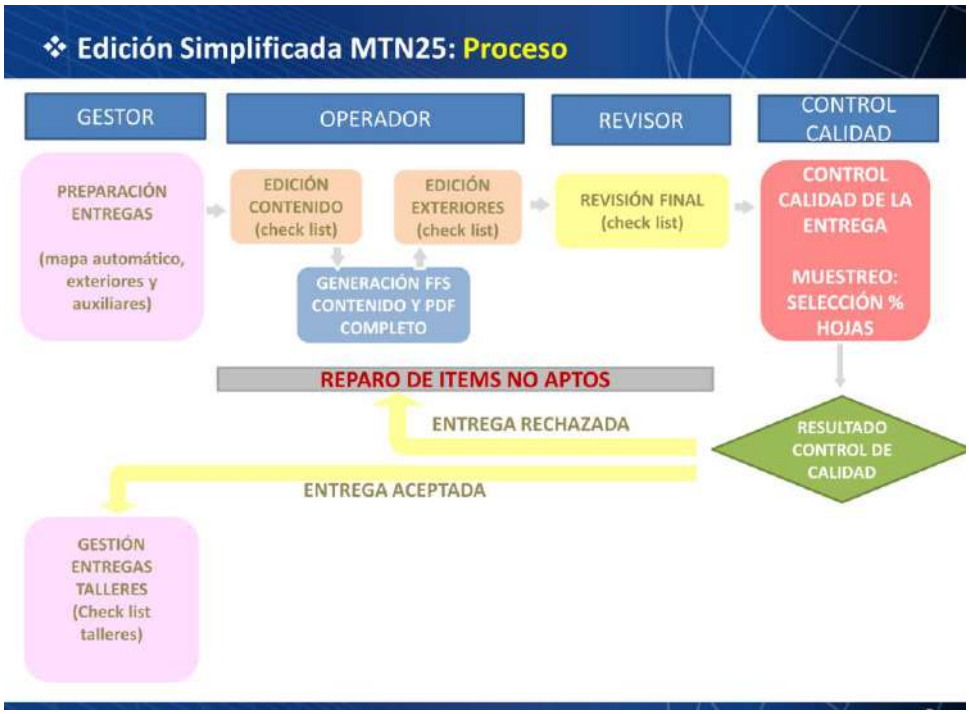


Ilustración 5 - 91: Flujo del proceso de Edición Simplificada

Registro

El sistema de registro mediante el que va a ser posible llevar a cabo el seguimiento y el trabajo que se está realizando, es por un lado el check-list, que se va rellenado durante el proceso y al finalizar el trabajo se guarda con el pdf final, contemplando todas las acciones realizadas en los ítems y el resultado del control de calidad. Por otro lado, se dispone del **sistema Redmine**, sistema de gestión de proyectos online, en el que se registran los diferentes tipos de tareas que se llevan a cabo durante el proceso (ejecución, validación, corrección y soporte), así como la documentación, noticias, foros e informes Gannt y de tiempos generados durante la ejecución.

5.3.3. Diferencias con el proceso tradicional

Una de las principales diferencias con el proceso tradicional de generación de hojas del MTN25 es que únicamente se efectúan trabajos de edición, no de actualización masiva, aunque como se ha mencionado anteriormente, en determinadas ocasiones sí es necesario puntualmente realizar este tipo de tareas.

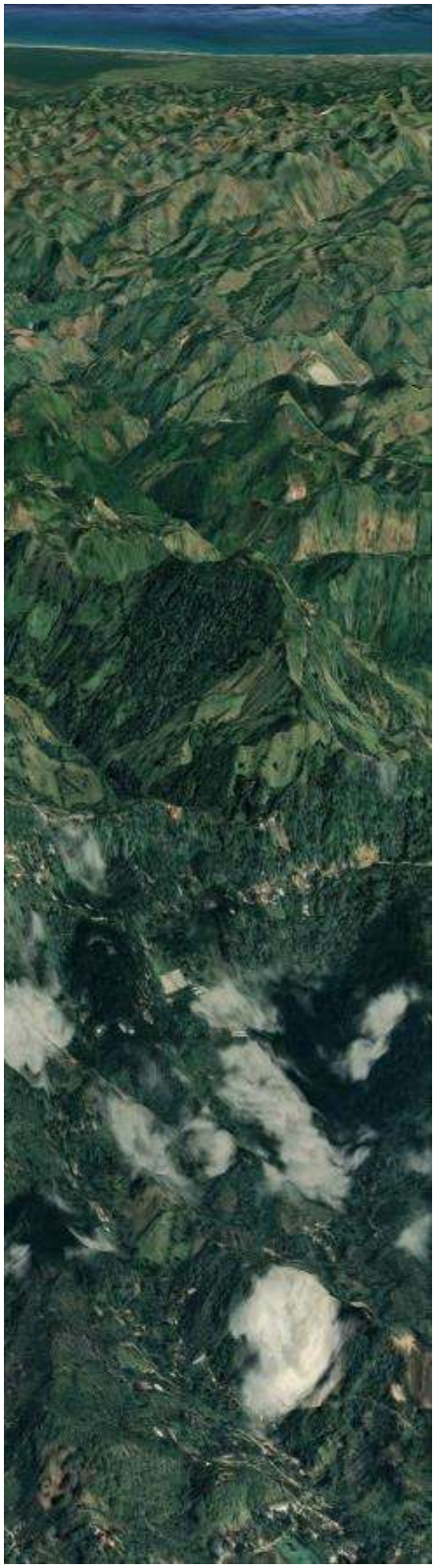
Entre otras **ventajas** respecto a la forma tradicional de hacer cartografía, se pueden destacar las siguientes:

- Menor carga de edición. Se parte de un producto en el que se ha llevado a cabo ya una edición y generalización automática. Los criterios de edición únicamente atienden a cuestiones de legibilidad y a los procesos de edición automáticos que no nos dan una solución deseable, y no a las especificaciones del MTN25 tradicional.
- Tiempo mínimo entre la actualización del dato y publicación en formato de cartografía impresa.
- Menor uso de recursos.

- Costes de producción menores.
- Mayor objetividad.
- Se dispone de un control de tiempos de proceso y resultados (registro de calidad).
- Posibilita la implantación de la Mejora Continua (indicadores de calidad)

Como principales **inconvenientes** de este proyecto se puede señalar:

- Una menor calidad de edición. Con la edición realizada de forma tradicional se conseguían unos resultados estéticos difíciles de igualar con la generación automática.
- Mayor riesgo de no detectar la información que contenga errores. Al revisarse sólo los conflictos que la herramienta automática no resuelve de forma satisfactoria y los que se refieren a cuestiones de legibilidad, el resto de la información que tenga errores pasará desapercibida.



CAPÍTULO 6: PRODUCCIÓN DEL MAPA II. VISUALIZACIÓN FINAL DEL PRODUCTO

Técnicas de simbolización, sombreado y rasterizado

Descripción breve

Contiene la descripción de los procesos que se aplican a los datos cartográficos generados (capítulo 5) hasta su publicación final en distintos formatos. Estos procesos, orientados a la visualización final del producto, contemplan: la simbolización, la incorporación del sombreado y los exteriores y el rasterizado del mapa.

Índice de contenidos

6.1. Introducción 1

6.1.1. Teoría del color 1

La Luz 1

Síntesis aditiva de colores-luz 2

El Pigmento 2

Síntesis sustractiva de colores-pigmento 3

Atributos del color 4

Tono 4

Saturación 5

Luminosidad 5

Modelos de color 5

El modelo RGB 6

El modelo CMYK 6

Uso de colores planos en la impresión offset 7

6.1.2. Simbolización cartográfica 8

Tipos de símbolos cartográficos según su geometría 8

Características visuales de los símbolos 9

Tipos de símbolos puntuales y su clasificación 11

Clasificación mediante textos. El rótulo como signo convencional 12

a) Cuantitativa: 13

b) Cualitativa: 13

c) De orden: 13

d) Sobre disciplinas diversas: 13

e) De localización: 13

f) Técnica: 13

6.1.3. Herramientas habituales aplicables en esta fase 18

Adobe Illustrator/Avenza Mapublisher 18

Adobe Illustrator 18

Avenza Mapublisher 19

Adobe Photoshop / Geografic Imager / Global Mapper 20

Adobe Photoshop 20

Geographic Imager 20
Global Mapper 20
FME (Feature Manipulation Engine) 21
6.2. Modelo Digital del Terreno (MDT) y Sombreado 22
6.2.1. Modelo Digital de Elevaciones (MDE) 22
Curvas de nivel 22
Red Irregular de triángulos (TIN) 23
Malla regular (grid) 24
6.2.2. El sombreado 26
Obtención del sombreado a partir del Modelo Digital del Terreno 27
Software utilizado 27
Proceso de obtención del sombreado 28
Realización de un mosaico de provincias en Global Mapper 35
Planteamiento 35
Ejecución 35
Exportación del sombreado por hojas 37
Exportación por hojas a escala 1:25,000 37
Exportación por hojas a escala 1:50,000 43
6.3. Procesos de simbolización e integración de sombreado 44
6.3.1. Proceso manual de simbolización para publicación impresa 44
Utilización básica de Adobe Illustrator 44
Creación y manipulación de objetos 44
Producción y dibujo de objetos 44
Manipulación y transformación de objetos 45
Alineación y distribución de objetos 46
Creación de objetos complejos (operaciones Boolenas) 46
Uso de Cuadrículas y Guías 46
Manejo del Color 47
Edición de Líneas 48
Edición de Textos 49
Estilos Gráficos y de Carácter 49
Símbolos 50
Pinceles 50
Capas 51
Uso de Plantillas 51
Utilización básica de Mapublisher junto con Adobe Illustrator 52

Importación de datos hacia la plantilla 53
Sistema de Coordenadas 53
Map Views 54
Map Attributes 56
Map Themes 57
Descripción de las fases del proceso 60
PASO 1. Montaje de la Hoja 60
PASO 2. Incorporación de interiores 68
PASO 3. Comprobación de encaje entre interiores y exteriores 73
PASO 4. Incorporación del sombreado 73
PASO 5. Guardado del documento 75
6.3.2. Proceso automático de simbolización y rasterización 77
Software de producción. Consideraciones previas 77
PostGis: almacenamiento de los datos cartográficos 77
Ventana de datos de un dgn en FME 78
Illustrator y Mapublisher, ejemplo de entidad 79
Interacción entre de FME y la plantilla de Mapublisher (Maptheme) 81
Proceso de simbolización y automatización (parte vectorial) 85
Especialidad elementos de comunicaciones (Carreteras, pistas, caminos y vías verdes) 90
Manipulación de elementos para el case de hojas 92
Proceso de simbolización y automatización (parte raster) 98
Lectura de archivo geoPDF 98
Georreferenciación del ráster embebido 98
Procedimiento de sombreado 100
Proceso integrado: geoPDF más ráster 104
6.4. Publicación de las diferentes salidas 108
6.4.1. Publicación impresa 108
6.4.2. Publicación archivos vectoriales a descarga: geoPDF 108
6.4.3. Publicación archivos raster para servicios y descarga: COG, ECW 111
6.5. Evolución de las salidas cartográfica para visores 112
6.5.1. Ejemplo de cartografía específica para visores: el mapa de Alta Resolución 113
Definición y características 113
Diseño de Implementación 114

Índice de ilustraciones

Ilustración 6 - 1: Espectro Electromagnético - CC BY-SA 3.0: Horst Frank, Jailbird..... 1
Ilustración 6 - 2: Síntesis aditiva de colores-luz 2
Ilustración 6 - 3: Reflexión en un objeto opaco y Transmisión en un objeto transparente..... 3
Ilustración 6 - 4: Síntesis sustractiva de colores-pigmento..... 4
Ilustración 6 - 5: Círculo cromático de tonos 4
Ilustración 6 - 6: Gradación de saturación en orden descendente 5
Ilustración 6 - 7: Gradación de luminosidad en orden descendente 5
Ilustración 6 - 8: Dispositivos que usan el modelo CMYK 6
Ilustración 6 - 9: Dispositivos que usan el modelo CMYK 6
Ilustración 6 - 10: Guía Pantone..... 7
Ilustración 6 - 11: Guía Pantone..... 7
Ilustración 6 - 12: Tipos de datos y su distribución en el diseño de símbolos cartográficos 9
Ilustración 6 - 13: Características visuales de los símbolos..... 10
Ilustración 6 - 14: Ejemplo de la variedad de rótulos del Mapa Topográfico Nacional 14
Ilustración 6 - 15: Rótulos completos sobre fondo uniforme 15
Ilustración 6 - 16: Efecto barrera en rotulación 15
Ilustración 6 - 17: Orientación de los rótulos..... 15
Ilustración 6 - 18: Rótulos con halo..... 16
Ilustración 6 - 19: Sentidos preferentes en la rotulación..... 16
Ilustración 6 - 20: Curvatura de los rótulos..... 16
Ilustración 6 - 21: Rótulos con letras separadas 17
Ilustración 6 - 22: Ubicaciones preferentes del rótulo..... 17
Ilustración 6 - 23: Evitar confusión por alineación de rótulo y símbolo 17
Ilustración 6 - 24: Evitar intersección de rótulo y líneas de cuadrícula 18
Ilustración 6 - 25: Evitar alineación de rótulos..... 18
Ilustración 6 - 26: Interface de Adobe Illustrator..... 19
Ilustración 6 - 27: Interface de Photoshop + Geographic Imager 20
Ilustración 6 - 28: Interface de Global Mapper 21
Ilustración 6 - 29: Interface de FME 21
Ilustración 6 - 30: Modelo del terreno con curvas de nivel 22
Ilustración 6 - 31: Perspectiva 3D de las curvas de nivel 22
Ilustración 6 - 32: Perspectiva 3D de un modelo TIN..... 23
Ilustración 6 - 33: Perspectiva 3D más cercana de un modelo TIN..... 23
Ilustración 6 - 34: Aspecto numérico de una malla regular 24
Ilustración 6 - 35: Aspecto gráfico de una malla regular interpretada por un SIG, con origen en esquina suroeste 25
Ilustración 6 - 36: Planta embalse y curvas de nivel 25
Ilustración 6 - 37: Perspectiva 3D del embalse 26
Ilustración 6 - 38: Vista parcial de la hoja 552-2 sin sombreado 26
Ilustración 6 - 39: Vista parcial de la hoja 552-2 con sombreado 27
Ilustración 6 - 40 : Vista parcial del sombreado de la hoja 552-2 27
Ilustración 6 - 41 Geotiff "normal" del MDT de Bizkaia 31

Ilustración 6 - 42 Geotiff "cenital" del MDT de Bizkaia 31
Ilustración 6 - 43 Conjunto de archivos que han intervenido en la obtención del sombreado.. 34
Ilustración 6 - 44: Archivos de sombreado resultantes por NUMERO de hoja de MTN25 42
Ilustración 6 - 45: Tabla hoja_c en PostGIS 77
Ilustración 6 - 46: Atributos implicados 77
Ilustración 6 - 47: Tabla 08p de textos 77
Ilustración 6 - 48: Atributos implicados 78
Ilustración 6 - 49: Datos de un archivo dgn en la ventana Inspector de FME..... 78
Ilustración 6 - 50: Configuración de la Mesa de Trabajo..... 80
Ilustración 6 - 51: Entidad de texto con sus atributos en Mapublisher 80
Ilustración 6 - 52: Regla para simbolización 80
Ilustración 6 - 53: Estilo gráfico correspondiente a la simbolización del Fondo de Autonómica de 1ª 81
Ilustración 6 - 54: Tabla de atributos y entidades. 81
Ilustración 6 - 55: Atributos implicados 81
Ilustración 6 - 56: Cruce de entidades geográficas con la tabla en Excel..... 82
Ilustración 6 - 57: Tabla de textos negros 83
Ilustración 6 - 58: Atributos implicados. 83
Ilustración 6 - 59: Regla para simbolización 84
Ilustración 6 - 60: Estilo gráfico correspondiente a la simbolización de sitios..... 84
Ilustración 6 - 61: Cruce de entidades de texto con la tabla de atributos de texto..... 84
Ilustración 6 - 62: Selección de textos..... 85
Ilustración 6 - 63: Parámetros de conexión a la Base de datos PostGIS 86
Ilustración 6 - 64: Parámetros del "reader" para leer los marcos de hoja. 86
Ilustración 6 - 65: Cuadro de diálogo para la elección del número de hoja..... 86
Ilustración 6 - 66: Parámetros del "reader" de tablas de entidades de PostGIS 87
Ilustración 6 - 67: Ejemplo de inicio de ejecución del proceso de simbolización 87
Ilustración 6 - 68: Marco con uno de los lados ampliados..... 88
Ilustración 6 - 69: Ejemplo de mapa simbolizado con zona ampliada 88
Ilustración 6 - 70: Proceso de generación de la cuadrícula de la proyección UTM 88
Ilustración 6 - 71: Parámetros para el cruce de la tabla hojac con la tabla correspondiente 89
Ilustración 6 - 72: Entidad de PostGIS correspondiente a la tabla 05p..... 89
Ilustración 6 - 73: Tabla de excel que contiene los campos nncccppe y la descripción necesaria para la simbolización..... 89
Ilustración 6 - 74: Parámetros para el cruce de tablas de PostGIS y Excel 89
Ilustración 6 - 75: Eje, caja y fondo correspondiente a Autovías y su composición 90
Ilustración 6 - 76: Flujo para generar entidades duplicadas y triplicadas..... 90
Ilustración 6 - 77: Parámetros Transformador AttributeManager 91
Ilustración 6 - 78: Parámetros Transformador StringConcatenator 91
Ilustración 6 - 79: Estilo de estanque, piscina. 92
Ilustración 6 - 80:Trazo y relleno para el estilo de estanque, piscina. 92
Ilustración 6 - 81: Maptheme de entidades lineales..... 92
Ilustración 6 - 82: Vista general del proceso de extracción de entidades que tocan el marco. . 93
Ilustración 6 - 83: Filtro de entidades que tienen trazo y relleno..... 93
Ilustración 6 - 84: Transformador que obliga a las entidades superficiales a ser líneas. 94
Ilustración 6 - 85: Transformador que "rompe" una línea cerrada en varios segmentos. 94
Ilustración 6 - 86: Parámetros buffer de los segmentos del marco. 95
Ilustración 6 - 87: Parámetros para hacer el filtro espacial 95

Ilustración 6 - 88: Writer de FME 96
Ilustración 6 - 89: Configuración general del writer..... 96
Ilustración 6 - 90: Ajustes del writer de Adobe Illustrator - Avenza MAPublisher..... 97
Ilustración 6 - 91: Lector de geoPDF 98
Ilustración 6 - 92: Parámetros del lector de geoPDF..... 98
Ilustración 6 - 93: Cálculo del centro del marco que nos servirá para georreferenciar el ráster 99
Ilustración 6 - 94: Escalado y desplazamiento del ráster embebido a su verdadera posición ... 99
Ilustración 6 - 95: Relación de la dimensión de la mesa de trabajo en dimensiones papel y los lados de la mesa de trabajo en dimensiones terreno, es decir, a escala 1:25000..... 100
Ilustración 6 - 96: Parámetro de Sistema de Coordenadas a través del valor del atributo **epsg.** 100
Ilustración 6 - 97: Transformador para ordenar las bandas y obtener el mapa en color verdadero (RGB). 100
Ilustración 6 - 98: RasterPaletteResolver 101
Ilustración 6 - 99: Parámetros que "obligan" a una cierta interpretación de las bandas del ráster. 101
Ilustración 6 - 100: Transformador para añadir una banda con su valor para todos los píxeles. 101
Ilustración 6 - 101: Transformador RasterExpressionEvaluator. 102
Ilustración 6 - 102: Configuración del transformador RasterExpressionEvaluator..... 103
Ilustración 6 - 103: Cambio de las componentes de color del fondo de las ramblas. 103
Ilustración 6 - 104: Transformador WorkspaceRunner..... 104
Ilustración 6 - 105: Parámetros del WorkspaceRunner para la simbolización y generación de archivos geoPDF. 104
Ilustración 6 - 106: Tabla correspondiente a las hojas que van a entrar en un determinado lote. 105
Ilustración 6 - 107: Transformador Decelarator. 105
Ilustración 6 - 108: Parámetros del transformador Decelerator. 106
Ilustración 6 - 109: "Esperador" con un sorter..... 106
Ilustración 6 - 110: Parámetros del transformador "Sorter". 106
Ilustración 6 - 111:WorkspaceRunner que llama al workbench de rasterizado. 107
Ilustración 6 - 112: Parámetros del WorkspaceRunner de generación del rasterizado. 107
Ilustración 6 - 113: Proceso completo de automatización para la producción de geoPDF y su rasterizado..... 107
Ilustración 6 - 114: Ventana de importación del archivo ffs 109
Ilustración 6 - 115: Exportación a geoPDF 109
Ilustración 6 - 116: Cuadro de diálogo que permite quitar el archivo AI (responsable del ráster) del archivo geoPDF..... 109
Ilustración 6 - 117: Ajuste para la automatización de cara a FME 110
Ilustración 6 - 118: Introducción del archivo .pes de ajustes..... 110
Ilustración 6 - 119: Parámetros para configurar si se quiere o no incluir el archivo ai dentro del geopdf 111
Ilustración 6 - 120: Analogías mapa de Alta Resolución con MTN25 automático 113
Ilustración 6 - 121: Diferencias con MTN25 automático..... 114
Ilustración 6 - 122: Flujo del Mapa de Alta Resolución a partir del de Mapa Automático a escala 1:25,000..... 114

6.1. Introducción

En este capítulo se describen los procesos y las técnicas que se aplican a los datos cartográficos generados en el proceso de producción descrito hasta ahora (resultados del capítulo 5), que están orientados a optimizar la calidad visual de los mismos en el producto final, al objeto de que el mapa comunique al usuario el contenido del territorio de una manera eficiente: simbolización, sombreado, rasterizado y generación del producto final tras incorporar los exteriores del mapa.

Dada la importancia que la componente "visualización" tienen en esta segunda fase de la producción cartográfica, el contenido del capítulo comienza con la explicación de nociones básicas y fundamentales sobre la teoría del color, para adentrarse después en las técnicas de simbolización cartográfica y de realce de la orografía y finaliza con la explicación de las distintas tipologías de salidas que se pueden generar.

6.1.1. Teoría del color

La reproducción de un original a color nos obliga a disponer de ciertos conocimientos y especialización en distintas técnicas que nos ayuden a que la consecución del trabajo sea lo más fiel posible al fin deseado, es por eso que se haga una breve introducción a la Teoría del Color y la Reproducción Gráfica.

La Luz

La luz es una de las formas de energía que puede ser captada por el ojo humano permitiendo ver los objetos que la emiten directamente como podría ser una bombilla, o bien indirectamente al reflejarla desde un objeto que la emita o al atravesar un objeto transparente.

La luz viaja a una velocidad de 300,000 km/segundo. Se podría considerar que la luz es como un rayo que se mueve en línea recta, pero puede pasar a través de un objeto (transparente) o ser parada por él (opaco). También existe el caso de objetos que dejan pasar sólo una parte de la luz (translúcidos).

La luz visible está constituida por radiaciones electromagnéticas de propagación rectilínea, que ocupa un intervalo reducido en el espectro electromagnético conocido. Dentro de éste se encuentran las ondas de radio y televisión, los rayos X, la radiación calorífica, radiación infrarroja, radar, microondas, etc., y, aunque todas ellas tienen una constitución ondulatoria como la de la luz, su longitud de onda es distinta y por ello el ojo humano no puede captarlas y en consecuencia son invisibles. El **espectro electromagnético visible por el ojo humano** queda comprendido entre **los 400 y los 750 nanómetros de longitud de onda**, siendo esta la distancia entre la cresta de una onda y la cresta de la onda sucesiva.

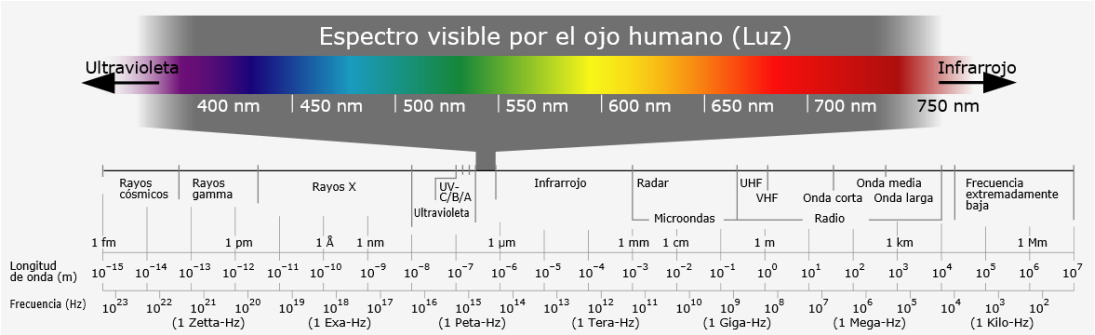


Ilustración 6 - 1: Espectro Electromagnético - CC BY-SA 3.0: Horst Frank, Jailbird

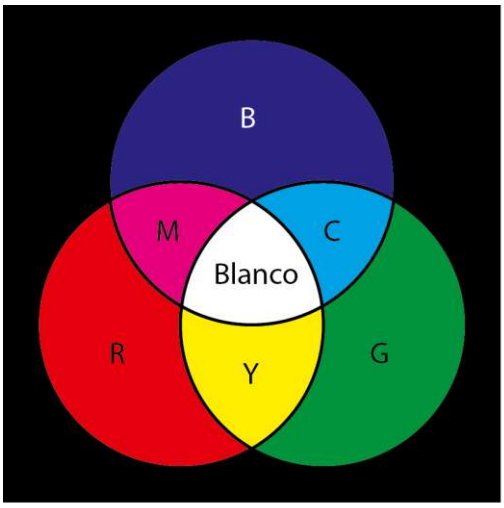
Si hiciéramos atravesar una luz blanca por un prisma de cristal, ésta se dispersaría en las distintas longitudes de onda haciendo visibles los colores que la componen dando lugar al espectro visible de la luz.

En este espectro se puede comprobar que las principales luces de color son la azul-violeta: aproximadamente entre 400-500 nanómetros; la verde: aproximadamente 500-600 nanómetros y la roja con aproximadamente 600-750 nanómetros. Estas luces de color se denominan **primarias.**

De igual manera, se da el fenómeno de que cuando todas las longitudes de onda del espectro visible estimulan simultáneamente y en partes casi iguales (o sea con la misma intensidad) el ojo humano, se percibe la luz blanca.

Síntesis aditiva de colores-luz

Tal cual hemos visto anteriormente, si proyectáramos simultáneamente las tres luces primarias una sobre otra tal cual se muestra en la siguiente figura, se obtendría la luz blanca. En las zonas donde se suman dos colores-luz primarios aparecen los colores-luz secundarios que son una aproximación hacia la luz blanca y que se denominan también "complementarios".



Colores-luz primarios

R= Red = Rojo
G= Green = Verde
B= Blue = Azul-violeta

Colores-luz secundarios

C= Cyan = Cyan
M= Magenta = Magenta
Y= Yellow = Amarillo

Ilustración 6 - 2: Síntesis aditiva de colores-luz

Tal cual se ha comentado anteriormente la luz blanca podría reconstruirse además de superponiendo los tres colores luz primarios, superponiendo una luz coloreada primaria (por ejemplo, la azul-violeta) sobre una secundaria (la amarilla por ejemplo, que también es la suma de las otras luces coloreadas primarias verde y roja).

En la siguiente tabla se reproduce un esquema de luces coloreadas complementarias:

Luz coloreada primaria	**Azul**	**Verde**	**Roja**
Luz complementaria correspondiente	**Amarilla**	**Magenta**	**Cyan**

El Pigmento

Los objetos que nos rodean aparecen coloreados a nuestra vista porque están constituidos o recubiertos por elementos y sustancias químicas denominados pigmentos, que tienen la

propiedad de absorber todas o algunas de las longitudes de onda de la luz y reflejan las restantes, que son las que percibe el ojo humano.

Por ejemplo, un objeto nos parece verde porque la luz blanca (azul + verde + roja) ha sido parcialmente absorbida (azul + roja) y parcialmente reflejada (verde) por el pigmento que constituye dicho objeto.

El mismo fenómeno se produce si el pigmento está incorporado a un cuerpo transparente, por ejemplo, una placa de cristal.

Dicho cristal nos parecería verde porque la luz blanca (azul + verde + roja) es parcialmente absorbida (azul + roja) y parcialmente transmitida (verde) por el pigmento incorporado al cristal.

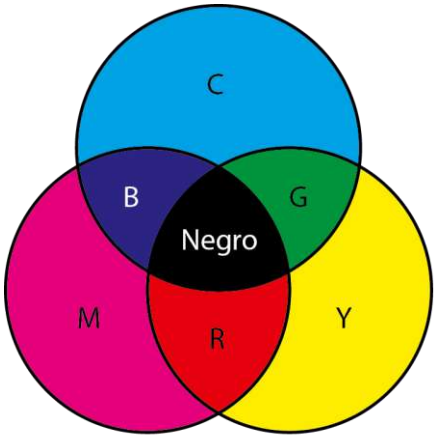
Ilustración 6 - 3: Reflexión en un objeto opaco y Transmisión en un objeto transparente

Por consiguiente, básicamente son dos tipos de objetos los que vemos coloreados a causa del pigmento que los constituye o recubre: objetos opacos (a los que pertenecen los papeles junto con las tintas de imprimir) y objetos transparentes (entre los cuales podríamos incluir los filtros coloreados de fotografía).

Los primeros reflejan y los segundos transmiten sólo las longitudes de onda que el ojo humano ve en ese momento gracias a los pigmentos, absorbiendo el resto. Este fenómeno se conoce como poder selectivo de los pigmentos.

Síntesis sustractiva de colores-pigmento

Cuando hablamos de pigmentos, el comportamiento es el contrario que con la luz: la suma de colores tiende hacia el negro. La suma de dos colores-pigmento primarios da lugar a uno secundario.



Colores-pigmento primarios

C= Cyan = Cyan
M= Magenta = Magenta
Y= Yellow = Amarillo

Colores-pigmento secundarios

R= Red = Rojo
G= Green = Verde
B= Blue = Azul-violeta

Ilustración 6 - 4: Síntesis sustractiva de colores-pigmento

El color-pigmento negro se podría reconstruir a través de la suma de los tres pigmentos primarios o bien de un color-pigmento primario más su complementario.

En la siguiente tabla se reproduce un esquema de colores-pigmento complementarios:

Color-pigmento primario	**Amarillo**	**Magenta**	**Cyan**
Color-pigmento complementario	**Azul**	**Verde**	**Rojo**

Atributos del color

Los atributos son los que hacen único a cada color. Básicamente son tres: tono, saturación y luminosidad.

Tono

Aunque lo hagamos de una manera coloquial, cuando nos referimos a un color realmente estamos pensando en el tono. También es conocido como matiz, tinte, croma o Hue (en inglés). El tono básicamente es la longitud de onda predominante del color que vemos y por la cual designamos los colores: rojo, violeta, azul, etc. Un buen referente visual para conocer lo que son los tonos sería el arcoíris.

Ilustración 6 - 5: Círculo cromático de tonos

Saturación

La saturación nos indica la intensidad cromática o pureza del color y queda determinada por la cantidad de luz que posee un color. Sus valores van desde su máximo, el color puro, hasta su mínimo, que sería un tono de gris. De esta forma, cuanto más gris tenga el color éste será más pálido y con menor saturación. Cuando decimos que un color es muy vivo o intenso esto significa que está muy saturado.

Ilustración 6 - 6: Gradación de saturación en orden descendente

Luminosidad

La luminosidad es la cantidad de luz emitida o reflejada por un objeto. En un color sería su oscuridad o claridad. Cuanta más luminosidad tenga un color más luz será reflejada y más se acercará al blanco. Y al contrario, cuanta menos luminosidad tenga se acercará al negro.

Ilustración 6 - 7: Gradación de luminosidad en orden descendente

Muy habitualmente hablamos de un color azul claro cuando éste tiene un matiz cercano al blanco; sin embargo, hablamos de un color azul oscuro cuando el azul se acerca al negro.

Modelos de color

Los modelos de color nos ayudan a clasificar y definir los colores en función de una serie de parámetros matemáticos. El uso de unos u otros modelos de color vendrá definido por el propósito de cada uno.

Podemos definir dos grandes modelos de color de uso común en el ámbito del diseño gráfico y la impresión de artes gráficas.

Estos son el **Modelo RGB** y el **Modelo CMYK**.

El modelo RGB

Como ya conocemos, el modelo RGB es un modelo de color aditivo formado por los tres colores primarios: Rojo, Verde y Azul. Con la suma de estos colores podemos recrear el espectro completo. El modelo de color RGB se utiliza generalmente en dispositivos electrónicos como pueden ser las pantallas de los teléfonos móviles o monitores de los ordenadores.

Es el utilizado en los archivos que se publican en la web (.png, .jpg, .gif, etc.) así como los archivos ráster de cartografía (.ecw, .cog, .geotiff, etc.)

El modelo CMYK

El modelo CMYK es un modelo de color basado en la síntesis sustractiva y está formado por los tres colores primarios: Cyan, Magenta y Amarillo, a los cuales se añade un cuarto color Negro.

Aunque en teoría la suma de los tres colores-pigmento primarios (Cyan, Magenta y Amarillo) daría el color negro, en la práctica el color así obtenido no es lo bastante intenso, dando como resultado un color pardo oscuro. Por dicho motivo, a la hora de imprimir se agrega pigmento negro que se encarga de potenciar las zonas más oscuras de la imagen (sombras).

Con la suma de estos cuatro colores podemos representar sólo una parte del espectro.

Ilustración 6 - 8: Dispositivos que usan el modelo CMYK

El modelo de color CMYK se utiliza generalmente en imprenta (bien sea digital u offset).

De ahí, por ejemplo, que en las impresoras de uso común de chorro de tinta o impresora digital tengamos cuatro cabezales (o cuatro cartuchos) correspondientes a cada uno de los colores primarios (CMY) más el negro (Key).

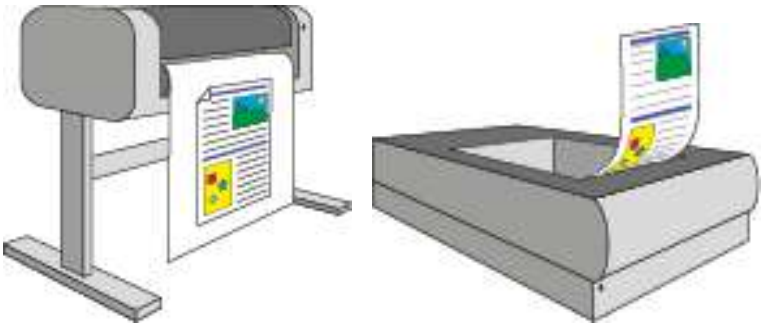


Ilustración 6 - 9: Dispositivos que usan el modelo CMYK

En imprenta tradicional también es el modelo de color estándar, por lo tanto, de ahí viene la separación de colores en cuatro planchas (C, M, Y, K)

Uso de colores planos en la impresión offset

Dada la complejidad que puede tener la reproducción de cartografía en imprenta tradicional (offset), existe un recurso que facilita la reproducción de determinados elementos. Este recurso es el uso de **tintas planas** (también conocidas como colores planos).

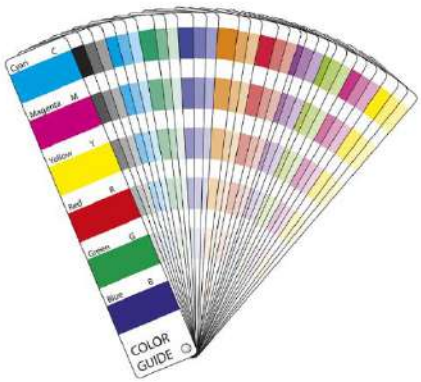


Ilustración 6 - 10: Guía Pantone

Una **tinta plana** es una **tinta** especial previamente mezclada que se utiliza en lugar, o además, de las **tintas** de cuatricromía y que requiere su propia plancha de impresión en una imprenta. Las **tintas planas** pueden reproducir con precisión los colores que están fuera de la gama de los colores de cuatricromía (por ejemplo: colores metálicos, colores con tonos pastel, fosforescentes, etc.)

El catálogo de tintas planas más común en artes gráficas es la Guía Pantone.

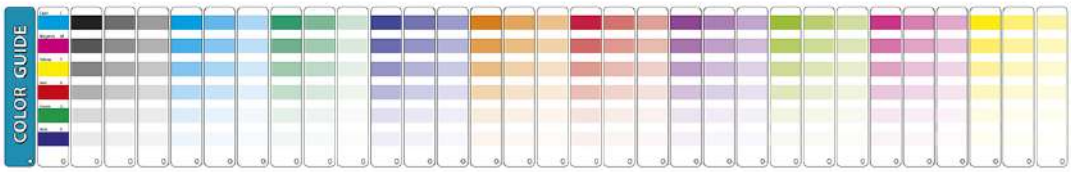


Ilustración 6 - 11: Guía Pantone

En cartografía, y más concretamente en algunas series del IGN de España, se utiliza como tinta plana (o quinto color) el siena de las curvas de nivel. De este modo, la separación de colores de una hoja de la escala 1:200,000 (provincial) sería: CMYK + Siena haciendo un total de cinco colores de imprenta. Esto facilita el ajuste del trabajo en las máquinas de impresión offset, sobre todo cuando los mapas son de tamaño muy grande.

Existen algunos dispositivos de impresión digital que tienen la posibilidad de trabajar con colores planos, pero por norma general no es así. De hecho, en los dispositivos de impresión digital, el problema de ajuste de colores se minimiza considerablemente, no exigiendo planchas adicionales para evitarlo.

Otra cosa sería el uso de colores con componentes metálicos (oro o plata), que sólo se podrían reproducir debidamente con “tintas especiales” si el dispositivo es capaz de gestionarlas.

6.1.2. Simbolización cartográfica

Los mapas se realizan teniendo en cuenta el objetivo fundamental de transmitir información geográfica, y los procesos de recogida de datos, elección de la escala y proyección, simbolización y diseño gráfico, se enfocan en función de conseguir este fin, constituyendo todos ellos el diseño cartográfico del mapa.

Dentro del diseño cartográfico el diseño gráfico es una parte vital de la cartografía, debido a que una comunicación efectiva requiere que los diversos signos (líneas, símbolos, rótulos, colores, tonos, etc.) se hayan elegido con cuidado y presenten una armonía de conjunto.

La cartografía automatizada y asistida por ordenador ha tenido un gran impacto en la realización de mapas, pero con mayor fuerza en cuanto a exactitud, velocidad de ejecución, coste, agilización de tareas tediosas y flexibilidad. Sin embargo, diversos aspectos del diseño gráfico se han visto afectados, unas veces de modo positivo y otras en negativo. En el lado positivo, se ha conseguido una mayor flexibilidad de procedimientos y una mayor sencillez en la realización de cambios, lo que ha facilitado el desarrollo de prototipos y el ensayo de nuevas opciones de diseño. En el lado negativo, debido a la dificultad de realizar una gran organización gráfica, se ha tendido a simplificar la simbología, resultando algunos mapas bastante pobres, con un mínimo de información geográfica asociada.

Tipos de símbolos cartográficos según su geometría

Existen muchas formas distintas de simbolizar (es decir, codificar) los datos geográficos, conceptos y relaciones que aparecen en un mapa, pero con el fin de considerar como pueden utilizarse los signos para simbolizar la variedad de datos, es útil clasificarlos. Podemos reconocer cuatro clases de elementos gráficos: Puntos, líneas, zonas y rótulos.

- **Símbolos puntuales**: son signos individuales, utilizados para representar datos posicionales o de lugar no representables en su dimensión real a la escala del mapa; como por ejemplo ciudades, la altura de un lugar, edificaciones singulares, el centro de alguna distribución, o un volumen en un lugar, como la población de una ciudad. A pesar de que un signo puntual pueda abarcar una cierta cantidad del espacio cartográfico, siempre será un símbolo de punto por referirse a una localización concreta.
- **Símbolos lineales**: son signos lineales individuales utilizados para representar datos geográficos que tienen una realidad lineal (carreteras, ríos, ferrocarriles, etc.), o para representar datos puntuales que tienen una misma cualidad (las curvas de nivel son líneas utilizadas para representar puntos de igual elevación).
- **Símbolos zonales**: son signos superficiales que se extienden sobre una determinada superficie del mapa para indicar que tal región posee algún atributo común, como por ejemplo agua, jurisdicción administrativa, un mismo cultivo, o alguna característica que se pueda medir. Cuando se utiliza de este modo, un símbolo zonal se extiende de manera uniforme sobre la superficie que representa. Otro tipo de signo zonal está formado por símbolos no uniformes, como por ejemplo una variación tonal, para representar variaciones continuas de algún fenómeno de un lugar a otro. A este tipo de símbolo zonales pertenecen los sombreados del relieve, que dan una idea de las pendientes del terreno (no de su altura).

La siguiente Ilustración 6 - muestra algunos ejemplos de la gran variedad de símbolos puntuales, lineales y zonales utilizados para representar algunos tipos de datos nominales, ordinales y de valores de intervalo.

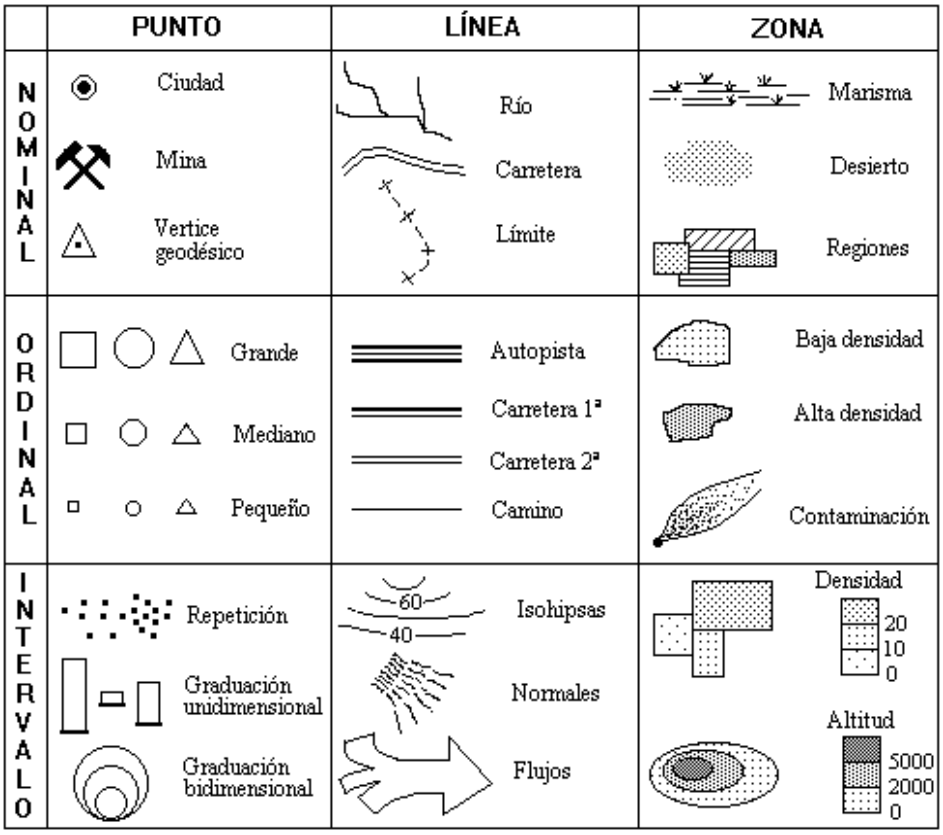


Ilustración 6 - 12: Tipos de datos y su distribución en el diseño de símbolos cartográficos

- **Los rótulos como símbolos**: los rótulos que aparecen en un mapa además de cumplir una labor informativa, al designar mediante el lenguaje escrito a los elementos del mapa, tienen también un valor muy importante como símbolo, ya que mediante la distinta tipografía pueden dar información cualitativa y cuantitativa de los elementos a los que nombran de manera verbal. Por ejemplo, mediante el distinto tamaño de los rótulos se puede indicar el número de habitantes de los núcleos de población, de la inclinación de las letras de determinados rótulos se puede deducir si se refieren a elementos naturales o artificiales del mapa, o con la forma de algunos rótulos se describe la conformación del elemento al que designan, como en el caso de los rótulos de los accidentes orográficos.

Características visuales de los símbolos

Las tres clases de símbolos, junto con los rótulos, componen los elementos gráficos básicos de la representación cartográfica; pero con el fin de representar los distintos datos de un modo significativo, se debe variar el aspecto de los símbolos mediante la utilización de lo que se podría denominar elementos visuales básicos, que son:

- **El color:** el color es una percepción visual tan importante como compleja, por eso le hemos dedicado un apartado específico en el capítulo anterior descubriendo todas sus particularidades aplicables al mundo de la cartografía.

- **El brillo**: como cualidad gráfica se refiere a la oscuridad o claridad de un signo, ya sea este negro o de cualquier otro color.
- **El tamaño**: los signos varían de tamaño (diámetro, área, anchura, altura, etc.) Normalmente, el tamaño de los símbolos va asociado a su importancia.
- **La forma:** la forma es la característica gráfica proporcionada por el aspecto distintivo de una figura regular, como por ejemplo un círculo o un triángulo; el perfil de un área irregular, como por ejemplo un estado o una isla; o el contorno de una característica lineal, como un río o una zona costera.
- **El espaciado:** cuando se diseña un signo mediante una determinada disposición de marcas, como por ejemplo una serie de puntos o líneas, su espaciado puede ser variado. A veces, variando el espaciado se pretende dar una idea cuantitativa o cualitativa del elemento representado.
- **La orientación**: la orientación se refiere a las disposiciones direccionales de una señal individual, o de las líneas paralelas encuadradas dentro de algún elemento o referidas a los límites del mapa.
- **La ubicación**: la ubicación sobre el mapa se aplica en general únicamente a aquellos componentes que pueden desplazarse, como por ejemplo los rótulos. Las ubicaciones de la mayoría de los símbolos vienen obligadas por su situación en el terreno y son inamovibles.

Los elementos visuales se pueden combinar de infinidad de maneras para una correcta comunicación de los contenidos del mapa. En la siguiente figura se han agrupado según se aplican a los elementos gráficos puntuales, lineales y zonales.

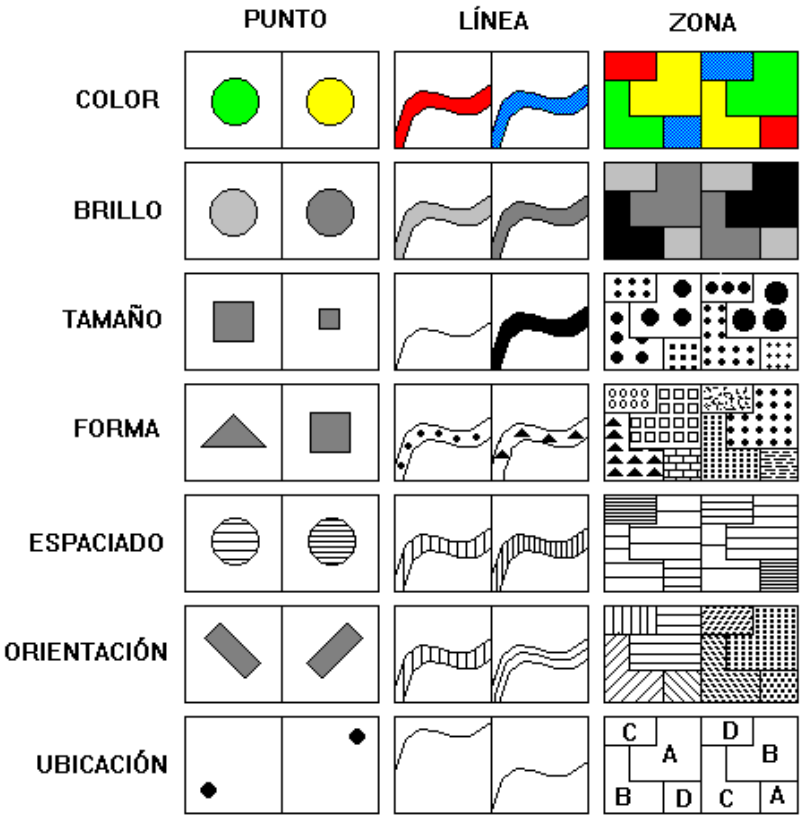


Ilustración 6 - 13: Características visuales de los símbolos

Tipos de símbolos puntuales y su clasificación

Los símbolos puntuales intentan referenciar fenómenos cuyas dimensiones no son representables a la escala del mapa. Dentro de toda la simbología de un mapa el diseño de los símbolos puntuales constituye el mayor reto para el cartógrafo. No hay que olvidar que los mapas son un instrumento de comunicación visual repletos de símbolos que pueden ser observados por multitud de usuarios con diferentes procedencias, culturas, creencias, nivel de conocimientos, etc. lo que dificulta el proceso de asociación mental de una idea a un símbolo.

En general el grado de simbolización puntual de un mapa depende de dos factores:

- La **escala del mapa**: en escalas grandes los símbolos puntuales afectarán a elementos que tengan una localización casi puntual sobre el terreno (pozo, vértice geodésico, entrada de una cueva, etc.), mientras que en escalas pequeñas podrán designar a elementos con una cierta entidad superficial (aeropuertos, hospitales, cascos de población, etc.). Vemos, pues, que la escala del mapa es determinante a la hora de decidir qué elementos se tendrán que representar mediante símbolos puntuales y cuáles no. Por regla general, cuanto más pequeña sea la escala de un mapa mayor carga simbólica puntual llevará, aunque la cantidad de símbolos puntuales depende también de la utilidad que se quiera dar al mapa.
- La **utilidad del mapa**: la cantidad de símbolos puntuales depende también del fin para el que se ha creado el mapa y a que usuarios va dirigido. En un mapa topográfico técnico el grado de simbolización será más bajo que en un mapa excursionista o turístico o temático.

El nivel de representación de un símbolo puede medirse según el grado de asociación instantánea que el usuario del mapa tenga al ver el símbolo en cuestión: si al observar un determinado símbolo el usuario reconoce inmediatamente el concepto correcto que el símbolo representa, estaremos hablando de un grado de asociación alto; por el contrario, si el usuario no es capaz de asociar símbolo y concepto, estaremos ante un grado de asociación inexistente.

La cartografía ha establecido una clasificación que divide en tres grandes grupos los símbolos puntuales respecto a su forma:

- **Símbolos pictóricos**: son los que más alto grado de asociación tienen, ya que la tendencia a la hora de diseñar el símbolo es la de realizar representaciones muy evocadoras del concepto original, incluso de tipo realista (en teoría no sería necesaria una leyenda para su comprensión). Este tipo de símbolos son recomendables para usuarios sin especial preparación cartográfica, aunque presentan inconvenientes técnicos, entre los cuales se encuentran una mayor complejidad de realización, un peor comportamiento en las reducciones de tamaño y un menor factor de precisión en la ubicación del fenómeno cartográfico.

Ilustración 6 - 3: Símbolos pictóricos de Estación espacial y Árbol singular.

- **Símbolos geométricos**: tienen un grado menor de asociación que los anteriores (normalmente necesitan de una leyenda para ser reconocidos) ya que reducen los rasgos significativos del fenómeno a representar a símbolos realizados mediante formas geométricas. Este tipo de símbolos son menos complejos, más flexibles en cuanto a procesos de reducción de tamaño y con mayor capacidad de precisión en la ubicación del elemento representado**.**

Ilustración 6 - 4: Símbolos geométricos de Cuevas (Habitada, Industrial y Natural)

- **Símbolos literales:** están formados por caracteres alfanuméricos, usualmente iniciales de palabras que hacen referencia directa al concepto que intentan representar. El problema más grave de estos símbolos es que el concepto no está representado mediante una abstracción simbólica, sino por un carácter que tiene todavía más restricciones de tipo cultural que los dos anteriores grupos ya que, evidentemente, la letra del símbolo hace referencia a una palabra en un idioma determinado que el usuario del mapa no tiene porqué conocer.

Ilustración 6 - 5: Símbolo literal de Helipuerto

En los mapas podemos encontrar símbolos que pertenezcan a una sola de las categorías que se acaban de ver o símbolos que son combinaciones de todas o algunas de ellas, por lo que el símbolo resultante disfrute y sufra de las ventajas y los inconvenientes que cada una de las categorías le aporte.

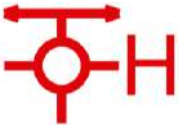


Ilustración 6 - 6: Símbolo pictórico y literal de Central eléctrica hidráulica

Clasificación mediante textos. El rótulo como signo convencional

La rotulación es uno de los componentes más significativos en la representación visual y en el diseño gráfico de un mapa.

El ordenador es de gran utilidad en la preparación de las distintas fuentes de texto que se han de utilizar y proporciona una agilidad insuperable en las tareas de colocación de rótulos, pero el arte de la ubicación de estos competerá al criterio de un buen operador con experiencia cartográfica.

La calidad gráfica de un mapa depende en gran medida de las características del tipo de letra, tamaño y colocación de los rótulos. Los mapas topográficos presentan normalmente una gran cantidad de nombres. El reconocimiento del fenómeno al que se aplica un nombre, el “tiempo de búsqueda” necesario para encontrar un nombre, y la comodidad con que dicho nombre pueda leerse, son de gran importancia en la funcionalidad del mapa.

La rotulación no debe contener información redundante, (es absurdo colocar al lado del símbolo de “ermita” el rótulo “ermita”), hay que procurar que la rotulación armonice con el conjunto, de tal modo que realce la comunicación sin atraer hacia sí misma una importancia excesiva.

Como todos los otros signos que aparecen sobre un mapa, el rótulo es un símbolo, pero es mucho más complicado en su funcionalidad que la mayoría de los otros símbolos. La rotulación, además de designar a los distintos elementos del mapa, nos aporta los siguientes tipos de información:

a) Cuantitativa:

- Tamaño de los rótulos (número de habitantes de las poblaciones, etc.).
- Caracteres numéricos (cotas, coordenadas, etc.).

b) Cualitativa:

- Tipo de la fuente de texto (comarcas, sitios, importancia política o administrativa de una población, etc.).
- Inclinación de las letras (elementos artificiales, naturales, etc.).
- Color (hidrografía, comunicaciones, altimetría, etc.).
- Orientación de los rótulos (sentido en que crece la pendiente del terreno, etc.).

c) De orden:

- Mayúsculas
- Minúsculas } (mayor o menor importancia de carreteras, ríos, etc.).
- Tamaño

d) Sobre disciplinas diversas:

- Geomorfológica (conformación del suelo: *Los Canchales, La Quebrada, El Torcal,* etc.).
- Agropecuaria (usos del suelo, naturaleza de la vegetación, cultivos desaparecidos, usos ganaderos: *La Vega, El Robledal, Los Linares, Las Majadas, Cañada Real*, etc.).
- Económica (explotación de recursos, industrias: *Cerros del hierro, Camino de los Lagares,* etc.).
- Lingüística (distintos idiomas o dialectos: *Playa - Platja - Praia - Hondartza, El Helechar - El Felechar - El Jelechar*, etc.).
- Cultural (costumbres, deportes: *Prado de la Romería, Plaza del Frontón,* etc.).
- Histórica (distintos pueblos que habitaron la zona, antiguas construcciones, sucesos acaecidos: *La Almunia, El Castillejo, Loma de la Reyerta,* etc.).

e) De localización:

- Puntual (pueblos, símbolos, etc.).
- Lineal (ríos, carreteras, etc.).
- Superficial (comarcas, sitios, etc.).

f) Técnica:

- Altitudes.
- Distancias.
- Coordenadas.

De un modo más sutil el rótulo colocado sobre un mapa puede utilizarse como indicación de escala. Mediante su contraste de tamaño respecto a otros factores tales como la anchura de las líneas y el tamaño de los símbolos podemos a menudo saber "con solo mirar" que la escala de un mapa es mayor o menor que la de otro mapa.

Debemos considerar el problema de elegir el tipo de fuente de texto a utilizar en un mapa del mismo modo que lo hacemos con respecto a otros elementos del diseño cartográfico, es decir, valorando adecuadamente los objetivos que debe cumplir el mapa eligiendo los tipos de letras en función de la armonía que debe regir el conjunto.

La frase "rotulación de un mapa" significa la preparación de todos los aspectos relacionados con esta fase de la cartografía, que incluye todo el material alfanumérico. Implica a todo el proceso, desde el comienzo hasta el final, y exige la cuidadosa consideración de muchos factores distintos. Cuanto más elaborado y complejo sea un mapa, más serán los elementos a tener en cuenta, pero en general existen por lo menos siete puntos principales a considerar. Las complejidades del mapa y sus propósitos añadirían subtítulos a la siguiente lista básica:

1- La familia de la fuente (Arial, Courier, etc.).

2- El tamaño de la fuente (mayúsculas, minúsculas).

3- El cuerpo de la fuente, es decir lo que mide su altura, en puntos o milímetros.

4- La inclinación de las letras de la fuente (vertical, inclinada)

5- El color de los rótulos.

6- El contraste entre rótulos y fondo.

7- La ubicación de los rótulos.

Ilustración 6 - 14: Ejemplo de la variedad de rótulos del Mapa Topográfico Nacional

De estos siete puntos es el último el que más problemas plantea. La disposición de los rótulos dentro de los mapas es de vital importancia y obedece a múltiples razones derivadas de su escala, y su propósito. Pero se puede decir que la cartografía admite una serie de reglas generales en cuanto a la ubicación de la rotulación de los mapas:

1. Los nombres han de situarse sobre tierra o totalmente sobre agua, con excepciones por problemas de espacio en mapas de pequeña escala.

Ilustración 6 - 15: Rótulos completos sobre fondo uniforme

2. En general, todos los detalles planimétricos que tienen "efecto barrera" (carreteras, ferrocarriles, ríos, etc.) no deberán interponerse entre el rótulo y el lugar que este designa.

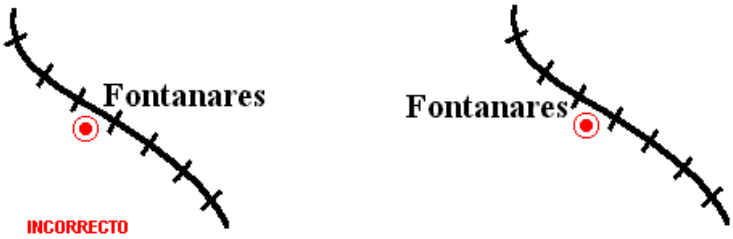


Ilustración 6 - 16: Efecto barrera en rotulación

3. La rotulación debe situarse, en general, de forma armonizada con la orientación de la estructura del mapa. En los mapas a gran escala esto implica el paralelismo con los bordes superior e inferior del marco; en los mapas de pequeña escala significa el paralelismo con los paralelos geográficos.

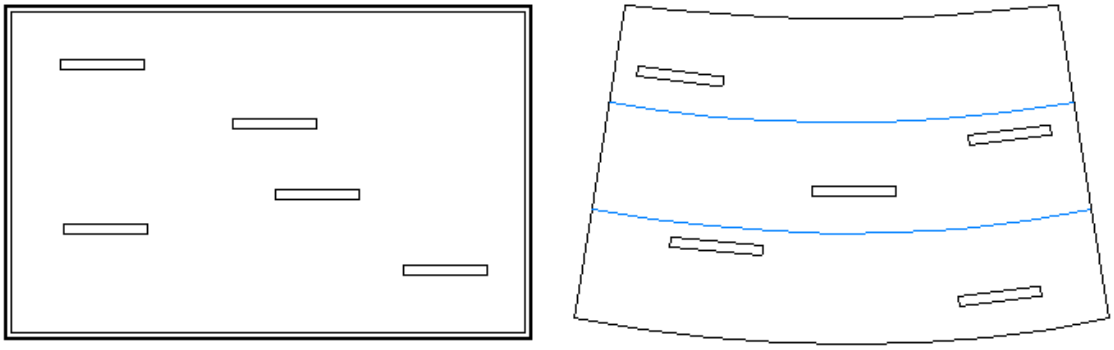


Ilustración 6 - 17: Orientación de los rótulos

4. Cuando los datos geográficos, como líneas y fondos, entren en conflicto con la rotulación, en el caso de que ambos sean del mismo color serán los datos, y no los rótulos, los que deban interrumpirse.

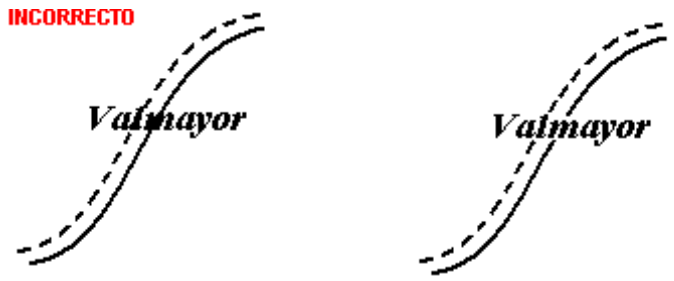


Ilustración 6 - 18: Rótulos con halo

5. Los rótulos no deben quedar nunca invertidos (como excepción, en algunos casos convendrá rotular invertidas las cotas de las curvas de nivel, para indicar el sentido de la pendiente ascendente).

6.

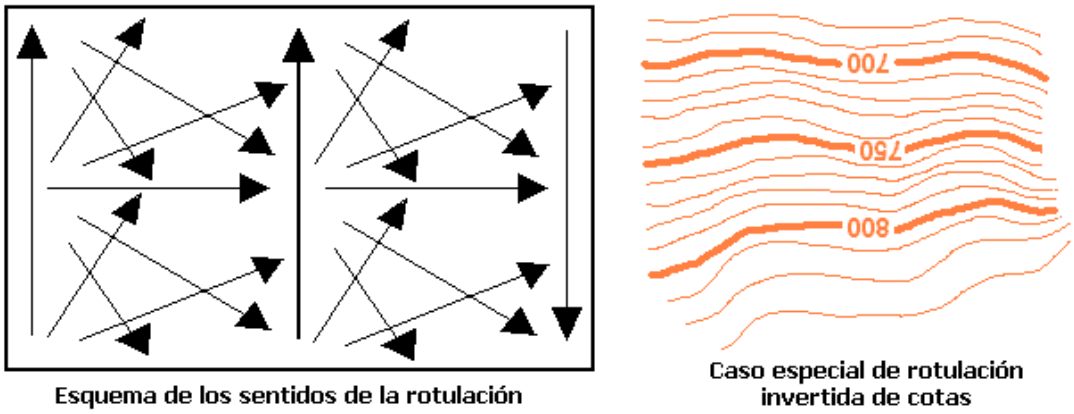


Ilustración 6 - 19: Sentidos preferentes en la rotulación

7. Los rótulos no se curvarán siguiendo elementos geográficos, a menos que sea necesario (en pequeñas escalas algunos rótulos se curvan para indicar la forma del elemento que designan).

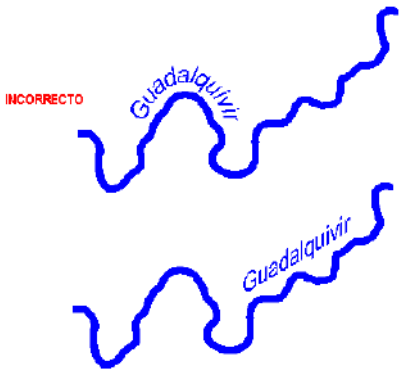


Ilustración 6 - 20: Curvatura de los rótulos

8. En escalas grandes los nombres deben mantener sus letras separadas por el menor espacio posible. En escalas pequeñas las letras se podrán separar para indicar, sobre todo, la extensión de un determinado accidente geográfico. Una posible regla es que la separación entre las letras sea como máximo cuatro veces su altura.

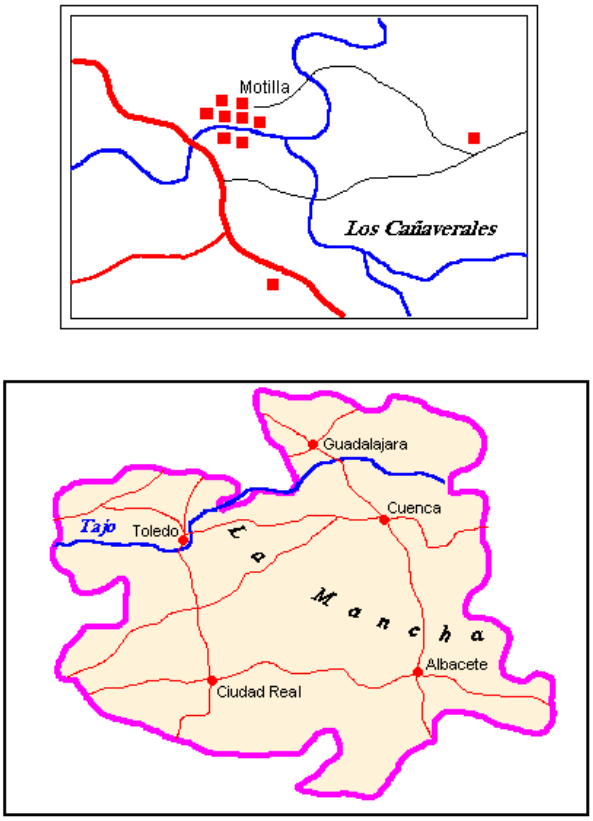


Ilustración 6 - 21: Rótulos con letras separadas

9. Los rótulos que designen detalles puntuales se colocarán preferentemente en el ángulo noreste o siguiendo el orden que se indica en la siguiente figura.

Ilustración 6 - 22: Ubicaciones preferentes del rótulo

Se ha de evitar que los rótulos se confundan, por alineación, con el detalle que designan.

Ilustración 6 - 23: Evitar confusión por alineación de rótulo y símbolo

10. Se evitará que los rótulos pisen las líneas de la cuadrícula o de la red de paralelos y meridianos.

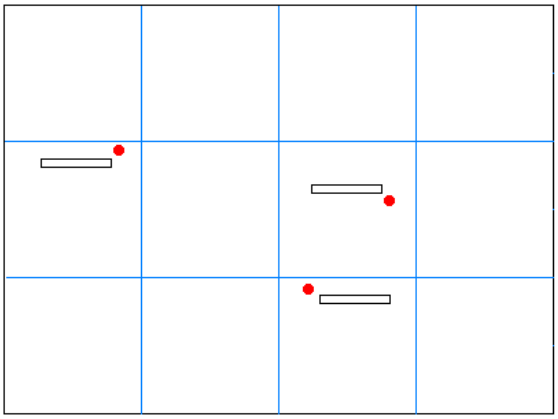


Ilustración 6 - 24: Evitar intersección de rótulo y líneas de cuadrícula

11. Los rótulos que designan diferentes detalles no deben estar exactamente alineados.

Ilustración 6 - 25: Evitar alineación de rótulos

6.1.3. Herramientas habituales aplicables en esta fase

A continuación se describe el software más relevante, en base a la experiencia del IGN de España en la producción de su cartografía básica oficial, que se emplea en esta fase de generación del mapa.

Adobe Illustrator/Avenza Mapublisher

Adobe Illustrator

Es un software dedicado al dibujo vectorial y al diseño de elementos gráficos casi para cualquier tipo de soporte y dispositivo, que tiene una interfaz de usuario sumamente amigable. Es la herramienta de diseño por excelencia y además cada vez está más extendido su uso como aplicación para la producción de cartografía impresa, de hecho en el IGN de España se ha convertido en una herramienta imprescindible puesto que todas las series cartográficas que se publican impresas, y muchos de los productos publicados en web, se elaboran con él.

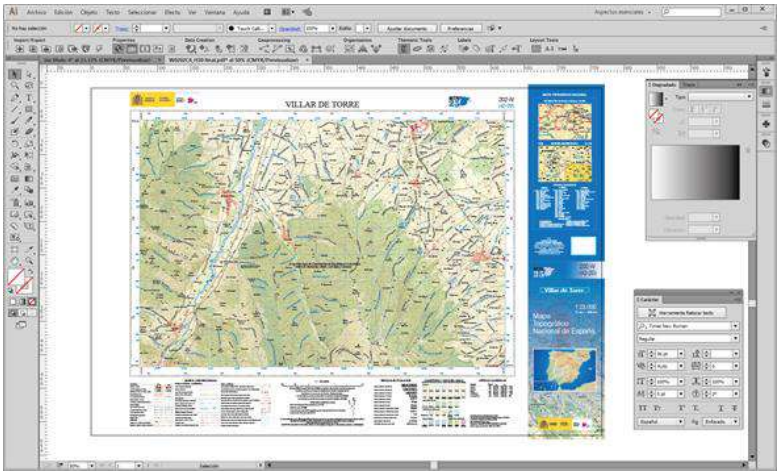


Ilustración 6 - 26: Interface de Adobe Illustrator

Adobe Illustrator sirve principalmente para crear dos tipos de archivos gráficos: vectores y mapas de bits. A grandes rasgos, y sin entrar en conceptos matemáticos, podríamos explicar estos dos conceptos de esta manera:

- **Vectores**: se trata de dibujos basados en puntos que se unen mediante líneas para que la imagen general mantenga la coherencia sin importar la resolución. Se usan para crear diseños de alta calidad que se adapten a cualquier medio donde deban ser utilizados.
- **Mapas de bits**: en este caso, el gráfico se basa en una retícula octogonal coloreada, es decir, una gran cantidad de píxeles que cada uno por separado contiene la información de color que conforma una imagen coloreada al verlos todos juntos. A diferencia de los vectores, los mapas de bits dependen de la resolución para poder tener coherencia, por lo que se debe tener esto en cuenta a la hora de emplearlos.

Con estas dos tipologías de gráficos se satisfacen prácticamente todos los requisitos de diseño de la producción de cartografía en el IGN de España.

Avenza Mapublisher

MAPublisher (de Avenza Systems) es un potente conjunto de complementos y herramientas para Adobe Illustrator que cierra la brecha entre los Sistemas de Información Geográfica (SIG) y el diseño gráfico de alta gama para la creación, la impresión de alta resolución y la publicación electrónica de mapas de alta calidad. MAPublisher permite realizar las tareas orientadas a la publicación cartográfica con las herramientas de gestión de datos SIG adecuadas para facilitar el proceso de producción de mapas, que le permiten importar y exportar desde/hasta infinidad de formatos estándar utilizados en cartografía tanto ráster como vectoriales: .shp, .kml, .gml, . dxf, .dwg, .gen, .e00, .dgn, .tab, etc.

Con este sistema rápido e intuitivo, la creación y publicación (impresa o web) de cartografía vectorial se beneficia de las posibilidades y potencia gráficas de Adobe Illustrator.

Adobe Photoshop / Geografic Imager / Global Mapper

Adobe Photoshop

Adobe Photoshop es un software especializado en la edición de imágenes y que se ha convertido de factor en un estándar a la hora de retocar y manipular fotografías y gráficos dentro del ámbito de la fotografía digital, artes gráficas y diseño gráfico en general.

Entre sus características se encuentran el uso de diversos modelos de color (RGB, CMYK, Cielab, multicanal), distintas profundidades de color (8, 16 y 32 bits), posibilidad de usar colores sólidos, así como semitonos, etc.

Photoshop utiliza sus propios formatos de archivo, aunque es capaz de importar y exportar a una gran cantidad de ellos, entre los cuales podemos encontrar el .jpg, .gif, .png, .tif, etc.

En la producción cartográfica impresa del IGN de España Photoshop encuentra su lugar a la hora de crear sombreados de alta calidad y también a la hora de editar/retocar imágenes ráster (estén georreferenciadas o no).

Geographic Imager

El software Geographic Imager (de Avenza) se instala sobre Adobe Photoshop como un Plug-in, y mejora la funcionalidad de éste para que el trabajo con imágenes espaciales sea mucho más rápido y eficiente. Agrega herramientas para importar, editar, manipular y exportar imágenes aéreas y de satélite.

Geographic Imager posibilita el uso de funciones nativas de Adobe Photoshop, como transparencias, filtros, edición de píxeles y ajustes de imagen como controles de brillo, contraste y curvas, mientras mantiene las referencias espaciales.

El IGN de España utiliza Geographic Imager como complemento de Photoshop para mantener los datos de georreferenciación de imágenes ráster durante la edición o retoque de las mismas.

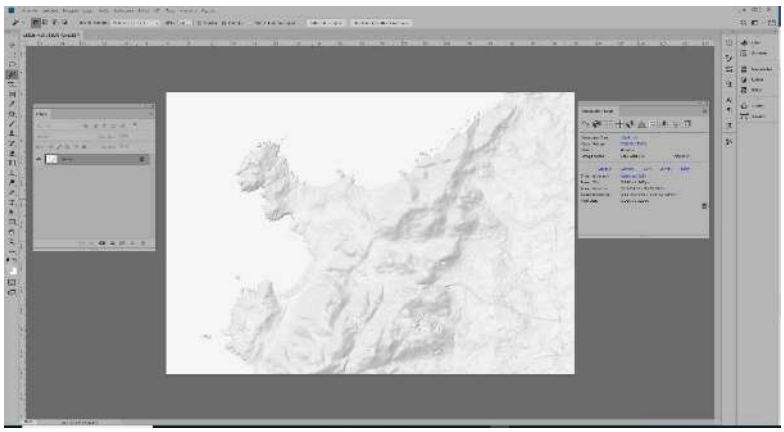


Ilustración 6 - 27: Interface de Photoshop + Geographic Imager

Global Mapper

Global Mapper (Danysoft) es un software SIG que proporciona una gama completa de herramientas de procesamiento de datos espaciales, con acceso a una variedad incomparable de formatos de datos utilizados en el mundo del CAD, SIG e Ingeniería.

Global Mapper es capaz de mostrar, convertir y analizar prácticamente cualquier tipo de datos geoespacial ya sea en 2D o 3D, de una nube remota o local, de datos raster o vectoriales, archivo plano o base de datos espacial. El software convierte, edita, publica, imprime, crea mosaicos, particiones, y permite al usuario realizar análisis espaciales avanzados.

En el ámbito de la producción cartográfica del IGN de España, Global Mapper se encarga de interpretar los Modelos Digitales del Terreno para producir a través de ellos los sombreados que se aplicarán a las distintas hojas y series cartográficas (MTN25, MTN50, MP200, etc.).

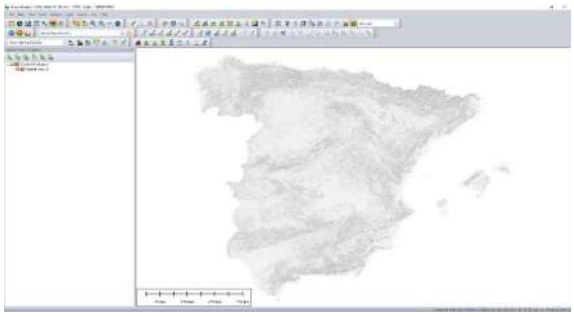


Ilustración 6 - 28: Interface de Global Mapper

FME (Feature Manipulation Engine)

A este software ya se le ha hecho referencia en capítulos anteriores, fundamentalmente en el capítulo 5, pues es la herramienta principal sobre la que se ha implementado todo el proceso de producción automática de los datos cartográficos con los que se generará el mapa. Igualmente, en esa segunda fase también se emplea en la automatización de los procesos asociados a la simbolización y rasterización, tal y como se explica en la sección 6.3.2.

El motor de manipulación de características (FME) es una plataforma que agiliza la traducción de datos espaciales entre formatos geométricos y digitales, es decir, carga los archivos en el formato de origen, los manipula a través de transformadores y, finalmente, los genera en el formato que se necesite. FME admite múltiples formatos, desde formatos de bases de datos hasta formatos vectorial y ráster de diversos tipos y su principal ventaja es su gran potencial de combinación y explotación de los datos desde un entorno visual que lo hace mucho más amigable que las consolas de otros lenguajes de programación.

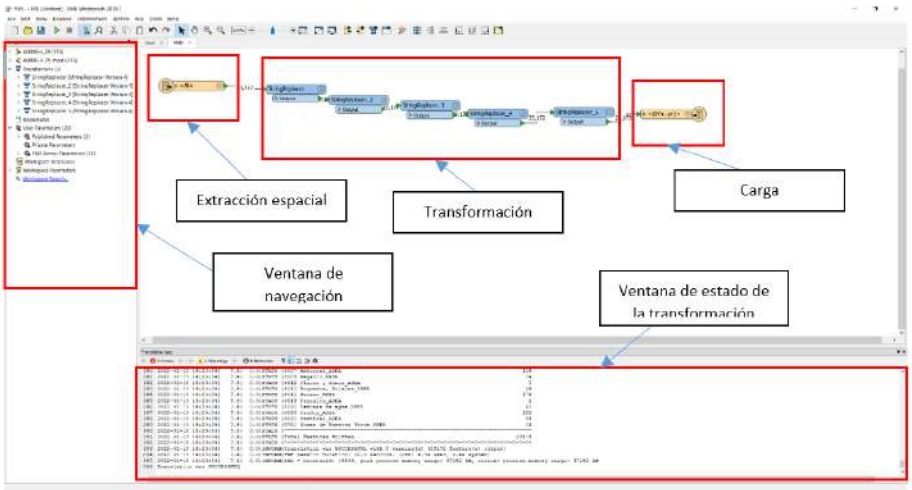


Ilustración 6 - 29: Interface de FME

6.2. Modelo Digital del Terreno (MDT) y Sombreado

Los Modelos Digitales de Terreno (MDT) constituyen uno de los modos básicos de representación digital de la superficie terrestre. Un MDT es una estructura numérica de datos que representa la distribución espacial de una variable cuantitativa y continua; cuando dicha variable es la elevación, entonces se trata del caso particular denominado Modelo Digital de Elevaciones (MDE).

6.2.1. Modelo Digital de Elevaciones (MDE)

Existen diversas formas de representar la elevación del terreno, ya sea por medio de datos vectoriales como son las curvas de nivel, por redes de triángulos irregulares o por medio de datos en formato raster, es decir, una representación matricial regular de la variación continua del relieve. Éste último es habitualmente el más empleado a la hora de definir los MDE, fundamentalmente debido a la facilidad con la que permite explotar su información.

Un MDE puede representarse de forma genérica mediante la ecuación: $z = f(x, y)$, que define un campo de variación continua. La imposibilidad de resolver la ecuación anterior para **todos** los puntos del territorio obliga a definir elementos discretos sobre el mismo que permitan simplificar la codificación de la elevación.

Curvas de nivel

Las curvas de nivel son líneas, definidas como una sucesión de pares de coordenadas que tienen como identificador, el valor de la elevación en cada uno de los puntos de la línea y son el resultado de intersectar el terreno con planos paralelos al nivel del mar y separados por un valor constante de equidistancia.

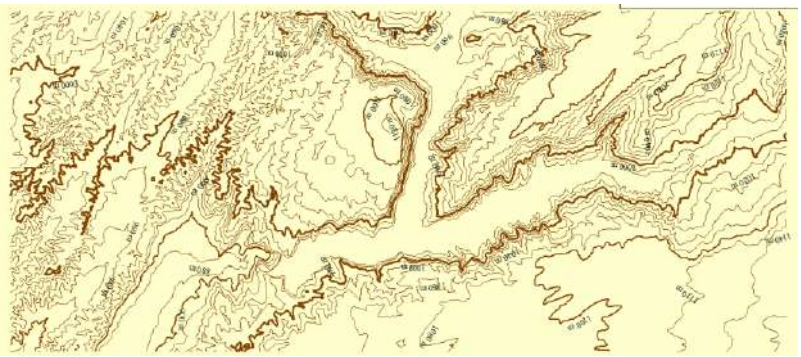


Ilustración 6 - 30: Modelo del terreno con curvas de nivel

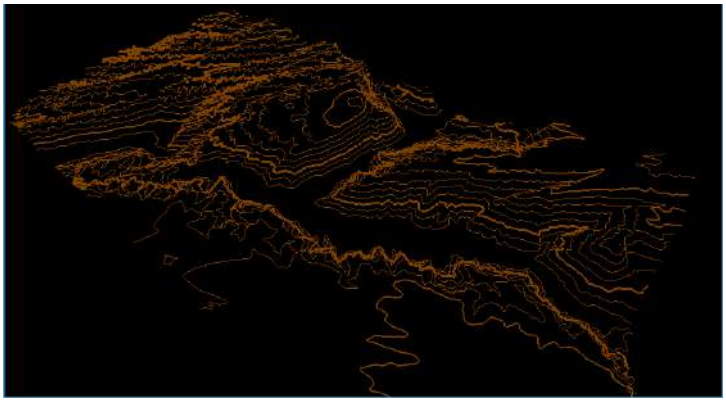


Ilustración 6 - 31: Perspectiva 3D de las curvas de nivel

Red Irregular de triángulos (TIN)

Una red TIN se genera a partir un conjunto de puntos de elevación conocida, al trazar triángulos formados por tripletas de puntos cercanos y no colineales, y formar así un mosaico. Hoy día lo ideal es partir de un conjunto de puntos LIDAR. Los modelos TIN tienen ventajas al adaptarse mejor a las irregularidades del terreno, ocupar menos espacio y dar muy buenos resultados a la hora de visualizar modelos en 3D o determinar cuencas visuales. Sin embargo, adolecen del inconveniente de necesitar un mayor tiempo de procesamiento y de resultar bastante ineficientes cuando se intenta su integración con otro tipo de información.

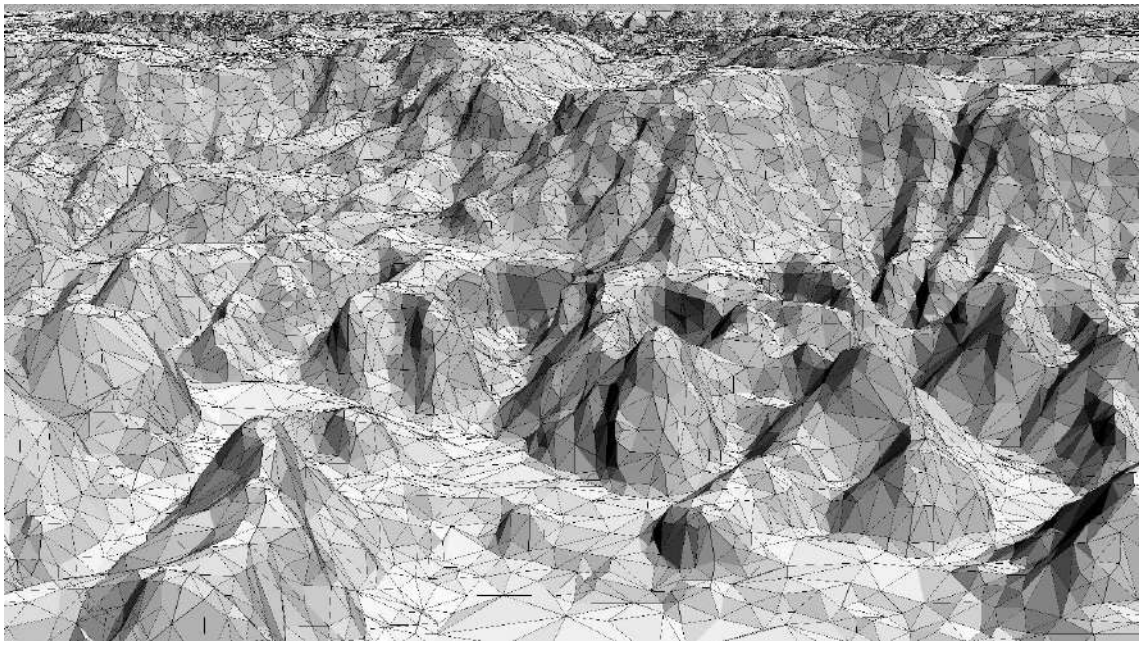


Ilustración 6 - 32: Perspectiva 3D de un modelo TIN

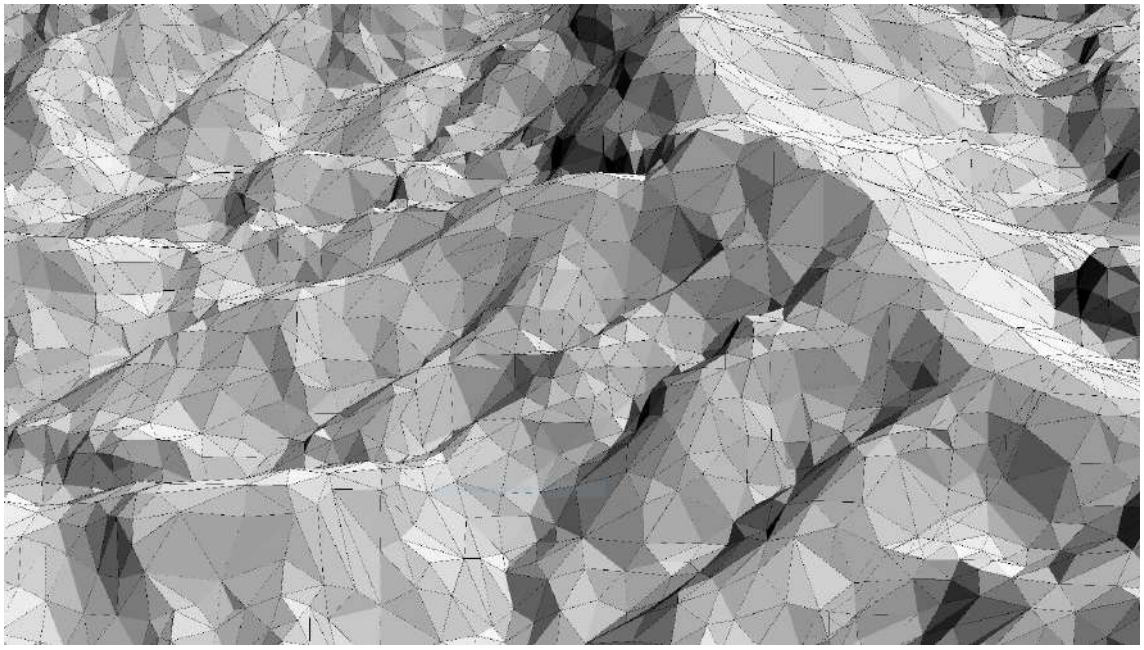


Ilustración 6 - 33: Perspectiva 3D más cercana de un modelo TIN

Malla regular (grid)

Como se ha comentado anteriormente, es muy adecuado para la integración de las elevaciones en un SIG porque así se pueden utilizar diversas herramientas para la obtención de nuevos resultados a partir del MDE.

En el caso del IGN de España, se emplea para obtener un sombreado de la zona cartografiada y superponerlo a ella, con el fin de ayudar al usuario en la comprensión del relieve del mapa que está leyendo.

A continuación, a título de ejemplo, se muestra y describe el aspecto numérico que presenta la malla regular generada a partir de la información altimétrica del Modelo Digital del Terreno (MDT05) que actualmente se tiene en el IGN de España, y que consta de una matriz de puntos separados 5 m entre sí.

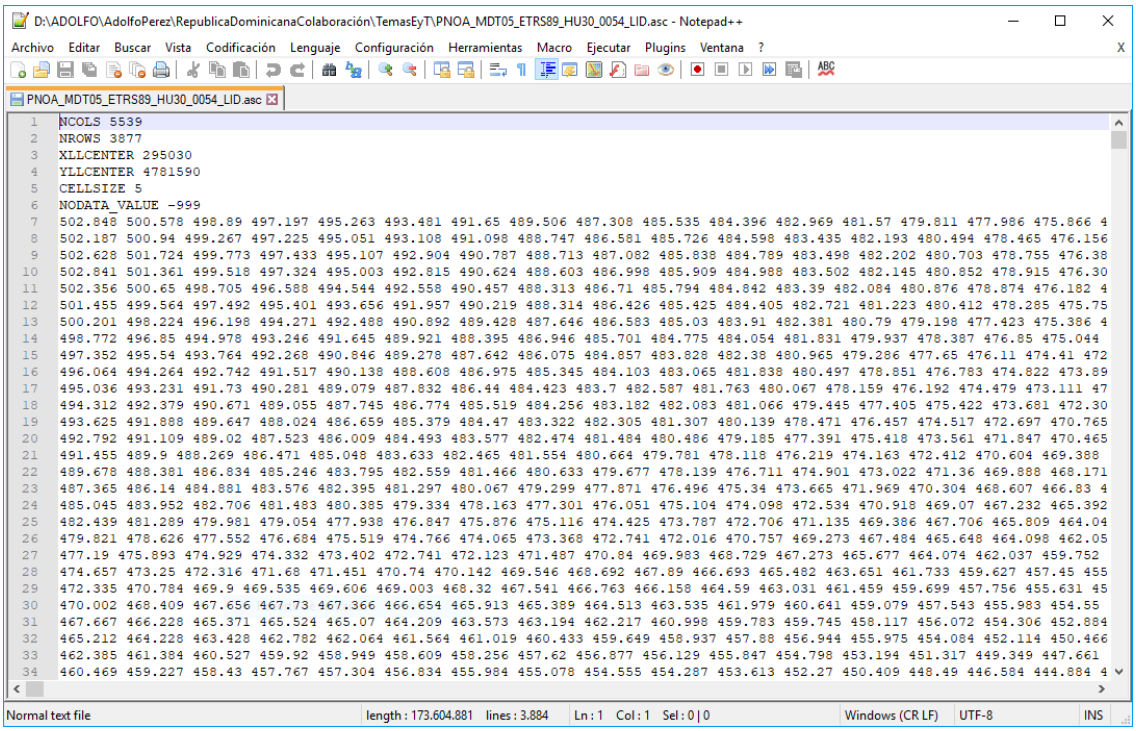


Ilustración 6 - 34: Aspecto numérico de una malla regular

La Ilustración anterior muestra la malla correspondiente a la hoja número 54 del Mapa Topográfico Nacional a escala 1:50,000 (MTN50); su cabecera indica que la coordenada X de la esquina inferior izquierda (SW) es 295,030 y la coordenada Y es 4,781,590, que la matriz de puntos consta de 5,539 columnas por 3,877 filas, y que el paso de malla es de 5 m. Después de declarar que -999 será el valor que se escriba cuando un punto no tenga Z conocida, viene la fila 7 del archivo que se corresponde con la fila inferior del MDT y muestra el conjunto de 5,539 valores de Z correspondientes a esa fila, teniendo el punto inicial un valor de 502,848. Así, los dos primeros puntos de la primera fila tendrían las siguientes coordenadas.

X_1= 295,030 ; Y_1= 4,781,590 y Z_1= 502,848

X_2= 295,035 ; Y_2= 4,781,590 y Z_2= 500,578

Y así sucesivamente hasta agotar todas las columnas de la fila 7 y luego acometer la siguiente fila (8) que tendría una coordenada Y constante de 4781595.

El primer punto de la fila 8 tiene las siguientes coordenadas:

X_{5540}= 295,030 ; Y_{5540}= 4,781,595 y Z_{5540}= 502,187

X_{5541}= 295,035 ; Y_{5541}= 4,781,595 y Z_{5541}= 500,94

De este modo, cualquier programa que interprete este formato será capaz de reproducir la forma del terreno, tal como se aprecia en la siguiente imagen, que se corresponde exactamente con los datos mostrados arriba de la hoja 54 del MTN50.

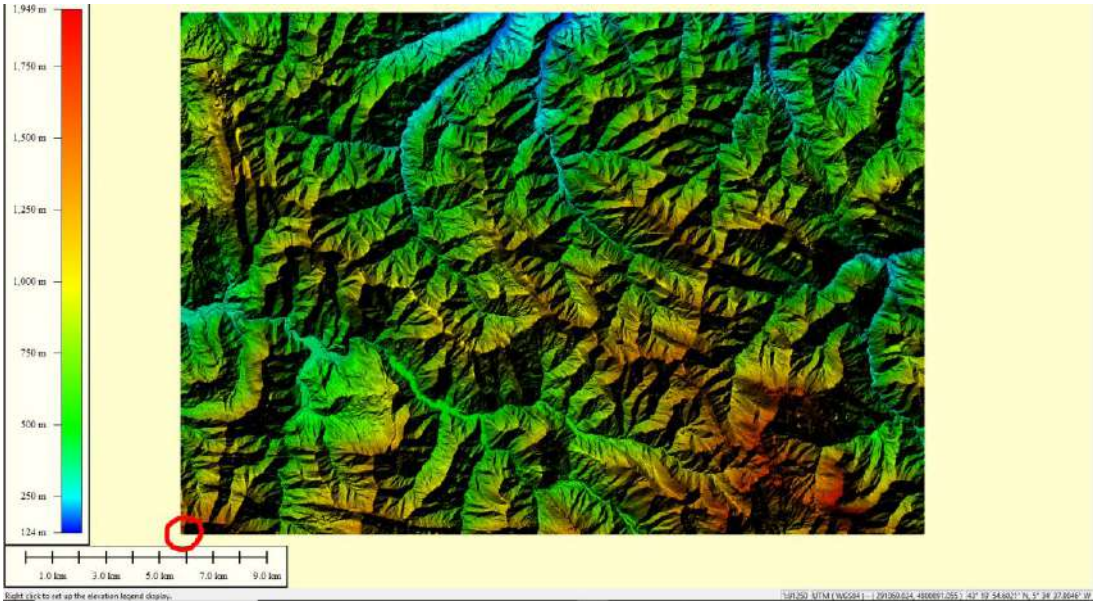


Ilustración 6 - 35: Aspecto gráfico de una malla regular interpretada por un SIG, con origen en esquina suroeste

A partir de un MDT se pueden realizar análisis morfológicos del terreno, por medio de programas SIG. En la imagen siguiente se visualiza la planta de un embalse a la que se han añadido las curvas de nivel con una equidistancia de 20 m, extraídas a partir del propio MDT.

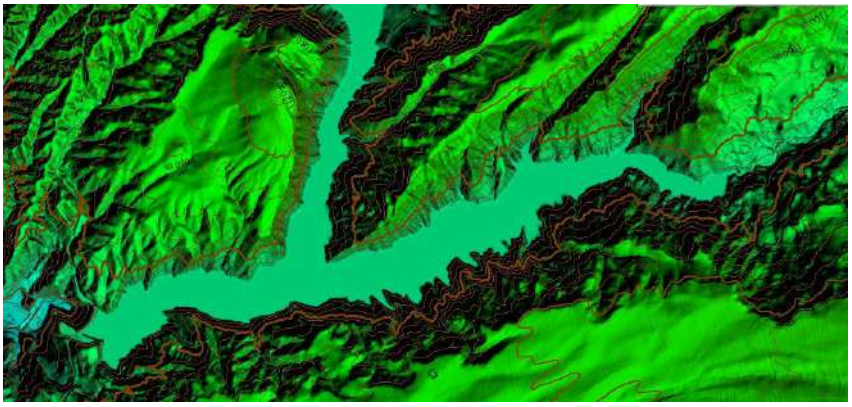


Ilustración 6 - 36: Planta embalse y curvas de nivel

Y en esta otra imagen una perspectiva 3D de la misma zona.

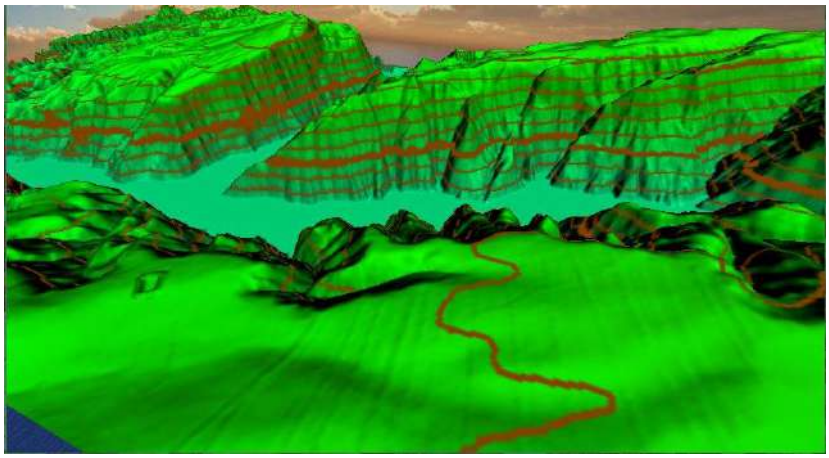


Ilustración 6 - 37: Perspectiva 3D del embalse

6.2.2. El sombreado

El sombreado cartográfico no es más que una representación bidimensional del modelo digital del terreno (modelo tridimensional), es decir, es una interpretación del MDT que, una vez integrada debidamente en el mapa, ayuda a la interpretación del relieve de la zona representada y realza y embellece la publicación, dotándola de individualidad y, por qué no, de una gran carga de dinamismo visual en las hojas en las que el relieve es más acusado.

Como ejemplo del realce visual obtenido, se muestra una hoja del MTN con y sin sombreado:

Ilustración 6 - 38: Vista parcial de la hoja 552-2 sin sombreado

Ilustración 6 - 39: Vista parcial de la hoja 552-2 con sombreado

Ilustración 6 - 40 : Vista parcial del sombreado de la hoja 552-2

Obtención del sombreado a partir del Modelo Digital del Terreno

Esta sección se documenta en base a la experiencia del IGN de España en el proceso de dotar de sombreado a la cartografía que produce.

La combinación del mapa con el sombreado se realiza gracias a que el ráster del sombreado está debidamente georreferenciado en el mismo sistema de coordenadas que la hoja a la que corresponde.

Software utilizado

Para la obtención de un sombreado óptimo, y en base a la experiencia adquirida a lo largo del tiempo, se utiliza el siguiente software:

- Global Mapper: que permite gestionar el Modelo Digital del Terreno y exportarlo como una imagen raster en escala de grises.

- Adobe Photoshop: editor de imágenes muy utilizado en diseño gráfico y que permite manipular y editar imágenes ráster con total fluidez.
- Geographic Imager (Avenza): plugin para Photoshop que nos permite importar, editar, manipular y exportar imágenes georreferenciadas a/desde Photoshop en diversos formatos estándar usados en cartografía.

  Con Geographic Imager se puede cambiar de sistema de coordenadas, escala y proyección un ráster georreferenciado con Photoshop, y exportarlo de nuevo con los cambios realizados.

  De este último Plugin se puede prescindir, aunque sin él no es posible hacer transformaciones ni cambios de escala de manera directa. Su instalación puede simplificar un poco la manipulación y edición del sombreado con Photoshop.

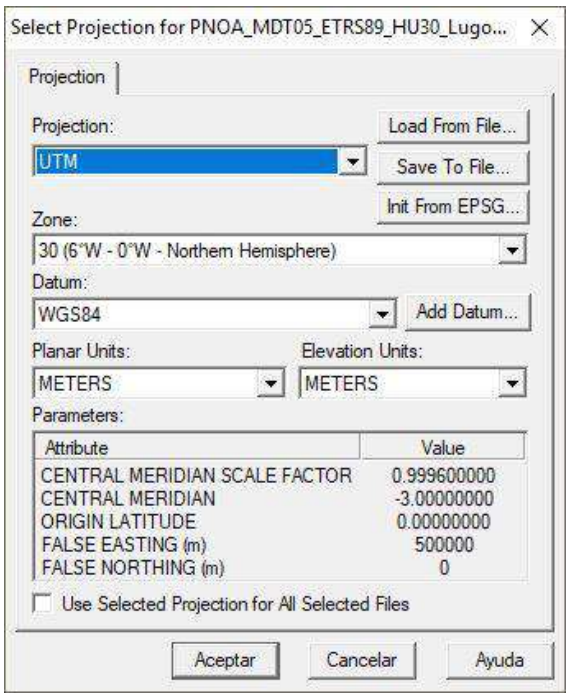
Proceso de obtención del sombreado

Los modelos digitales utilizados para la producción del sombreado tienen como unidad una provincia y son de una resolución de captura de 5 metros por pixel, que es más que suficiente para poder aplicarla a los mapas impresos y publicados en servicios de visualización que genera el IGN DE ESPAÑA.

El proceso de obtención del sombreado se compone de tres fases que se describen a continuación.

1. **Importación del MDT a Global Mapper y exportación de archivos .tiff georeferenciados**

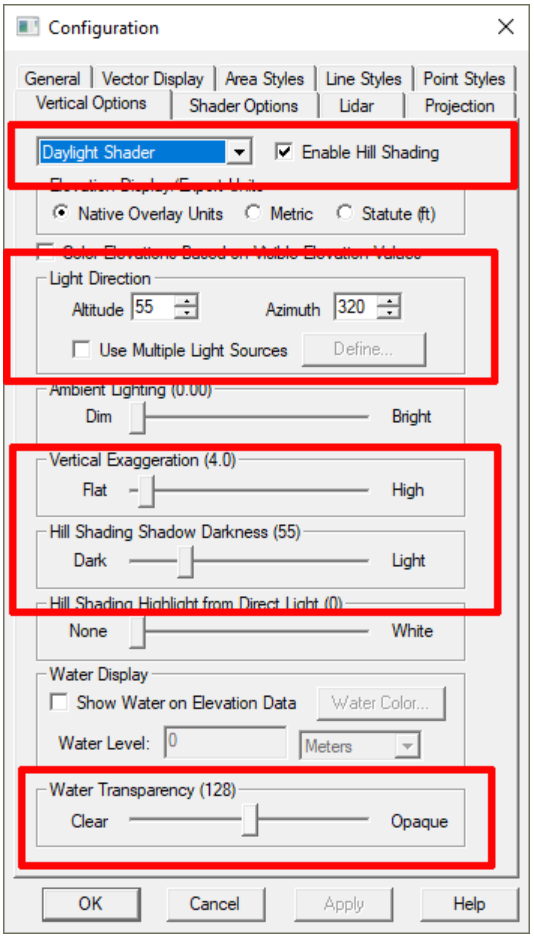
Se carga el MDT en el programa añadiéndolo a través de los menús (Open datafiles) o arrastrando el archivo .asc elegido a la ventana principal de Global Mapper. A la hora de cargar el modelo digital, el software nos solicita que confirmemos el sistema de referencia utilizado. Aquí deberemos indicar la proyección, la zona y el datum utilizado, así como las unidades de medida (planas y de elevación).

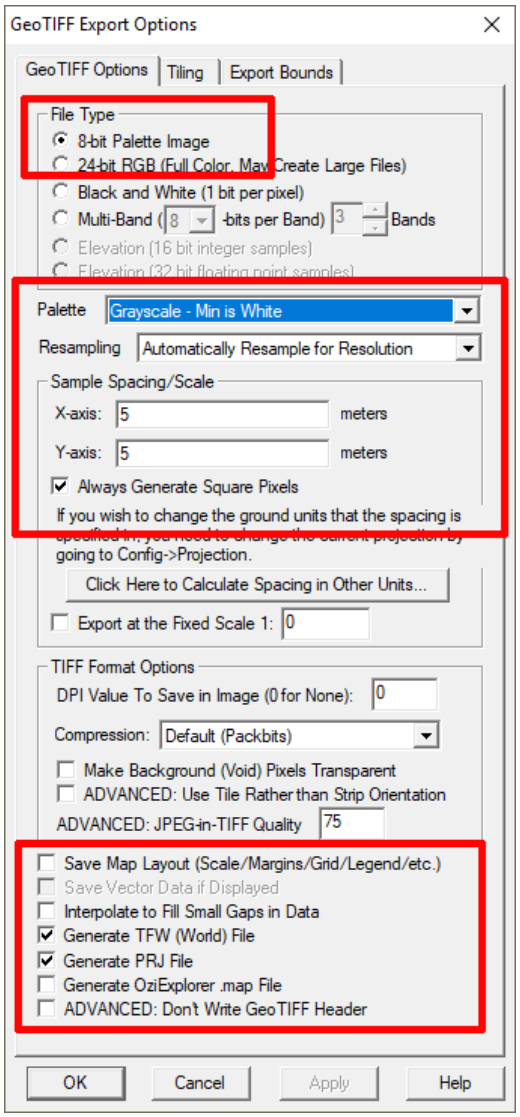


El archivo MDT (Surfer grid o ASCII) se carga en el Global Mapper y a partir de éste se le aplican dos shaders (sombreados) a través del menú **Tools→Configure**

En un primer paso aplicaremos al modelo digital el **Daylight Shader** (mapa de sombras convencional) con los parámetros que se muestran en la siguiente Ilustración 6 - (ventana Configuration).

Una vez aplicado el shader, procederemos a generar un primer archivo **Geotiff** desde el menú **File→Export→Raster image format** aplicando los parámetros que se muestran en la captura de pantalla (ventana Geotiff Export Options).

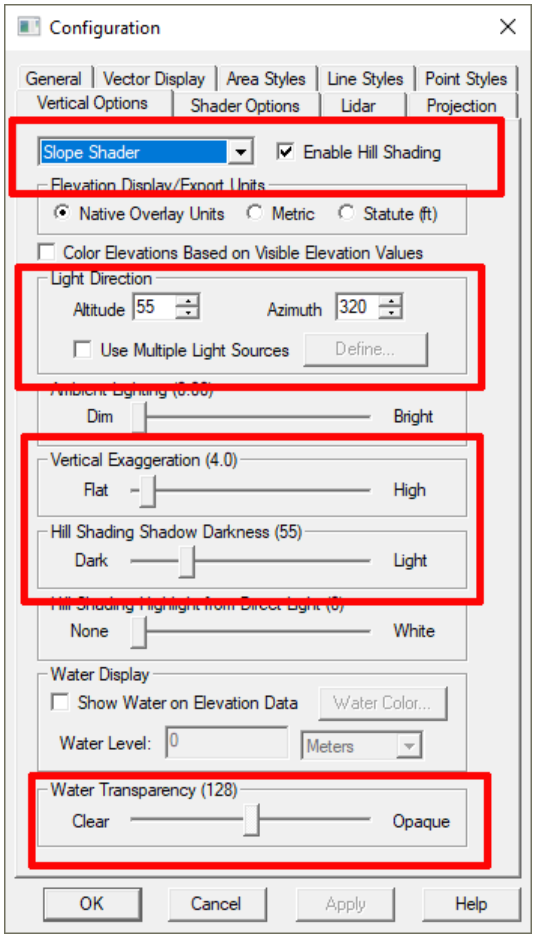


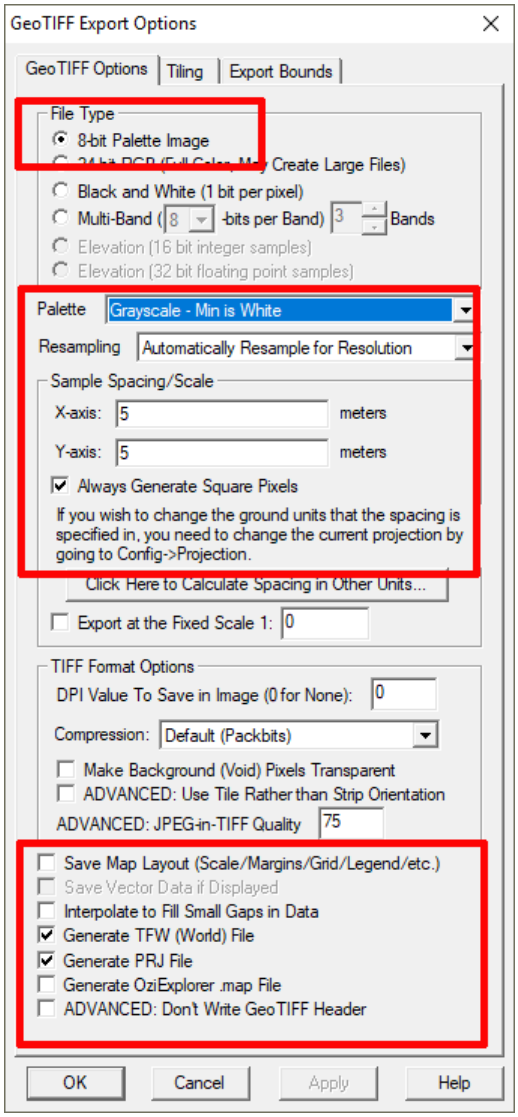


Para evitar problemas a posteriori debemos asegurarnos de que el archivo tiff va a ser exportado en escala de grises. A este archivo lo denominaremos **(nombre de archivo)+"normal"** y el resultado de la exportación deberá ser un GEOTIFF con sus archivos de proyección y georreferenciación (tres archivos en total).

Como segundo paso aplicaremos al modelo digital el **Slope Shader** (mapa de pendientes) con los parámetros que se muestran en la siguiente Ilustración 6 - y que nos sirve para aclarar las zonas llanas.

El archivo **Geotiff** lo generaremos desde el menú **File→Export→Raster image format** aplicando los parámetros que se muestran en la captura de pantalla (los mismos que en el geotiff anterior).





El Geotiff resultante lo denominaremos como (nombre de archivo)+"cenital" para diferenciarlo del anteriormente obtenido. Asimismo, y como en el caso anterior, deberemos tener la precaución de exportar a la vez los archivos .tfw y .prj, indicándolo en la ventana de exportación.

Estos dos archivos de proyección y georreferenciación (documentos de texto) son muy importantes si se va a hacer la edición de las imágenes con Photoshop (sin tener instalado Geographic Imager), puesto que contienen los datos de georreferenciación de la imagen. Deben tener el mismo nombre, aunque distinta extensión (prj o tfw) y deberán acompañar siempre al TIFF para que el software en el que vayamos a utilizar posteriormente la imagen reconozca sus coordenadas.

Una vez que hayamos obtenido los dos Geotiff con sus correspondientes archivos de georreferenciación procederemos a obtener el sombreado con Photoshop, que es el software que nos va a permitir obtener unos resultados óptimos.

A continuación, se muestran los resultados de la exportación de los dos archivos Geotiff (normal y cenital) de la provincia de Bizcaia (España). Se puede comprobar que cada uno de ellos tiene un aspecto bastante distinto del otro, aunque ambos parten del mismo modelo digital. Lo que los diferencia es la aplicación **de dos shaders diferentes (Daylight Shader y Slope Shader** respectivamente).

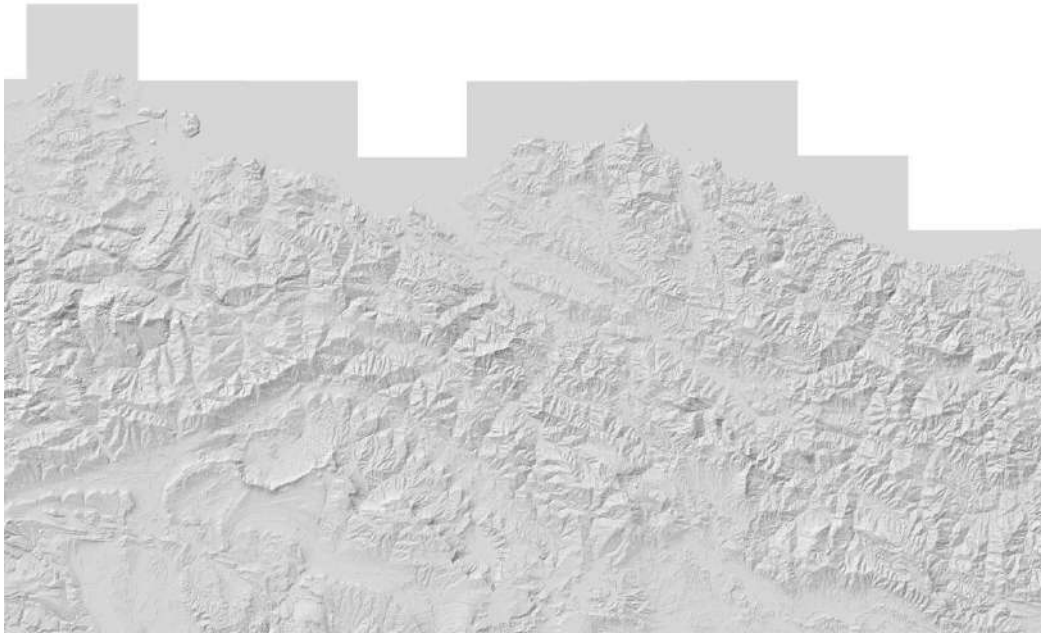


Ilustración 6 - 41 Geotiff "normal" del MDT de Bizkaia

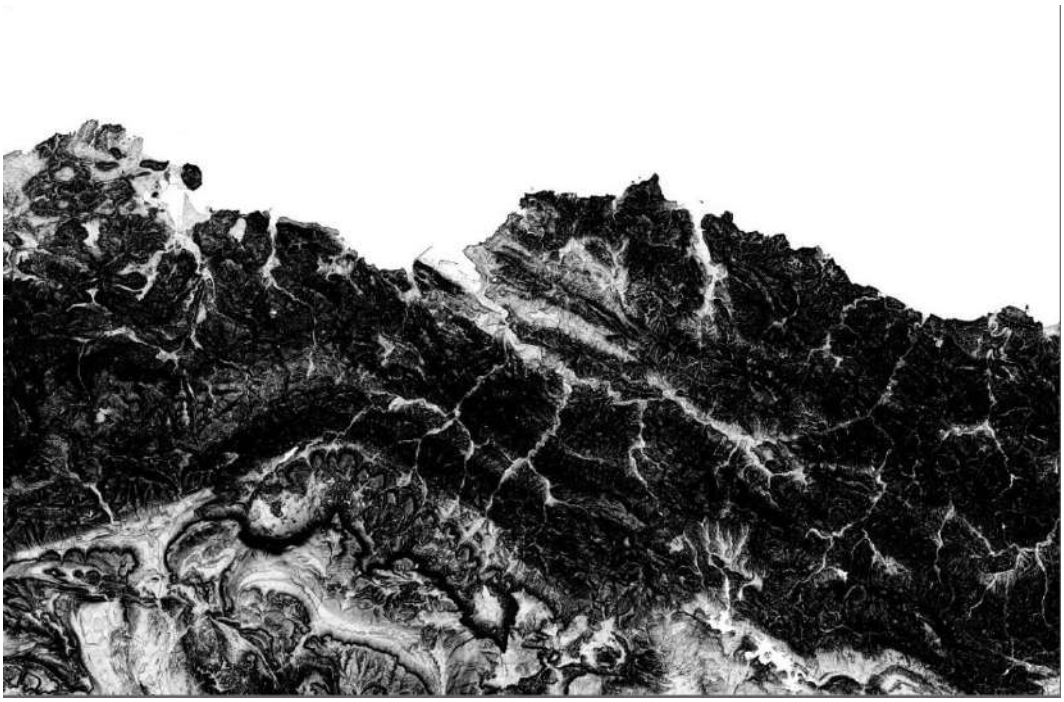
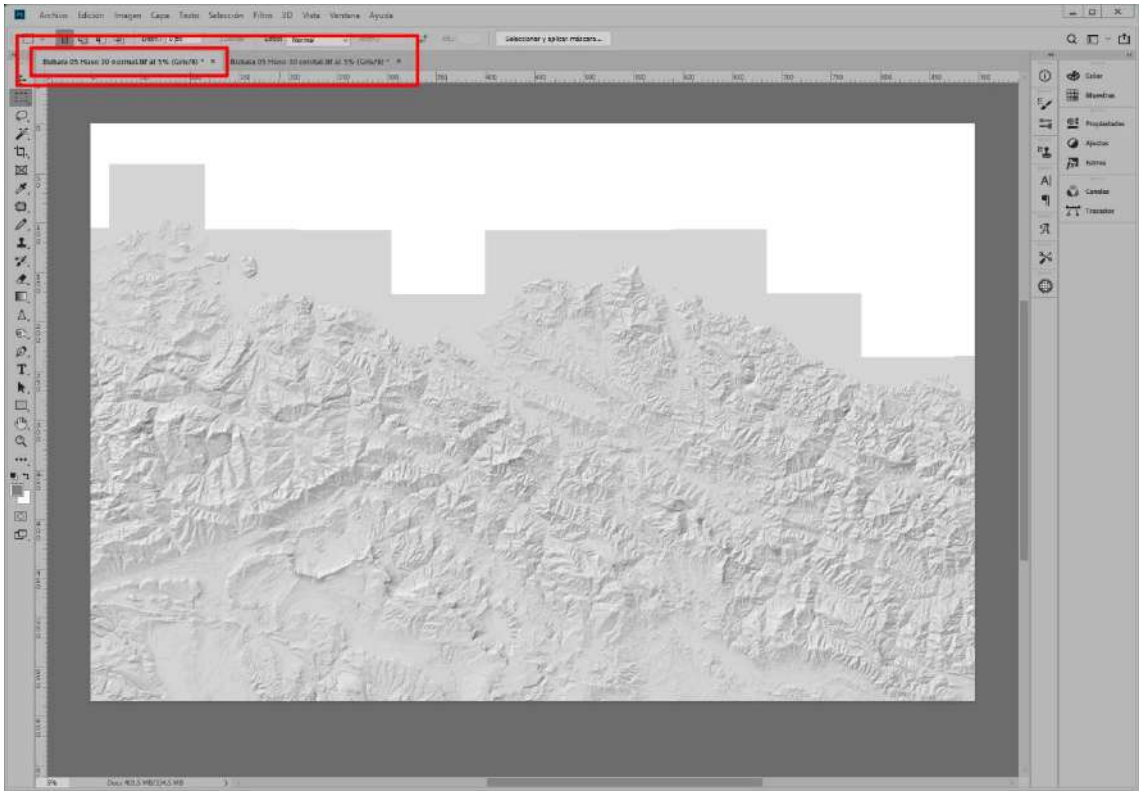


Ilustración 6 - 42 Geotiff "cenital" del MDT de Bizkaia

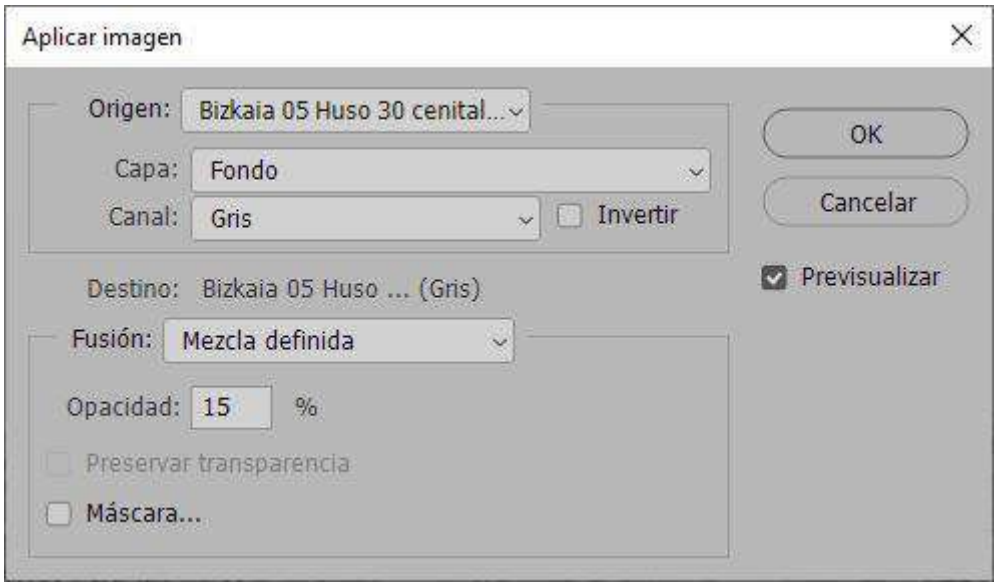
2. Manipulación de archivos Geotiff en Photoshop

Se insiste en que ambos archivos Geotiff deben estar en el modo de color Escala de Grises. Si no fuera así, podemos cambiarlos ahora desde el menú Imagen→Modo→Escala de Grises de Photoshop. Esto nos asegura en todo momento que los resultados obtenidos sean los ideales.

Lo primero que deberemos de hacer es abrir los dos geotiff con Photoshop a la vez, pero teniendo en primer plano el archivo "normal".



Después, deberemos ir al **Menú Imagen→Aplicar Imagen** donde nos aparecerá una ventana similar a la siguiente:

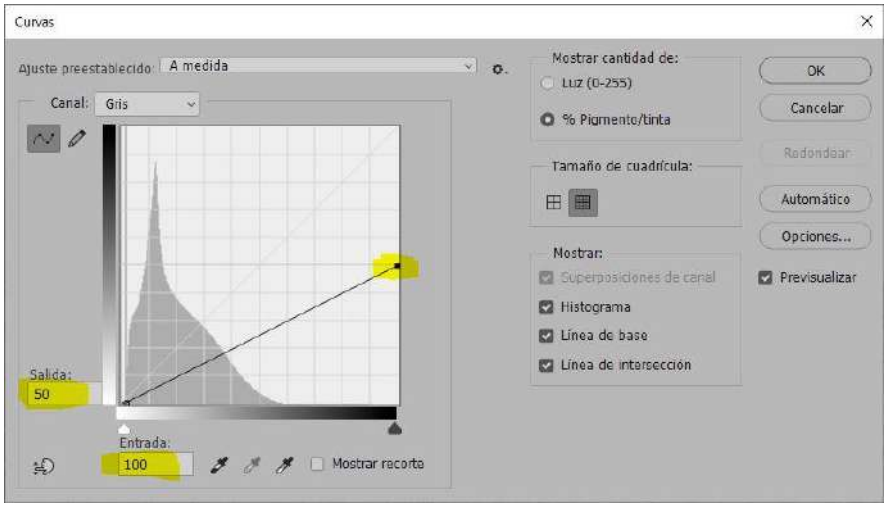


Ahora estableceremos los siguientes valores:

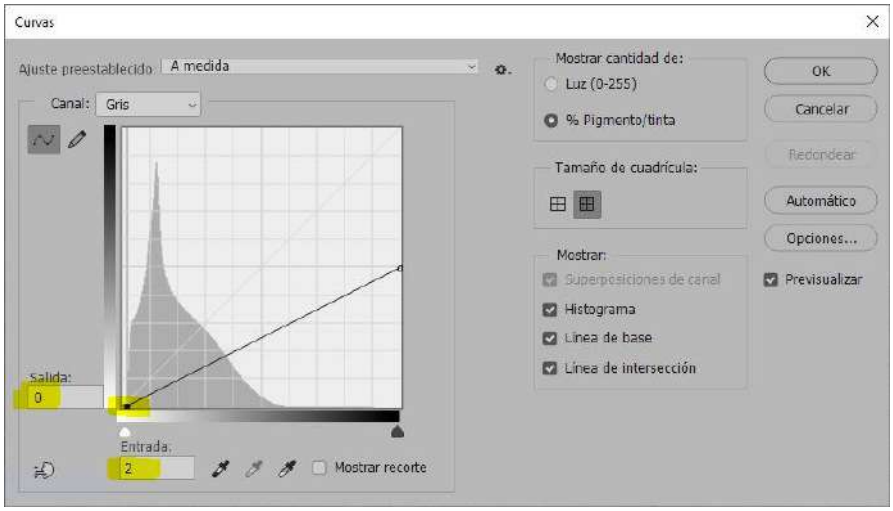
- En el menú desplegable **Origen** pondremos el archivo **"cenital".**
- En el menú desplegable **Fusión** pondremos **Mezcla definida.**
- En Opacidad pondremos 15%

Como se puede comprobar, el resultado de esta manipulación es evidente, el sombreado ha tomado una apariencia bastante buena, y casi podríamos decir que nos valdría para aplicarlo en los mapas necesarios.

Ahora deberíamos aplicar una curva de Photoshop para que el tono de las zonas más abruptas y las pendientes más inclinadas baje a un nivel aceptable. Eso lo haremos desde el Menú **Imagen→Ajustes→Curvas**, donde nos aparecerá una ventana similar a ésta y donde pondremos los siguientes parámetros modificando la curva:



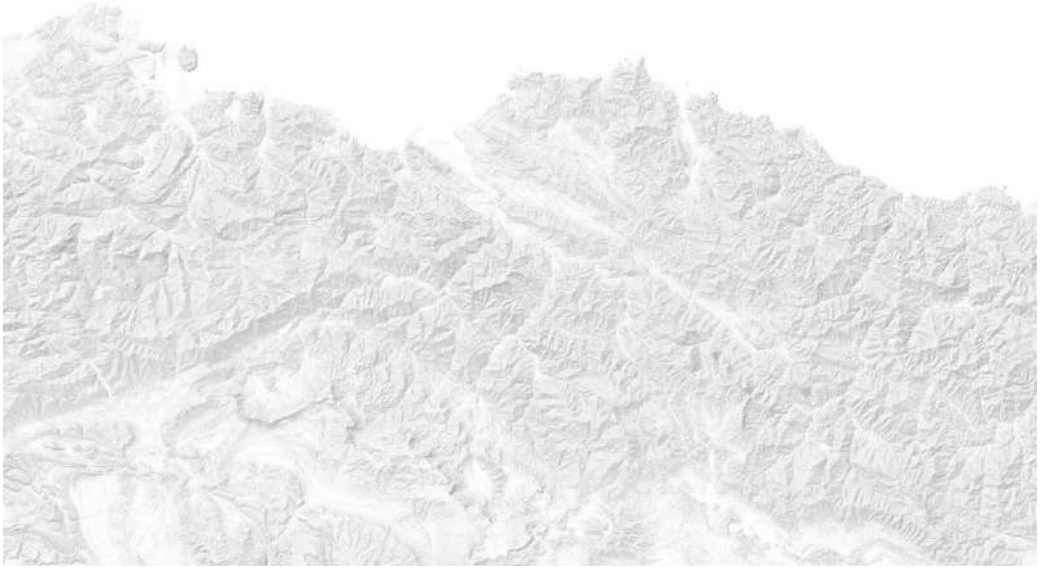
Sombras: Entrada 100 y Salida 50



Luces: Entrada 2 y Salida 0.

La curva se puede guardar para ser reutilizada posteriormente en otros sombreados.

El resultado de manipular la curva del sombreado es el siguiente:



El archivo resultante lo salvamos como **(nombre de archivo)+"final"** (para diferenciarlo de los anteriores "normal" y "cenital"), y acto seguido copiamos los archivos de georreferenciación de cualquiera de las imágenes originales y los renombramos igual que el final.

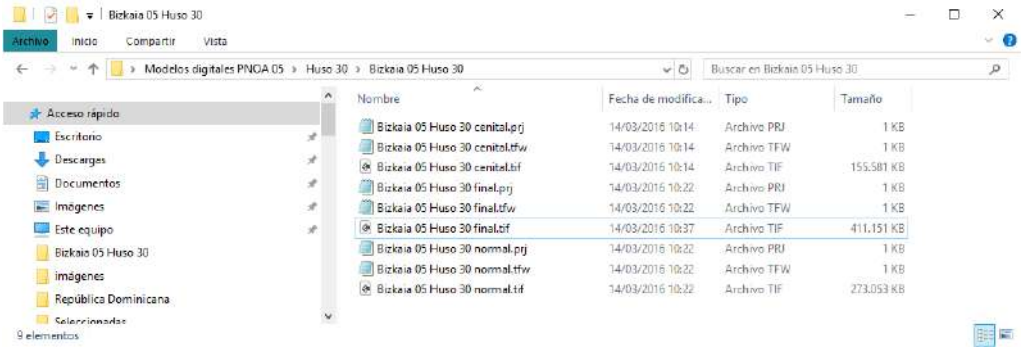


Ilustración 6 - 43 Conjunto de archivos que han intervenido en la obtención del sombreado

Ahora podemos decir que el sombreado está terminado y podríamos utilizarlo en su mapa correspondiente sin mayor problema, gracias a los archivos de georreferenciación que hemos creado y mantenido durante la manipulación del ráster.

Realización de un mosaico de provincias en Global Mapper

Planteamiento

Tal y como se ha comentado anteriormente, la unidad de producción del MDT en el IGN de España es la provincia, pero la unidad de publicación generalmente es la hoja a escala 1:25,000.

Teniendo en cuenta que una provincia da cabida a muchas hojas a escala 1:25,000, lo lógico sería que en cada hoja publicada se colocara simplemente el ráster de sombreado con el ámbito necesario para ésta, evitando incorporar datos ráster que no se van a utilizar, puesto que lo único que se consigue con ello es entorpecer la labor y el resultado final.

El fin perseguido con este planteamiento es tener un sombreado individual para cada una de las hojas de la escala 1:25,000, con su misma numeración y ámbito geográfico.

En el caso de España, surge cierta complicación en la organización del trabajo por el hecho de que la zona peninsular y el archipiélago Balear se encuentran en tres husos UTM: 29, 30 y 31. Por su parte, Canarias se encuentra mayoritariamente en el 28. Para simplificar el trabajo con las más de 4,000 hojas que componen la serie 1:25,000, se emplea el huso 30 extendido para la península y Baleares, manteniendo el 28 para Canarias.

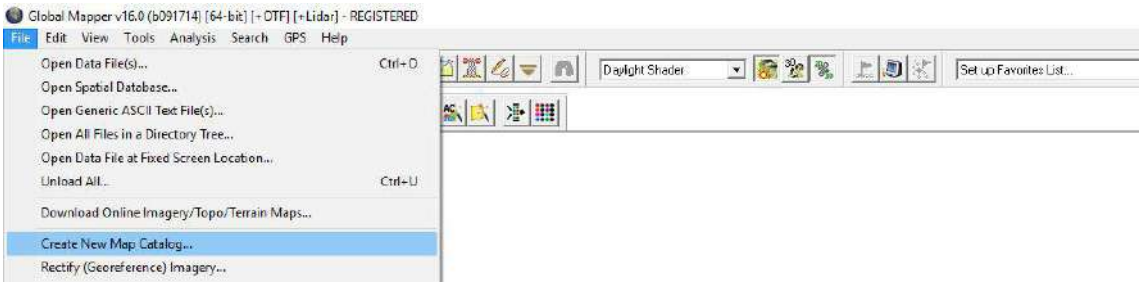
En el caso de República Dominicana, donde la mayor parte de su territorio se encuentra en el huso 19, se recomienda utilizar este huso de forma extendida.

Ejecución

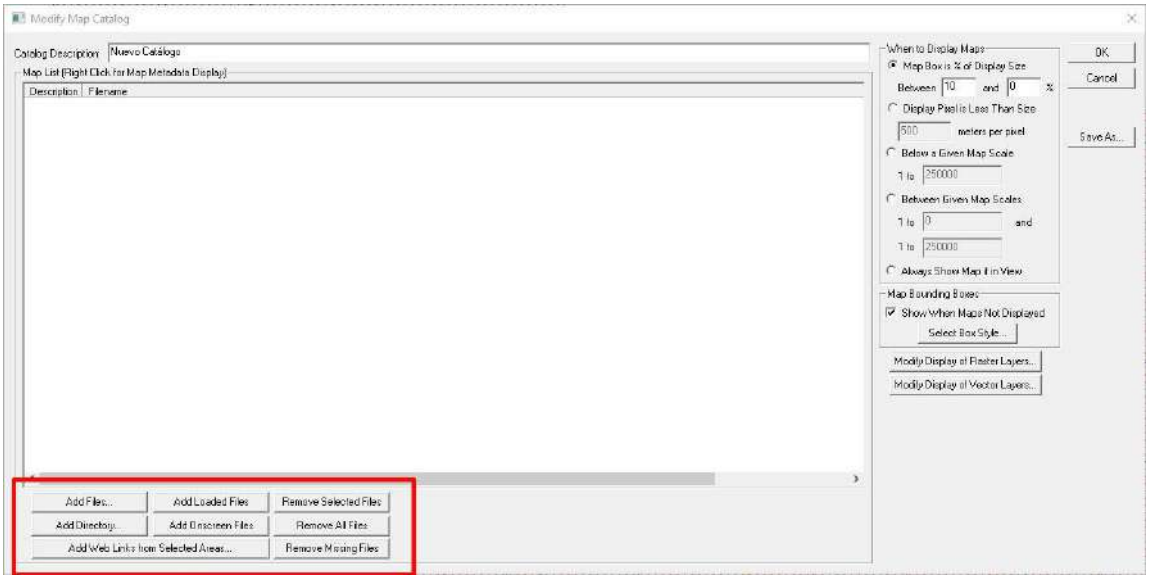
Una vez obtenidos los raster “finales” de todos los sombreados por provincias, lo que resta es ir agregandolos a un “catálogo” de Global Mapper para generar un mosaico de sombreados.

El catálogo es un tipo de archivo que gestiona los vínculos o enlaces a ciertos elementos/archivos (en este caso los ráster del sombreado), haciendo que el manejo y gestión de la información sea fluido y eficiente al no tenerlos “cargados” completamente dentro de un espacio de trabajo.

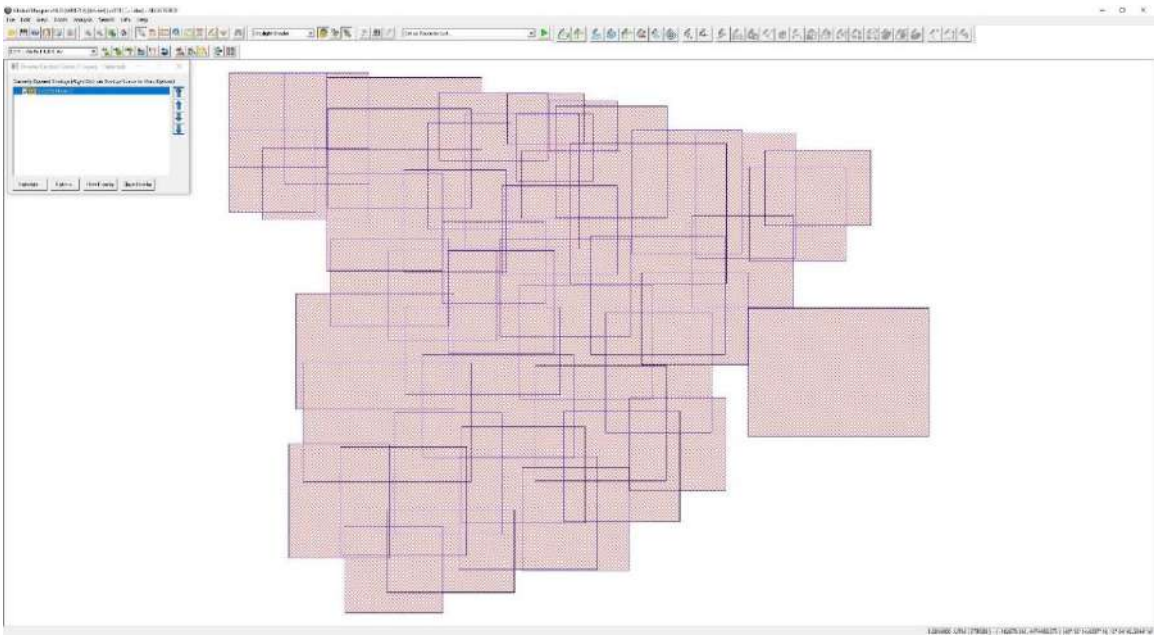
Para crear un nuevo catálogo, deberemos ir al menú **File→Create New Catalog**.



El programa nos solicita el nombre del catálogo y directorio donde lo queremos guardar y acto seguido nos muestra la siguiente ventana, desde la cual podemos ir cargando y gestionando los vínculos a los archivos que conformarán nuestro catálogo, utilizando los menús de la zona inferior izquierda:

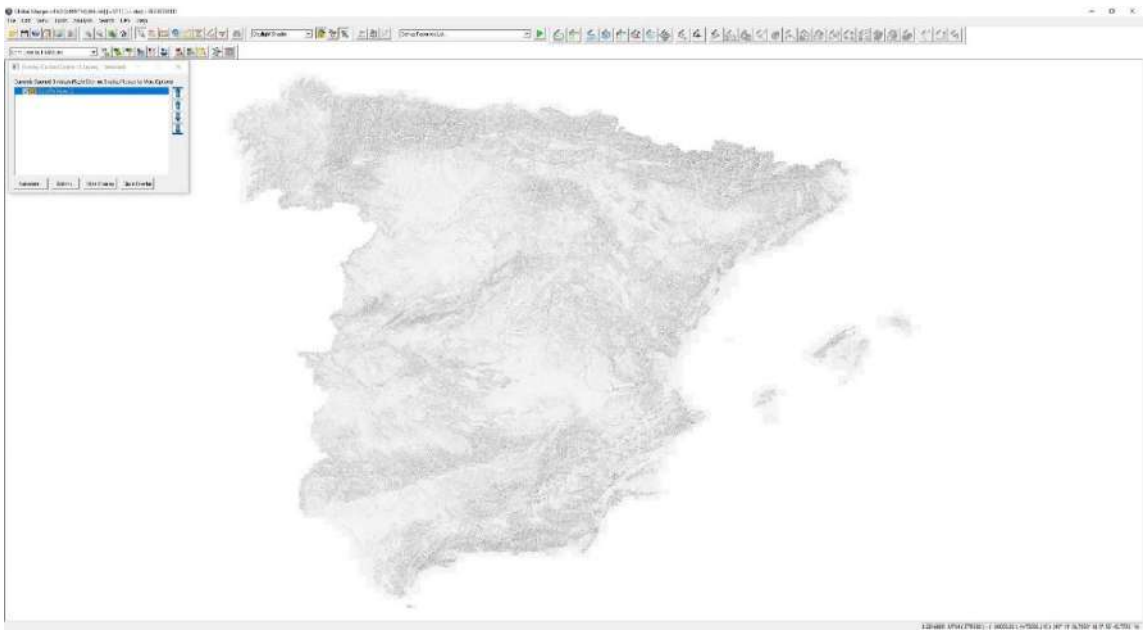


Cuando ya tenemos cargados todos los raster necesarios cerramos la ventana. En este momento Global Mapper en su espacio de trabajo nos muestra el siguiente aspecto, donde se puede apreciar de manera visual (pero esquemática) el conjunto de los archivos añadidos al catálogo.



Gracias a los archivos .tfw y .prj que acompañan a cada uno de los archivos de sombreado, se consigue que todos los elementos se coloquen en sus coordenadas reales, dando lugar al mosaico de imágenes georreferenciadas.

La configuración de esta visualización se puede cambiar a través de opciones, y hay que decir que depende del zoom aplicado, la vista será esquemática o real. A continuación, se muestra el aspecto con una configuración más “amigable” del conjunto del catálogo/mosaico.



Como se puede apreciar, no hay saltos de apariencia entre unas provincias y otras, gracias a la utilización de los mismos parámetros a la hora de generar los sombreados. Para realizar el mosaico de las hojas correspondientes al huso 28, la forma de trabajo sería la misma: se crearía un catálogo y se añadirían las provincias correspondientes.

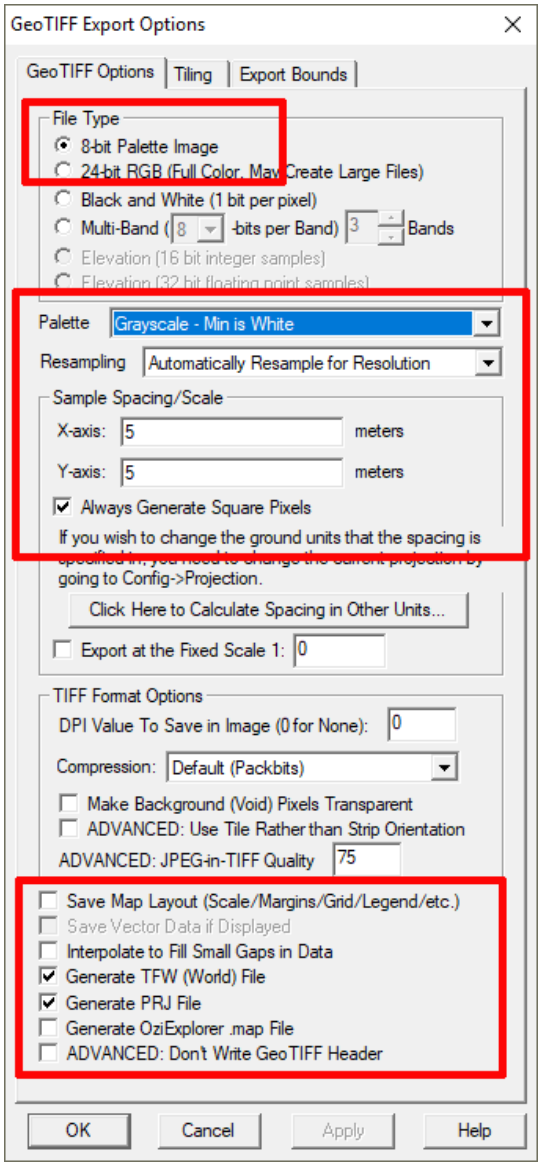
Exportación del sombreado por hojas

Exportación por hojas a escala 1:25,000

Una vez generado el mosaico del sombreado, lo siguiente que deberíamos realizar es la explotación del mismo extrayendo de aquí los sombreados individuales correspondientes a cada una de las hojas de la escala 1:25,000 del MTN25. Podemos proceder de dos maneras distintas, dependiendo del interés:

1. **Exportar desde el catálogo directamente los sombreados de las hojas del MTN25** sabiendo que todos van a estar en huso 30 (huso 19 en el caso de República Dominicana) mediante una “malla” en formato .shp que incluye el ámbito geográfico (coordenadas) de cada una. Posteriormente, si resulta necesario, esos sombreados de hojas se pueden reproyectar individualmente a otros husos horarios mediante Global Mapper.
2. **Exportar el catálogo completo a un nuevo archivo raster Geotiff** (completamente unido) del conjunto de sombreados de las hojas de la Península en huso 30.

Los parámetros de exportación a Geotiff del catálogo serían los siguientes:



El archivo exportado tendrá un “peso” considerable, por lo que hay que valorar si nos interesa esta opción.

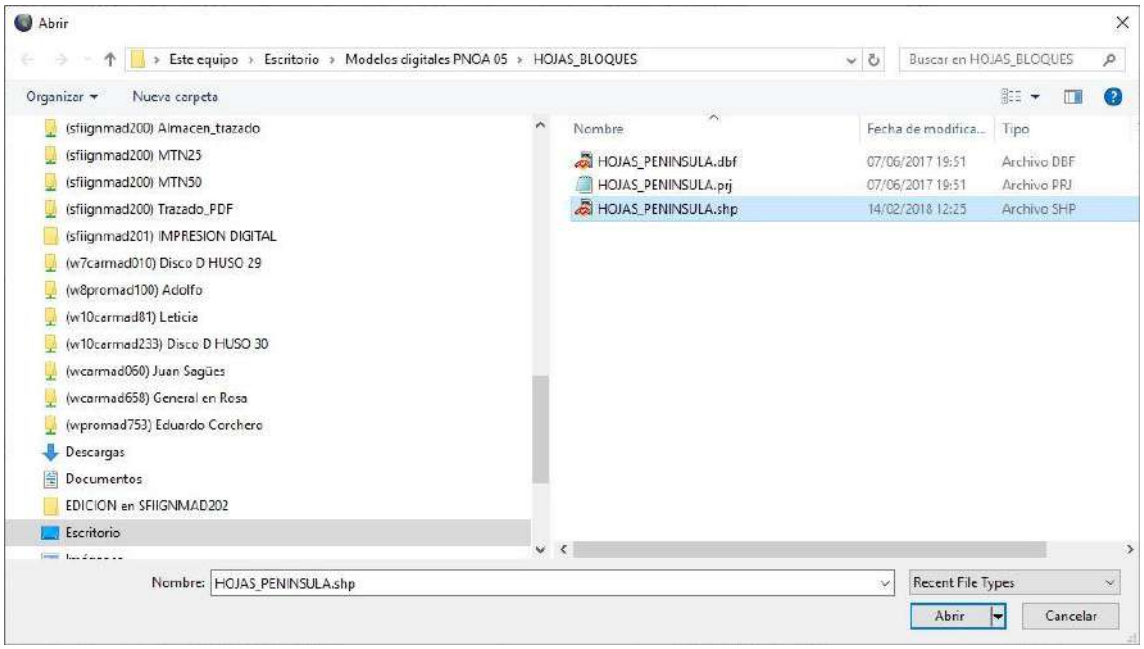
El archivo resultante podría ser aprovechado para ser reproyectado posteriormente en otros dos: uno en huso 29 y otro en huso 31. Posteriormente, de cada uno de ellos se podrían extraer únicamente los sombreados de las hojas correspondientes a cada huso mediante “mallas” en formato .shp que contengan los ámbitos de las que queramos exportar.

Finalmente se optó por la segunda opción, generando tres grandes archivos de sombreado, uno de cada huso (29, 30, y 31), y luego, aparte, el del huso 28 correspondiente a las Canarias.

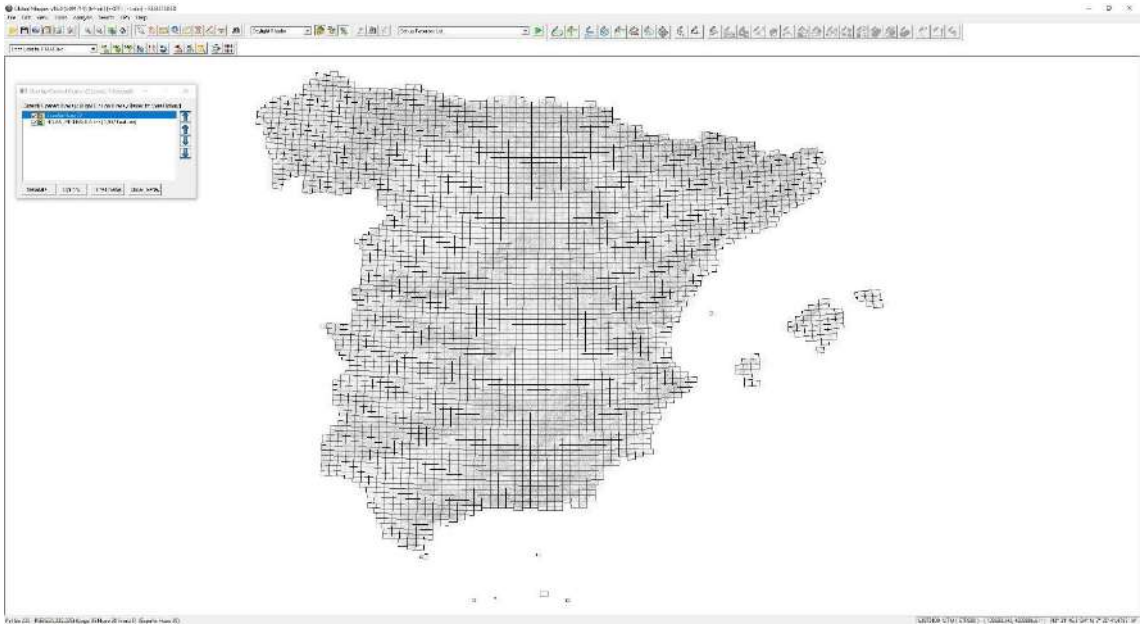
El procedimiento de exportación es el siguiente para ambos casos y se explica a continuación.

Teniendo abierto el catálogo de sombreados, o en su defecto el raster de todas las provincias obtenido de la segunda opción explicada anteriormente, deberemos incorporar a Global Mapper una malla (generalmente en formato .shp) con la totalidad de las hojas que pretendamos extraer.

Esto lo haremos cargándola desde el **Menú File→Open Data Files**

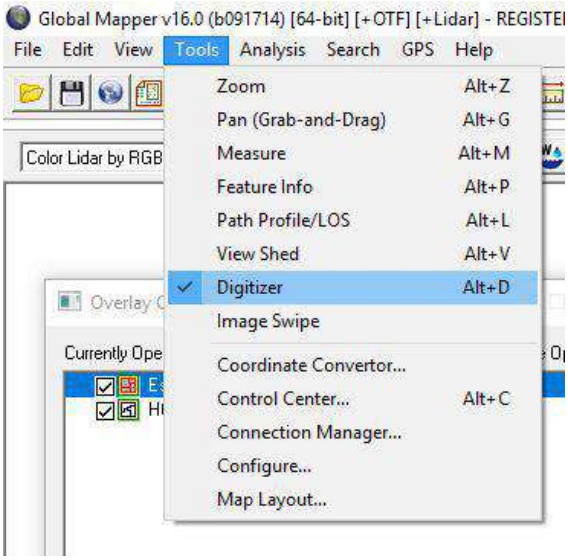


Esto nos da como resultado algo similar a la siguiente captura. Se puede apreciar que en la ventana **Overlay Control Centre** existen dos "capas": una es el catálogo de sombreados (o el raster único del sombreado en huso 30) y otra es la correspondiente a la malla que acabamos de cargar.

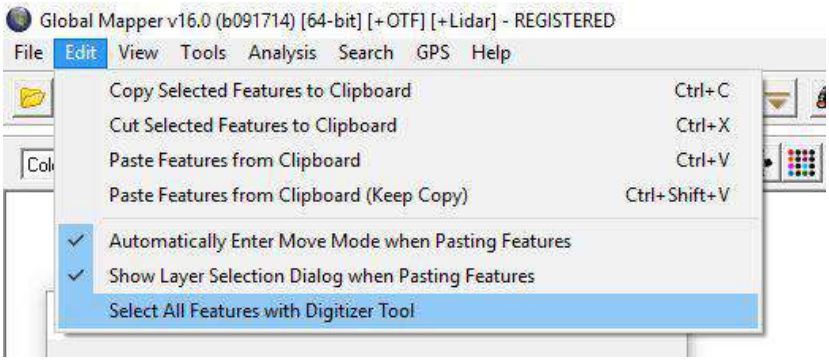


Una vez cargada la malla con los ámbitos de las hojas, procederemos a exportar los sombreados individualmente de la siguiente manera:

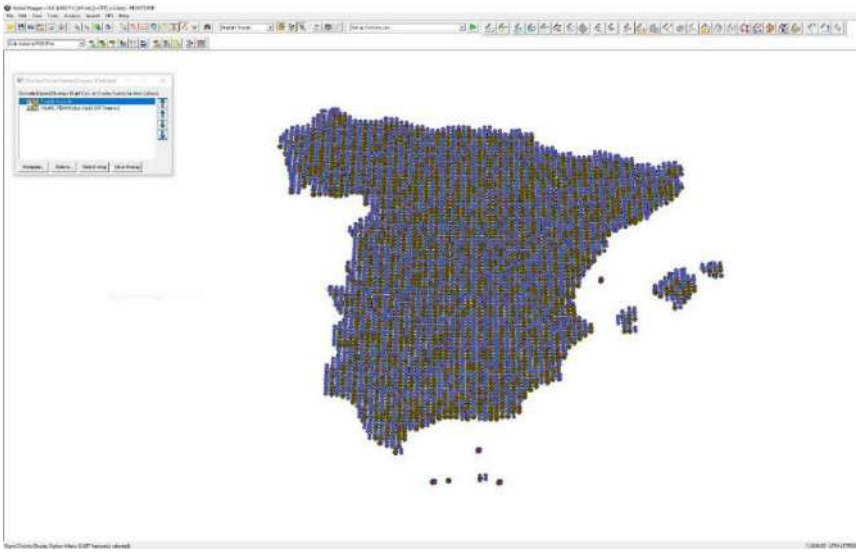
- Nos dirigiremos al **Menú Tools→Digitizer**



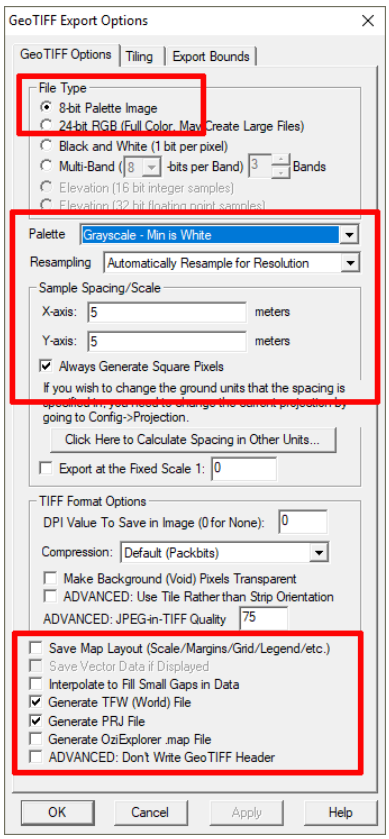
- Aceptamos el diálogo que aparece y a continuación vamos al **Menú Edit→Select All Features with Digitizer Tool**.

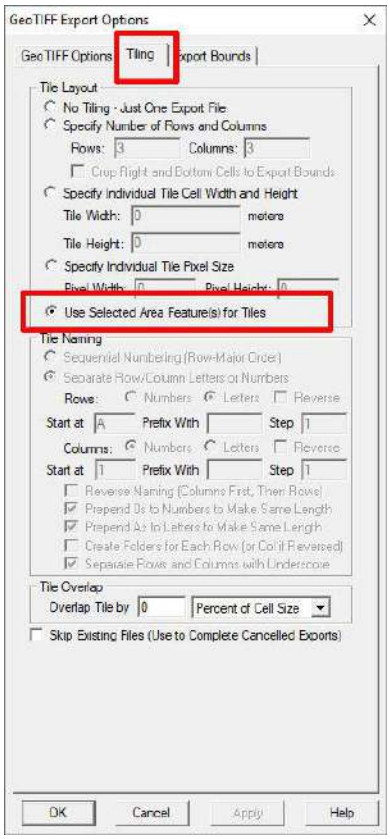


- Esto seleccionará todos los elementos de la malla de la siguiente manera:



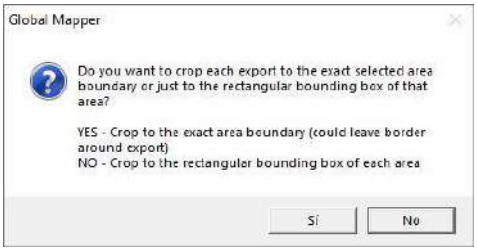
- Posteriormente accederemos de nuevo al **Menú Export→Raster/Image Format→Geotiff** utilizando los valores ya conocidos, pero en esta ocasión la salvedad es que utilizaremos también la pestaña **Tiling**, donde activaremos el siguiente parámetro, que lo que hace es indicar que vamos a utilizar las áreas seleccionadas para exportar y que coinciden con las hojas del MTN25.



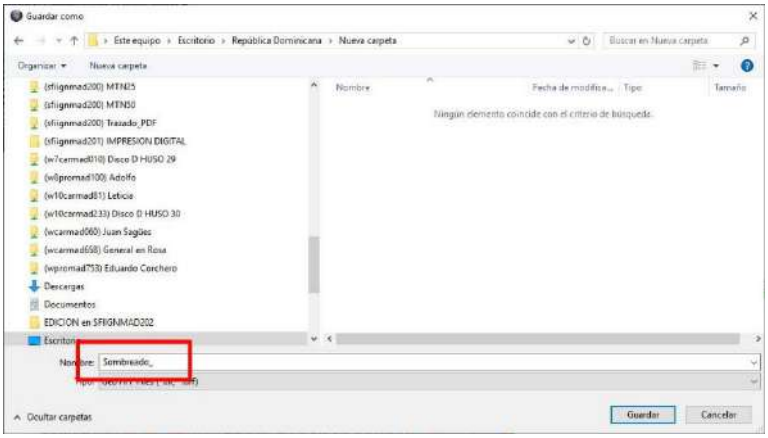


- En el momento en que pulsemos Ok, Global Mapper nos preguntará que atributo queremos utilizar para nombrar cada uno de los raster de sombreado que vamos a exportar. Evidentemente, en el archivo .shp que contiene debe existir algún atributo que nos facilite esta labor. En el caso del IGN DE ESPAÑA, entre otros atributos, se tiene uno denominado NUMERO y que contiene el número de clasificación de cada hoja del MTN25. En su defecto, se puede elegir también un número secuencial.

- Cuando confirmemos, Global Mapper nos pregunta si queremos cortar cada área ajustándose a sus dimensiones con el riesgo de que deje bordes de color alrededor de la exportación, o si por el contrario, queremos cortar por el bounding box que la contiene (esto tiene que ver con el giro de la convergencia de las hojas).



- En nuestro caso la respuesta es **NO**, puesto que queremos que se incluya la hoja + los solapes de las hojas adyacentes y de este modo tendremos continuidad entre las "tiles" o piezas que exportemos.
- Una vez indicado que queremos que el programa corte cada pieza por el área del rectángulo que las contiene, Global Mapper nos solicita el directorio donde las vamos a guardar, y un nombre para cada una de ellas. El nombre será una combinación entre lo que pongamos en la ventana más el campo que hemos elegido anteriormente. En el ejemplo ponemos Sombreado_ y Global Mapper lo completa con el número de Hoja tomado del atributo Número.



- El proceso dará comienzo y no terminará hasta que no se exporta el último sombreado. Esto puede llegar a tardar mucho tiempo dependiendo de la resolución, el modo de color, el número de hojas a exportar, etc. (si hay remuestreos, el tiempo de procesamiento también aumenta).

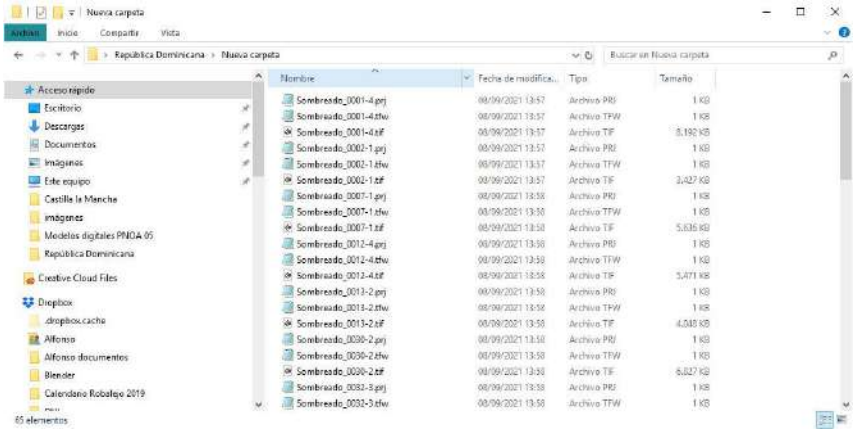


Ilustración 6 - 44: Archivos de sombreado resultantes por NUMERO de hoja de MTN25

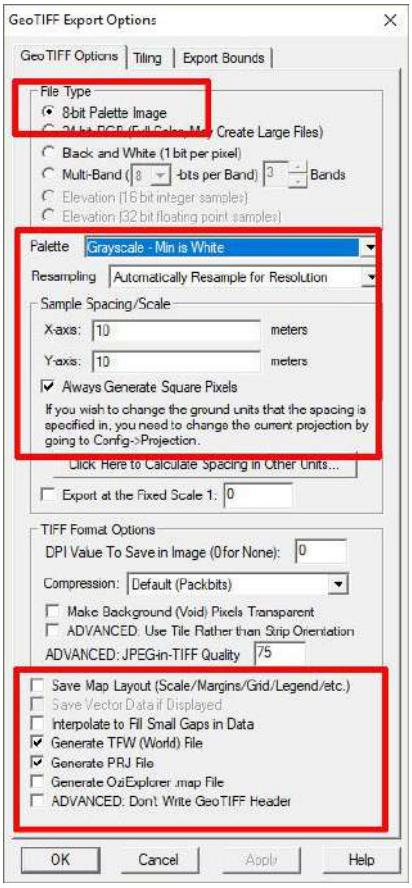
- Al igual que hemos realizado la exportación de los sombreados en huso 30, habrá que realizar el mismo procedimiento para las hojas del huso 28.

Exportación por hojas a escala 1:50,000

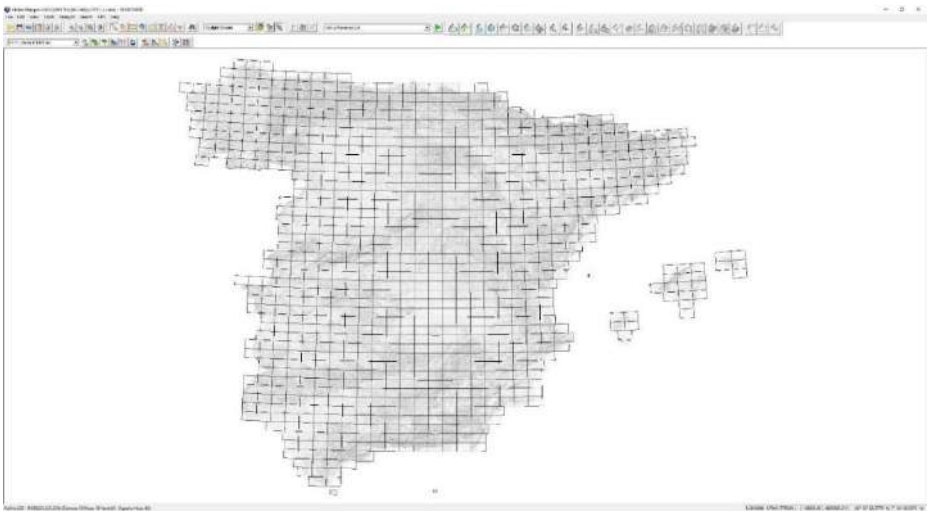
Utilizando como base el mosaico de sombreados de provincias o el Geotiff del sombreado a escala 1:25,000, es posible obtener también el sombreado por hojas con una resolución acorde a la escala 1:50,000.

El procedimiento básico sería el mismo que se ha utilizado, pero cambiando los parámetros de exportación como hojas sueltas en base a lo siguiente:

- La resolución de salida ha de variar de 5m/pixel a 10m/pixel



- La malla a utilizar deberá corresponder a la distribución de hojas a escala 1:50,000 (que en total son 1,073)



6.3. Procesos de simbolización e integración de sombreado

En esta sección se describen dos metodologías que se pueden llevar a cabo para el proceso de simbolización e integración del sombreado: de forma manual, que es la que se aplica en el IGN de España cuando se trata de publicación de cartografía impresa, y de forma automática, en los otros casos.

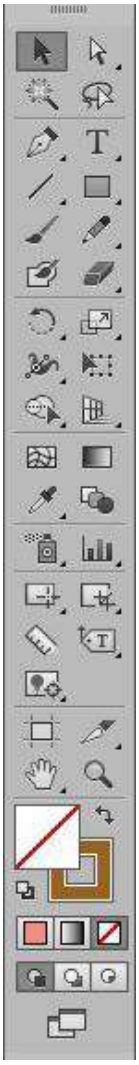
6.3.1. Proceso manual de simbolización para publicación impresa

En base a la metodología de trabajo implementada en el IGN de España, este proceso se realiza con el software Adobe Illustrator y con Mapublisher. Para facilitar la comprensión del proceso, a continuación se muestra la descripción de las principales funcionalidades de dichas herramientas que es necesario conocer a la hora de desarrollar dicho proceso.

Utilización básica de Adobe Illustrator

Creación y manipulación de objetos

- Panel de Herramientas:



El panel Herramientas aparece a la izquierda de la pantalla la primera vez que inicia la aplicación. Las Herramientas se utilizan para crear, seleccionar y manipular objetos en Illustrator.

Algunas herramientas disponen de opciones que aparecen al hacer doble clic sobre de ellas. Un pequeño triángulo en la esquina inferior derecha del icono de herramienta indica la presencia de herramientas ocultas que se muestran al hacer clic en él.

- Herramientas de Selección

Las herramientas de selección y selección directa nos servirán para seleccionar los objetos o su contenido (en el caso de la herramienta de selección directa).

Con la herramienta de selección (cursor negro) se pueden seleccionar objetos pinchando sobre ellos. Si a la vez se pulsa la tecla de mayúsculas se añadirán o restarán objetos a la selección, dependiendo de si están o no seleccionados previamente. Y arrastrando el cursor mediante una ventana se seleccionarán los objetos que toquen a esa ventana. Para mover o transformar los elementos enteros utilizaremos esta herramienta.

La herramienta de selección directa se puede utilizar para modificar nodos o puntos de ancla dentro de un trazado.

Producción y dibujo de objetos

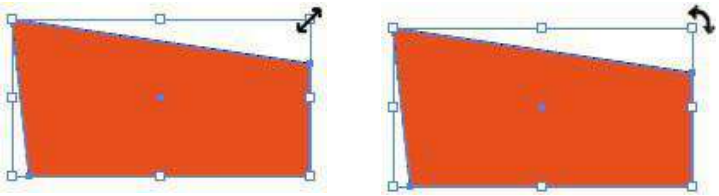
Podemos elegir entre varias herramientas para la producción de objetos, aunque posteriormente los podremos manipular y editar para obtener los resultados deseados.

En la imagen se ven las siguientes: herramienta Pluma para dibujar trazados, herramienta de Texto para editar textos, herramienta de Líneas para trazar líneas rectas, herramienta Rectángulos y Elipses para dibujar rectángulos y elipses (ésta última oculta), Herramienta Pincel, Lápiz y Pincel de Manchas para dibujar/pintar y herramienta Borrador.

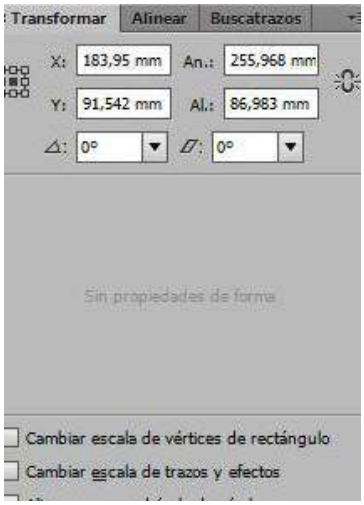
Manipulación y transformación de objetos

- Escala y Giro de objetos

Si seleccionamos un objeto y pinchamos con el cursor negro sobre cualquiera de los manejadores que le rodean, podremos deformarlo en anchura y altura. Para transformar su tamaño proporcionalmente deberemos tener pulsada la tecla Mayúsculas. Si nos alejamos un poco sobre los manejadores de la selección, la herramienta nos permitirá girar el objeto. Si pulsamos la tecla de mayúsculas a la vez que giramos, los ángulos de giro se restringen a 45 y 90 grados.

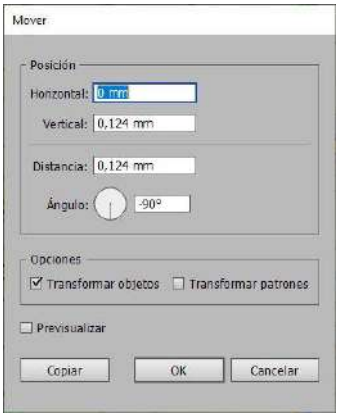


Para realizar estas transformaciones también se puede utilizar la herramienta Transformar (**Menú Ventana-->Transformar**), desde donde además podemos indicar desde qué punto del objeto se producirá la transformación.



- Movimiento de objetos

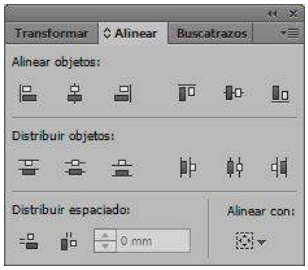
Para mover objetos podemos arrastrarlos con la herramienta Selección; también podemos utilizar los cursores del teclado o introducir valores concretos en el panel o un cuadro de diálogo **Menú Objeto-->Transformar/Mover**



Además, se puede usar la tecla Mayúsculas para restringir el movimiento de los objetos en la dirección horizontal, vertical o diagonal exacta relativa a la orientación actual de los ejes x e y. La tecla Mayúsculas sirve también para girar los objetos en incrementos de 45°.

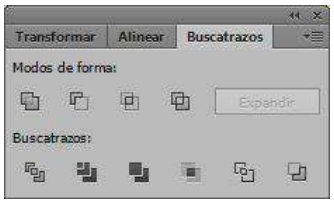
Alineación y distribución de objetos

Para alinear o distribuir los objetos en relación con los demás objetos seleccionados o con la mesa de trabajo se utiliza el panel Alinear (**Menú Ventana-->Alinear**).



Creación de objetos complejos (operaciones Boolenas)

Se pueden obtener objetos complejos desde objetos simples con la herramienta Buscatrazos mediante operaciones Booleanas (unificar, menos frente, formar intersección, excluir, dividir, cortar, combinar, recortar, contornear, menos fondo). Mediante los modos de forma estos objetos simples mantienen la información de su trazado hasta que pulsamos Expandir.



Uso de Cuadrículas y Guías

- Uso de la cuadrícula

La cuadrícula aparece detrás de la Ilustración y nos sirve de referencia para la colocación o creación de objetos; es un elemento de ayuda visual y no se imprime.

Para visualizar la cuadrícula, deberemos ir al **Menú Ver--> Mostrar cuadrícula**. Para ocultar la cuadrícula, deberemos ir al **Menú Ver--> Ocultar cuadrícula**.

Para ajustar objetos a las líneas de cuadrícula, deberemos ir al **Menú--> Ver--> Ajustar a cuadrícula**. De esta forma al mover o crear los objetos se hará desde puntos exactos de dicha cuadrícula.

- Uso de las guías

Las guías son unas líneas que no se imprimen y que permiten alinear texto y objetos gráficos. Podemos crearlas arrastrando desde las reglas de página.

Para mostrar u ocultar las guías, iremos al **Menú Ver--> Guías--> Mostrar guías o bien Ocultar guías**.

Para bloquear las guías, deberemos acceder al **Menú Ver--> Guías--> Bloquear guías**. De esta manera no se moverán accidentalmente facilitando el trabajo.

- Guías inteligentes

Las guías inteligentes facilitan la alineación o transformación de elementos u objetos con relación a otros. Para activar las guías inteligentes deberemos acceder al **Menú Ver--> Guías inteligentes**.

Manejo del Color

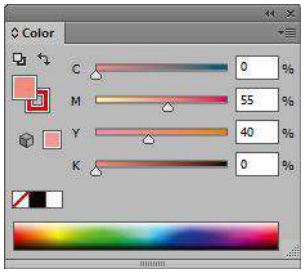
Al crear un nuevo documento, Illustrator nos pregunta el modo de color que utilizaremos para trabajar. En el caso de un documento dirigido a la impresión el modo será CMYK, y si está enfocado para web u otros usos, RGB.

- Color de Línea y Relleno

En la parte inferior de la barra de herramientas tenemos los iconos que nos muestran el color de relleno y de contorno de los objetos que vamos a crear o del objeto que tenemos seleccionado, para acceder a las modificaciones de color haremos doble clic sobre el icono y aparecerá la paleta de color para que elijamos un nuevo color.

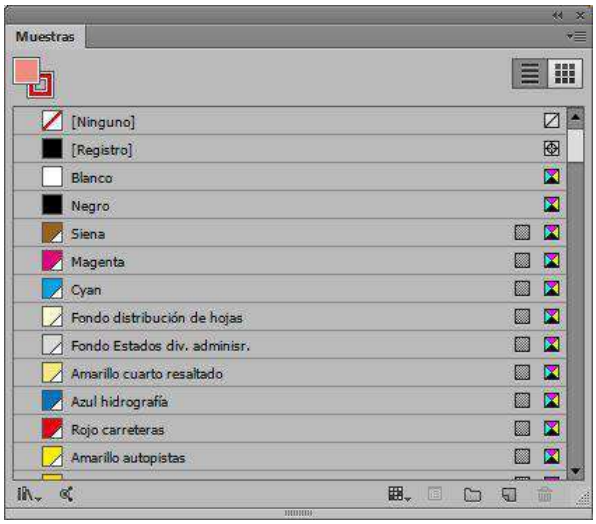
- Ventana de Paleta de color

Podemos indicar los valores de color numéricamente, gráficamente con los deslizantes o directamente pinchando sobre la muestra inferior, donde encontramos además la opción “sin color” y el blanco y negro puros. A esta paleta también se puede acceder desde el **Menú Ventana--> Color**.



- Ventana de Muestras de color

A esta ventana se accede desde el **Menú Ventana--> Muestras**, y es el lugar donde podemos ir agregando los colores que vayamos creando en el proyecto para poder reutilizarlos siempre que queramos. En el proyecto de un mapa del IGN DE ESPAÑA se utilizan unas 100 muestras entre muestras de color y patrones de cultivos (que también se guardan en la misma paleta).



Al generar un nuevo documento de Illustrator, en la paleta de muestras aparece una serie de muestras de color predeterminadas que podemos modificar. También podemos añadir colores para trabajar más cómodamente o, por el contrario, podemos eliminar los que no vayamos a utilizar. En todo caso es muy importante tener las muestras controladas para evitar problemas.

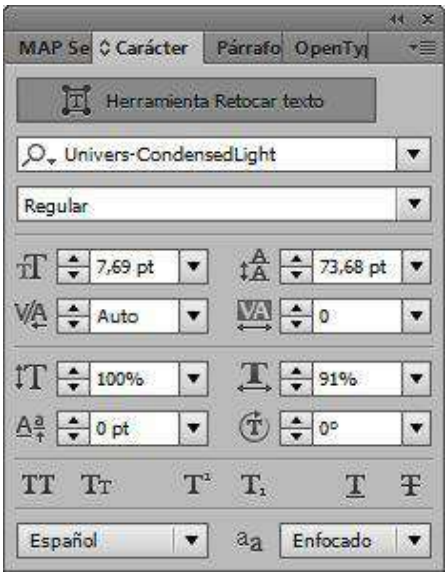
Edición de Líneas

Para la edición de líneas o trazos que previamente hayamos dibujado con la herramienta línea o pluma, utilizaremos la paleta que existe al efecto y a la cual accederemos desde el **Menú Ventana--> Trazos**.

En esta paleta, y teniendo una línea (o varias) seleccionada podremos definir una serie de aspectos “estéticos” de las mismas: el grosor de las líneas, la forma de sus remates y vértices, su discontinuidad indicando valores de guión o hueco, así como añadir puntas de flecha si lo deseamos.

Edición de Textos

Para la edición de textos contamos con una paleta con la que se pueden modificar todos los aspectos relativos a la tipografía, a la que accederemos a desde el **Menú Ventana--> Texto--> Carácter**.



Desde aquí podemos cambiar la fuente, su estilo, su tamaño o cuerpo, su interlineado, su interletreado, su escala horizontal o vertical, etc. También podemos indicar si va todo en mayúsculas, versalitas, en superíndice o subíndice, tachada, subrayada y, por supuesto, el diccionario (en caso de usarlo para particiones de palabras).

Estilos Gráficos y de Carácter

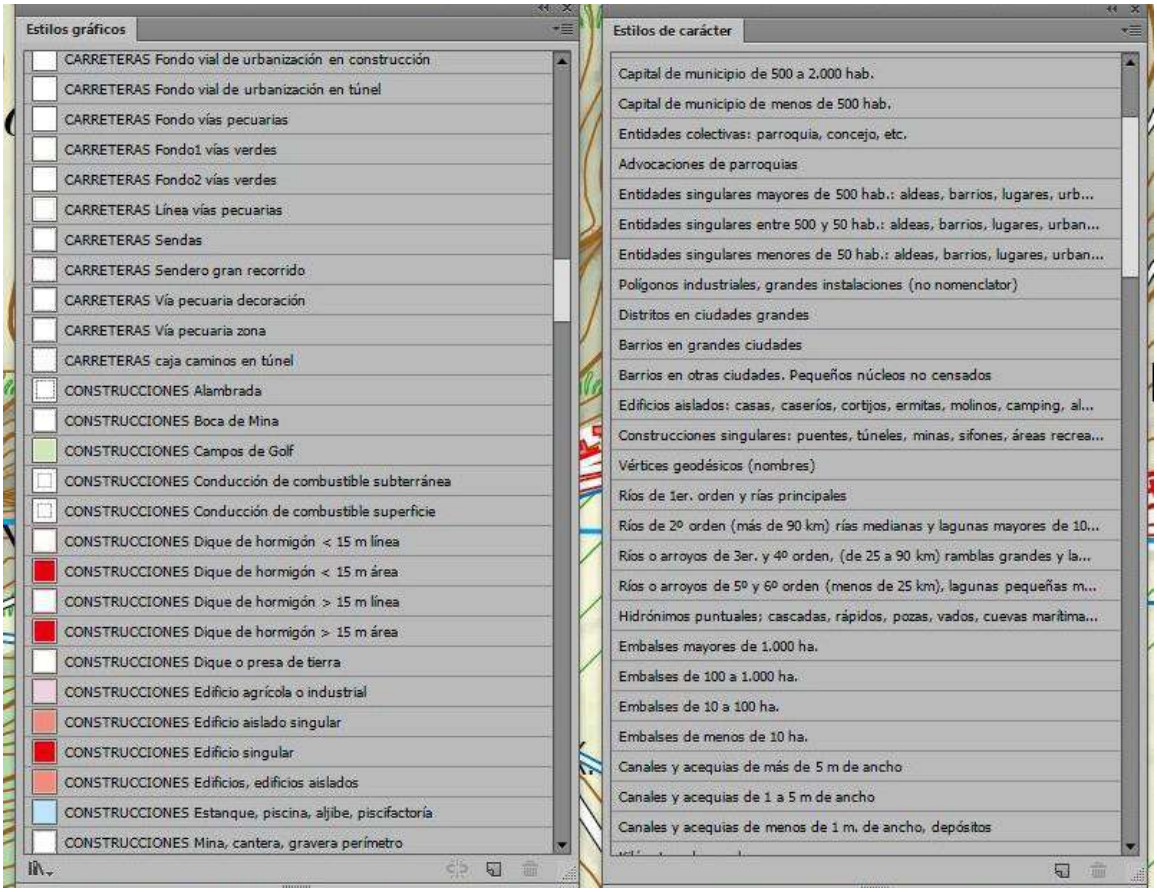
Los estilos son atributos de apariencia (color de línea, color de relleno, patrón de relleno, patrón de línea, grosor de línea, etc.), recogidos en un panel desde el cual podemos reutilizarlos cada vez que los necesitemos.

- Estilos Gráficos

Son los más habituales y se puede acceder a ellos desde el **Menú Ventana--> Estilos gráficos**. Un ejemplo de estilo en la cartografía del MTN sería el de las carreteras nacionales, donde llevan la misma apariencia: un determinado grosor de línea y un determinado color de línea.

- Estilos de Carácter

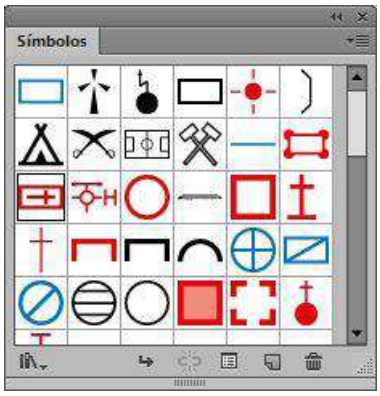
Al igual que existen estilos gráficos en los que se guarda la apariencia de los objetos, también existen los estilos de carácter que guardan la apariencia de los textos. En el caso de un mapa se utilizan sobre todo para homogeneizar la toponimia, categorizándola visualmente a través de su apariencia. Se accede a ellos a través del **Menú Texto--> Estilos de Carácter**.



Símbolos

Los símbolos son elementos gráficos guardados en el panel Símbolos. Podemos replicar tantas veces como queramos un símbolo en un documento; de este modo se crean instancias. Para acceder a ellos deberemos ir al **Menú Ventana--> Símbolos**.

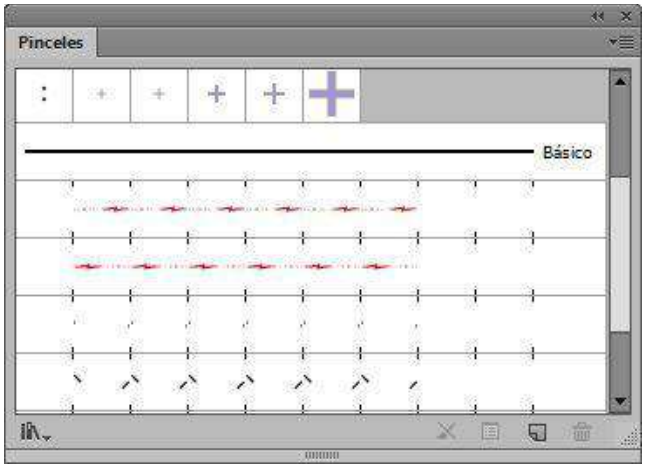
Un ejemplo de su uso en cartografía son los signos puntuales de un mapa.



Pinceles

Los pinceles son trazos u elementos que se aplican a los contornos de los trazados (cerrados o abiertos) o sobre las líneas del documento. Están clasificados en cuatro categorías: caligráficos, de dispersión, de arte y de motivo. Podemos acceder a ellos a través del **Menú Ventana--> Pinceles**.

Un ejemplo de su uso en cartografía son las líneas con patrones repetitivos.



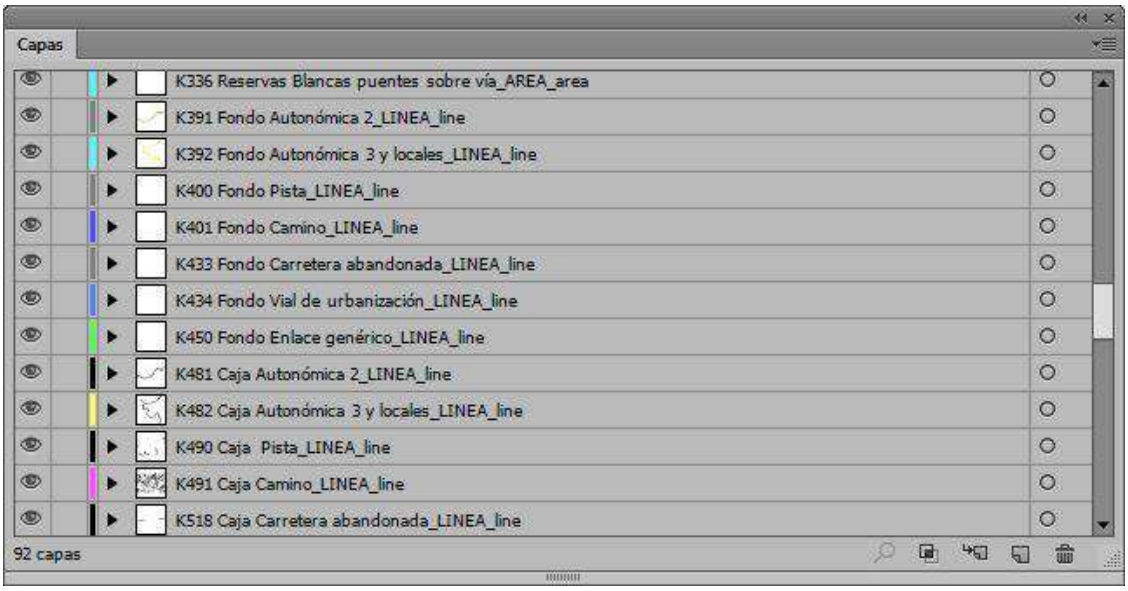
Capas

Las capas son unos elementos de la Ilustración que actúan como carpetas transparentes con contenido. Pueden contener cualquier tipo de ilustración.

Las capas se pueden organizar para que todo el contenido de una de ellas se sitúe por encima o por debajo del contenido de otra. Cuando creamos un nuevo documento en Illustrator, generalmente tendrá una sola capa. Podemos crearlas, eliminarlas y ordenarlas para que la información del documento quede ordenada visualmente.

En cartografía su uso es prácticamente obligatorio y está muy extendido. Un documento Illustrator de una hoja del MTN25 puede llegar a tener cientos de ellas.

Se puede acceder a ellas a través del **Menú Ventana--> Capas**.



Uso de Plantillas

Illustrator permite crear documentos a partir de plantillas suministradas con el programa o que nosotros mismos hayamos elaborado previamente. Esto permite, por una parte, ganar tiempo

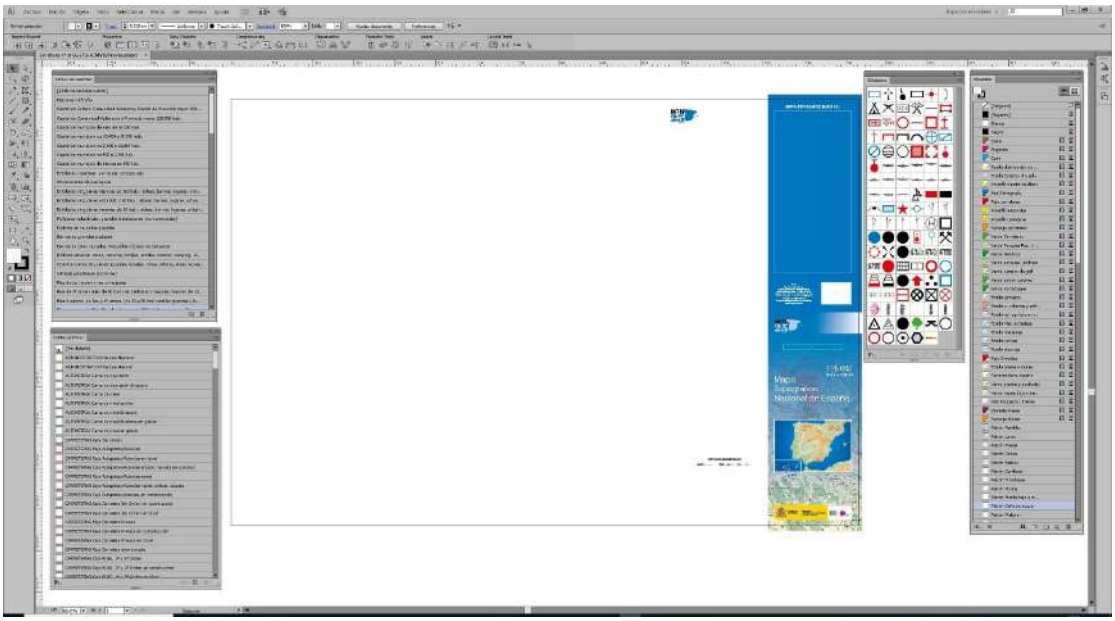
y, por otra, unificar documentos corporativos, ya que se utilizarán siempre los mismos estilos, las mismas muestras, las mismas fuentes, etc...

En la producción del IGN de España esto toma una gran dimensión pues es lo que permite homogeneizar el aspecto y apariencia de las todas las hojas de la serie que se publican. Evidentemente, cada serie cartográfica está dotada de su correspondiente plantilla.

En principio, y según la abrimos, el aspecto de la plantilla no nos lleva a imaginar toda la información que contiene, pero sin lugar a dudas es el documento que recoge toda la casuística y las posibilidades de simbolización cartográfica de una determinada serie cartográfica.

El uso intensivo de la plantilla para producción cartográfica se ha visto facilitada en gran manera por la utilización del software MaPublisher, que hace que Illustrator se vea implementado con funcionalidades SIG como son el uso de elementos gráficos georreferenciados, la posibilidad de cambiar de proyección y escala en el momento, la utilización de los atributos para poder simbolizar elementos puntuales, de tipo lineal y de tipo superficial así como textos, la importación y exportación en formatos estándar usados por los SIG, la posibilidad de publicar servicios de mapas vía web, la importación y exportación de imágenes ráster georreferenciadas y, por supuesto, la creación de archivos GEOPDF, en los cuales se pueden visualizar capas temáticas, realizar operaciones de medición, consultar los atributos de un elemento, etc.

En la siguiente imagen se muestra el aspecto de la Plantilla de las hojas del MTN25 antes de recibir los datos de un mapa a través de Mapublisher.



Utilización básica de Mapublisher junto con Adobe Illustrator

El software de cartografía MAPublisher se instala sobre Adobe Illustrator como un Plug-in que al que dota con más de cincuenta herramientas de cartografía SIG que facilitan la elaboración y publicación de mapas.

MAPublisher ofrece paneles, funciones y herramientas complementarios al entorno de diseño de Adobe Illustrator que permiten que los archivos de datos SIG se utilicen como base para producción cartográfica destinada a contenidos impresos o servicios web.

Trabajando desde un documento tradicional o una plantilla de Adobe Illustrator, los usuarios pueden crear o editar sus datos organizados a través de capas de mapa o Map Layers.

Importación de datos hacia la plantilla

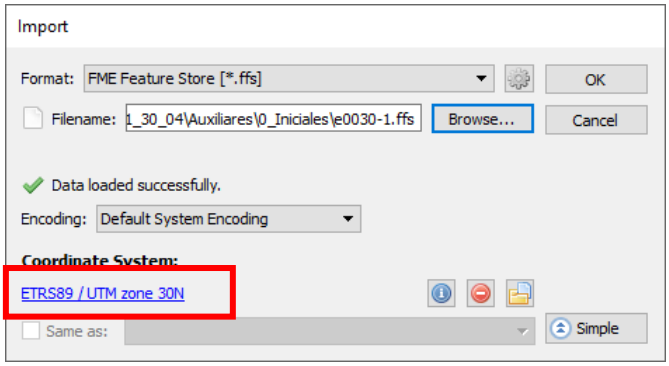
Las funciones de importación simple y de importación de datos múltiples de MAPublisher permiten la importación de los principales formatos SIG y CAD (ArcGIS, MapInfo, AutoCAD, MicroStation, Kml, etc.) en Adobe Illustrator, conservando su geometría nativa (Punto, Línea, Texto, Área), estructura de capas, ubicación geográfica y datos de atributos.

En la siguiente imagen, se destacan los iconos de las herramientas de importación simple e importación de datos múltiples de Mapublisher (recuadradas en rojo) dentro del apartado Import/Export.

Sistema de Coordenadas

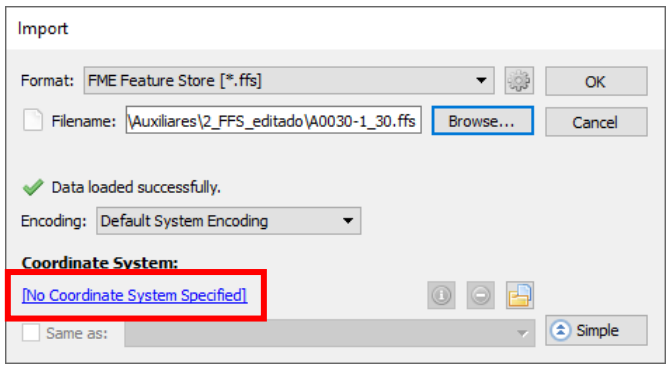
Una vez importados los datos, se presenta una ventana en la que hemos de indicar si el sistema de coordenadas es correcto. Si no es así, o si los datos no lo llevan implícito se podría cambiar haciendo clic en el apartado Coordinate System.

En esta imagen se muestra una ventana de importación que **lleva sistema de coordenadas integrado**.

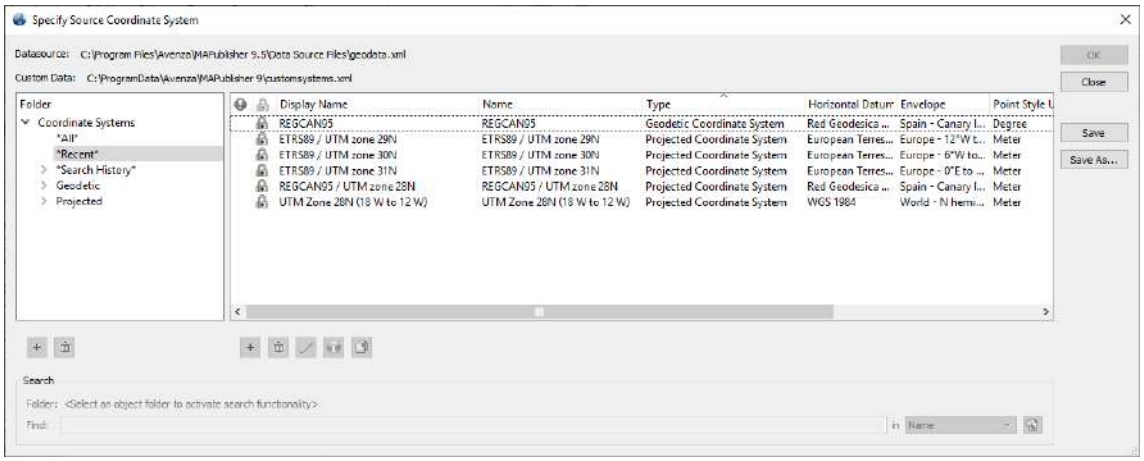


Si el Sistema de coordenadas que aparece es correcto, podremos continuar, pero si no lo es deberemos cambiarlo para asignar el que corresponda haciendo clic en el texto azul.

En esta otra imagen se muestra una ventana de importación que **no lleva sistema de coordenadas integrado**.



En este caso para asignar el sistema de coordenadas que corresponda haremos clic en el texto azul y se nos presentará una ventana desde la cual podremos elegir el correcto, o, en su defecto, añadir alguno que tengamos especificado como propio.



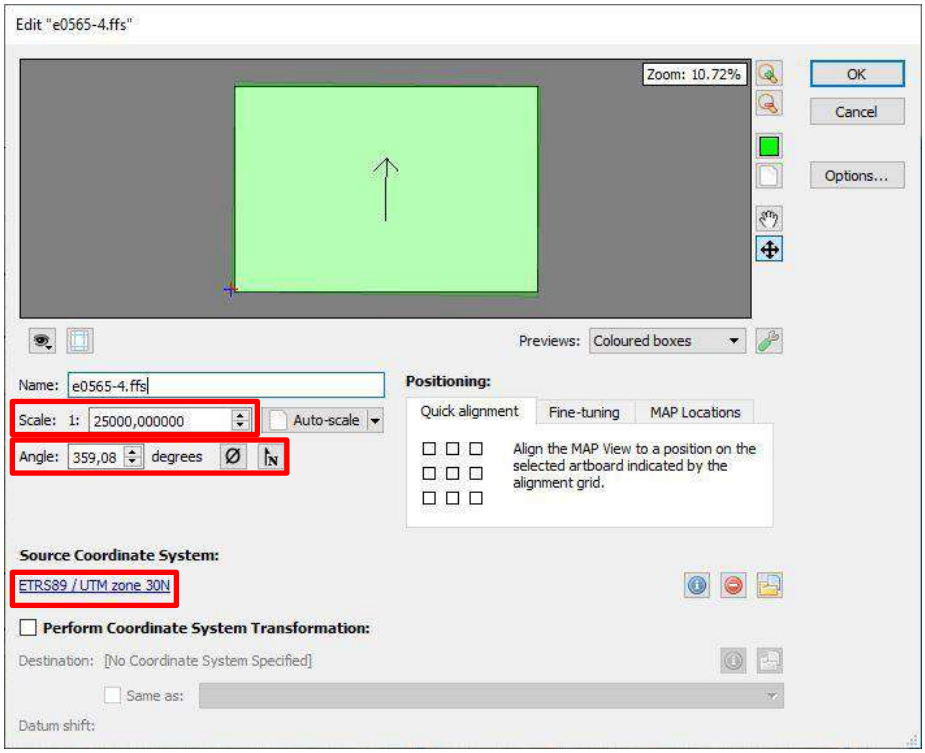
Map Views

El espacio de trabajo cartográfico se gestiona a través de MAP Views (vistas de mapa) que controlan los parámetros geoespaciales: sistema de coordenadas, escala, rotación y anclaje del mapa. Las capas organizadas en la misma MAP View comparten el mismo marco geográfico de referencia.

En la siguiente imagen, se puede apreciar el icono de la Map View (recuadrado en rojo) dentro del apartado Properties

Si hacemos clic en el icono de Map Views, se abre la ventana donde podremos asignar o cambiar parámetros.

La siguiente imagen corresponde al detalle de la Map View de una hoja de cartografía del IGN de España donde se pueden apreciar la escala, el ángulo de convergencia y el sistema de coordenadas (recuadrados en rojo).



Las capas con elementos georreferenciados y atributos (MAP Layers) deben residir dentro de la MAP View para ser administradas por las funciones de MAPublisher y conservar así su información espacial y sus atributos; sin embargo, las capas que no están relacionadas con una MAP View se pueden utilizar para fines de presentación (como título, escala y leyenda) puesto que carecen de georreferenciación y de atributos.

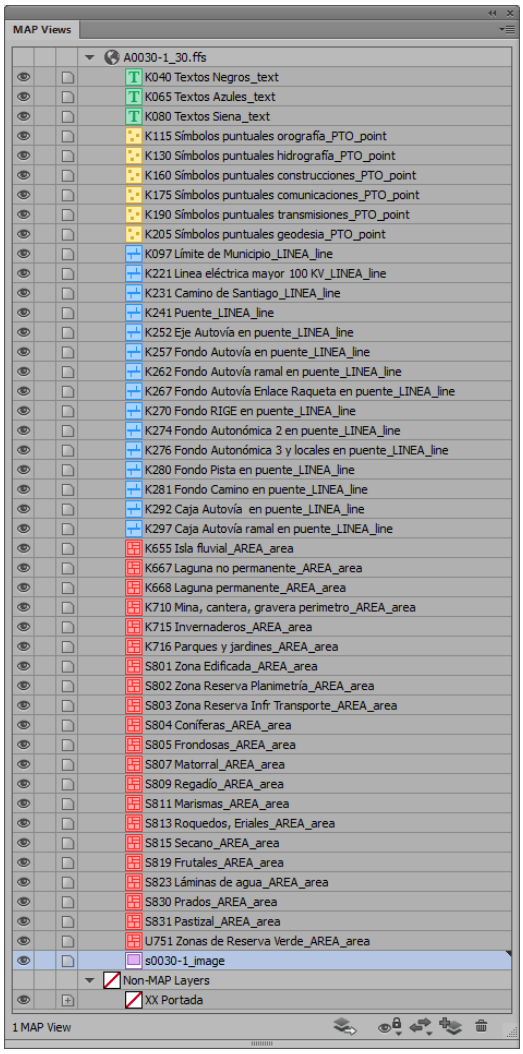
Hay que decir que las capas “de mapa” que gestiona MaPublisher y las que gestiona Illustrator de manera nativa, reciben tratamientos diferentes, aunque en sí coinciden en número y aparenten ser las mismas.

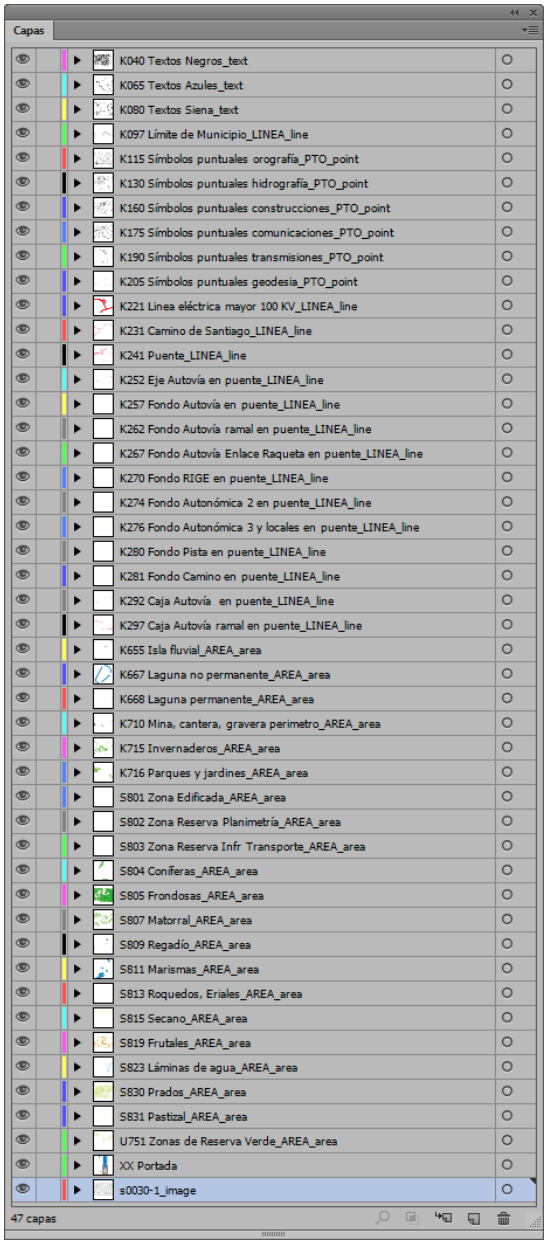
Un documento de Adobe Illustrator puede contener más de una MAP View. Por ejemplo, un mapa con un mapa insertado a mayor escala para representar un área de interés ampliada.

En la imagen de la izquierda, se muestra la Ventana MapView de una hoja de cartografía del IGN de España, donde se pueden apreciar los distintos tipos de capas “de mapa”:

- En verde: capas de texto
- En amarillo: capas de punto
- En azul: capas de línea
- En rojo: capas de área
- En morado: capas de imagen (en este caso del sombreado de la hoja)
- También existe una capa no georreferenciada correspondiente a la portada del mapa

Por el contrario, en la imagen de la derecha se muestra la ventana de capas de Illustrator, donde no se puede distinguir el tipo de geometrías que están alojadas en cada capa, pero donde sí que se puede apreciar que todas están ordenadas alfabéticamente (menos la que corresponde a la imagen del sombreado que está la última).



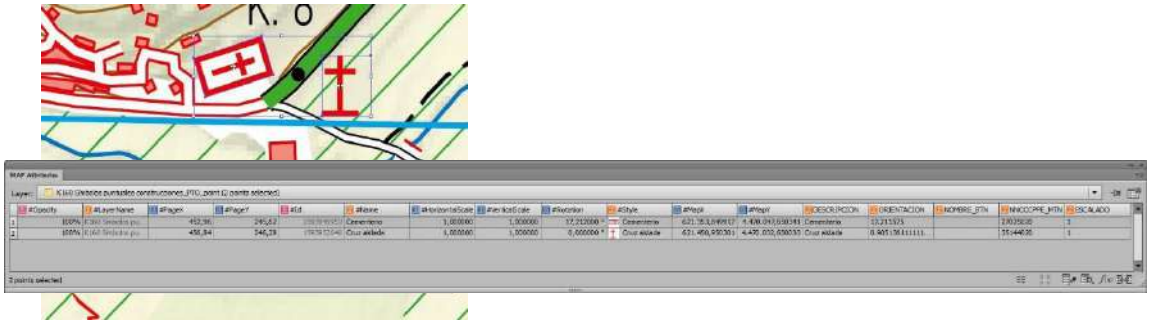


Map Attributes

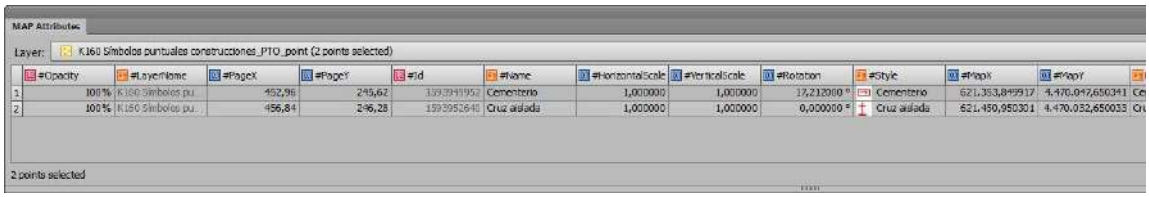
El panel de MAP Attributes (atributos de mapa) permite crear, editar y consultar información de los atributos de los elementos geográficos representados. Por ejemplo, para una capa de línea que contiene datos de la calle, cada segmento de línea se puede vincular a una tabla con información relativa al nombre y categoría de calle. Usando el panel de MAP Attributes, esta información se puede ver, consultar, editar o agregar, además se pueden agregar columnas de datos a través de su interfaz.

En la siguiente imagen se puede apreciar donde se encuentra el icono de los Map Attributes (recuadrado en rojo) dentro del apartado Properties

A continuación se muestra una ventana MAP Attributes donde se pueden visualizar los atributos de dos elementos puntuales.



Todos aquellos atributos cuyo nombre va precedido de una almohadilla (#) son calculados por MAPublisher automáticamente.



Map Themes

Los MAP Themes (temas de mapa) hacen uso de los datos de atributos para aplicar estilos gráficos, símbolos o estilos de caracteres de Adobe Illustrator a características del mapa en función de los valores de atributo de un objeto. Se puede elegir entre tres tipos de Map Theme: Hoja de estilo, Gráficos y Densidad de Puntos.

Localización del Icono de los Map Themes (recuadrado en rojo) dentro del apartado Thematic Tools

El Map Theme es la utilidad de Mapublisher que asocia estilos gráficos y de texto (propios de Illustrator) con atributos de elementos (propios de un SIG), de tal manera que en función de los atributos se simbolizan los elementos importados siempre atendiendo a su tipo de geometría: líneas, puntos, áreas o textos.

En la siguiente captura se muestra la ventana Map Themes de una hoja de cartografía del IGN de España, donde se pueden apreciar los distintos tipos de temas del mapa:

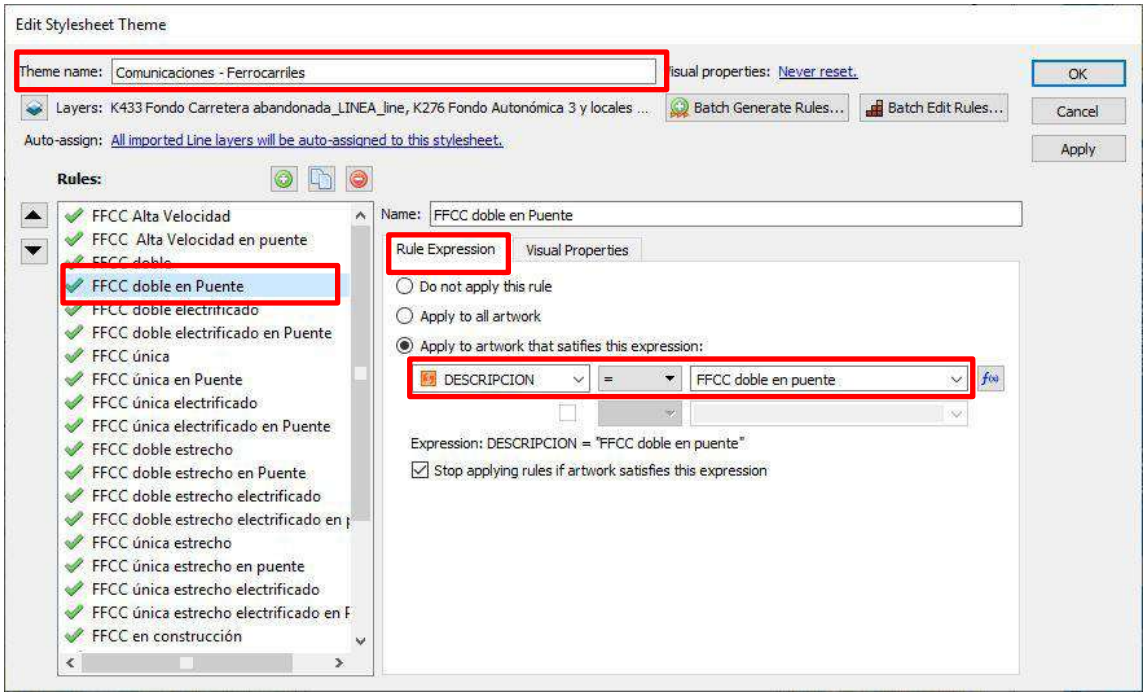
En verde: temas de texto

En amarillo: temas de punto

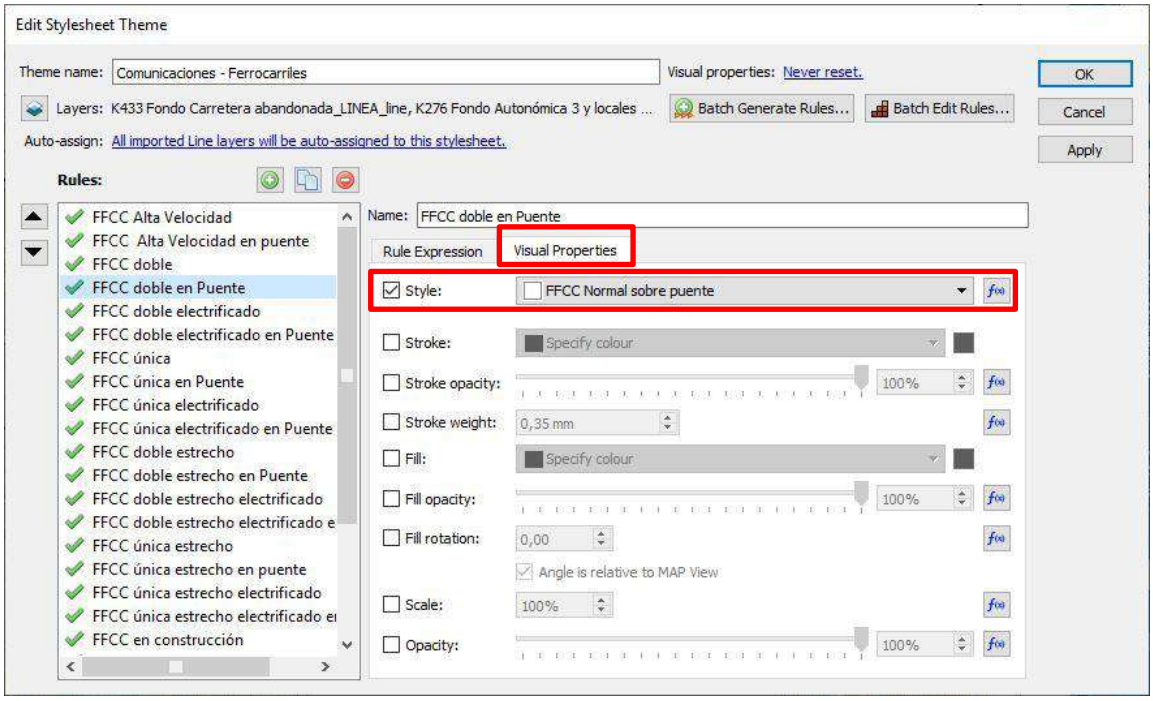
En azul: temas de línea

En rojo: temas de área

A continuación se muestra la ventana de Edición de una hoja de Estilo del Map Theme de Comunicaciones - Ferrocarriles (Tipo Línea) de una hoja de cartografía del IGN DE ESPAÑA, donde se encuentran recuadrados en rojo el nombre del Map Theme (Comunicaciones -Ferrocarriles), el elemento a simbolizar (FFCC doble en puente), la pestaña Rule Expresión (reglas de simbolización) donde se establece el atributo a través del cual se simboliza (DESCRIPCIÓN) y el valor de dicho atributo (FFCC doble en puente)



En la siguiente captura se observa la misma ventana tras haber accedido a la pestaña de propiedades visuales (Visual Properties), en la cual se establece el estilo gráfico de Illustrator con el que se debe representar el elemento (FFCC Normal sobre puente).



MAPublisher también ofrece otras muchas funciones y herramientas para administrar datos SIG desde Adobe Illustrator entre las que se encuentran:

- Importar los principales formatos SIG y CAD (Importación simple e Importación de datos múltiples de MAPublisher)
- Exportar a PDF geoespacial (GeoPDF)
- Exportar capas a los principales formatos SIG y CAD (Exportar MAP Layers, exportar MAP View)
- Exportar imágenes georreferenciadas en distintos formatos (Exportar imagen seleccionada, Exportar documento a imagen)
- Capacidad para dibujar símbolos puntuales a través de sus coordenadas (MAP Point Plotter)
- Manejo de datos de georreferenciación (MAP Locations, Georreferenciador)
- Recortar la extensión geográfica del mapa sin que los elementos pierdan sus atributos (Recorte vectorial, Recortar para dar forma)
- Unir líneas o puntos que comparten un mismo valor de atributo, simplificar líneas, crear un buffer alrededor de una línea (Zonas de influencia), cambiar el sentido de las líneas, unir áreas, unir líneas, unir puntos, trazar líneas a través de coordenadas, generalizaciones.
- Etiquetar elementos automáticamente según datos de sus atributos (MAP LabelPro, Label Features y MAP Tagger Tool)
- Seleccionar elementos a través de atributos (MAP Selection)
- Dibujar objetos con dimensiones geográficas a escala (MAP Plotter rectangle y MAP Plotter Ellipse)

- Crear cuadrículas, retículas e índices (Graticules, Measured grid e Index grid) a través de capas especiales de leyenda (Legend)
- Crear escalas gráficas y flechas de norte (Scale bar, North arrow)

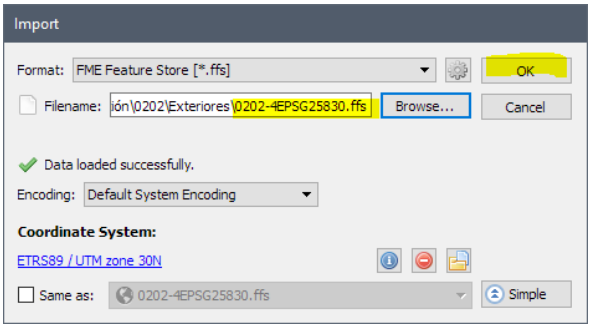
Descripción de las fases del proceso

El proceso manual de simbolización se compone de 5 pasos que son los que se describen a continuación.

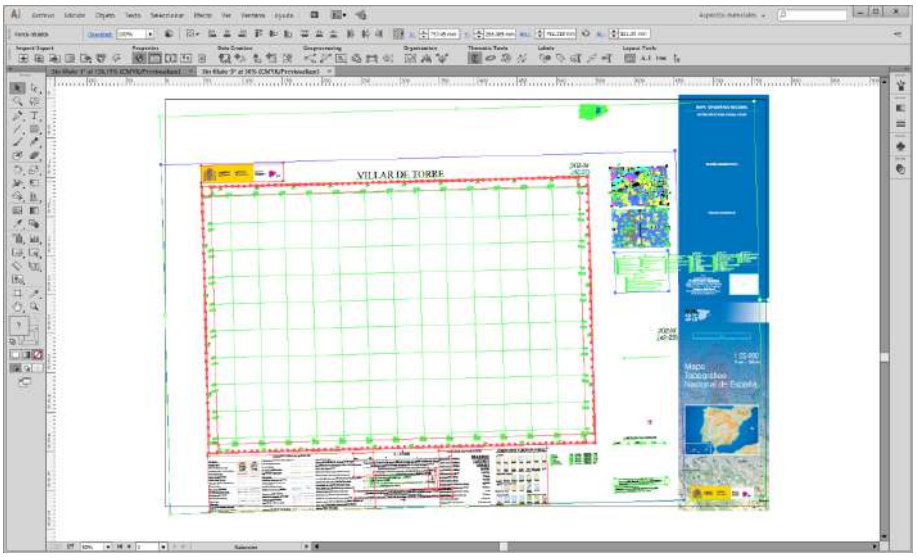
PASO 1. Montaje de la Hoja

Abrir Adobe Illustrator/Mapublisher con la plantilla BTN mediante el archivo “Plantilla BTN.ait”

Una vez abierto el documento/plantilla importaremos el archivo ffs de exteriores (por ejemplo “0202-4EPSG25830.ffs”)

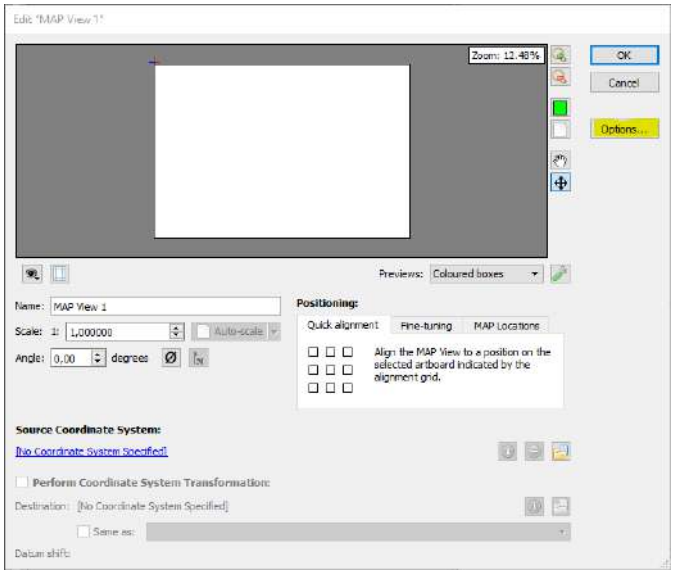


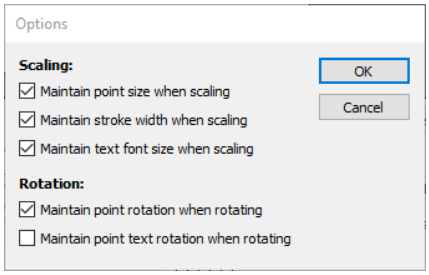
Una vez importados los datos de exteriores (que entre otras cosas son los que delimitan las dimensiones finales del trabajo) el aspecto de la hoja de cartografía será semejante a la siguiente imagen.



Configuración importante

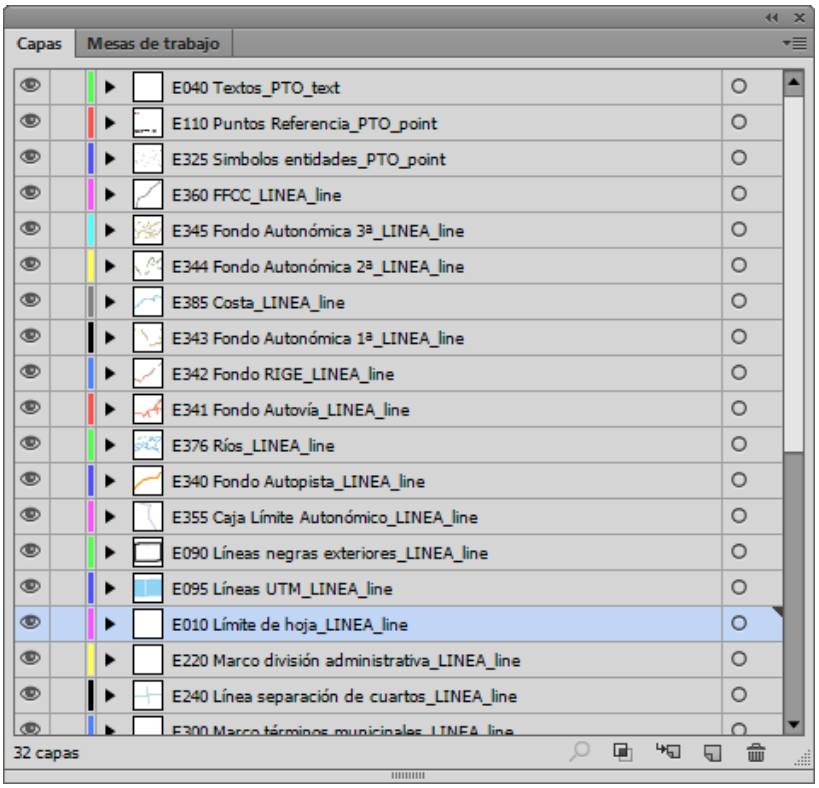
Antes de incorporar los elementos interiores, es muy importante que la configuración de Map Views de Mapublisher tenga contemplado lo siguiente (accedemos desde el apartado Options):





Esta configuración deberá ser la habitual para evitar problemas de giro en los textos. Hay que decir que una vez establecidos estos parámetros la configuración no cambia hasta que no se vuelva a editar.

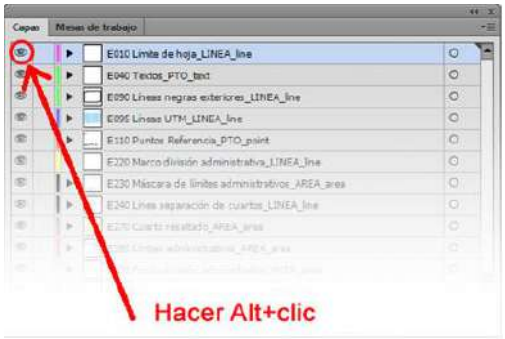
A continuación deberemos localizar la capa **E010 Límite de hoja_LINEA_line.** Para acceder a ella lo haremos mediante el menú **Ventana→Capas**



Una vez localizada la capa **E010 Límite de hoja_LINEA_line** , pulsaremos **Alt+clic** sobre el icono del ojo que precede a su nombre para visualizar solamente dicha capa.

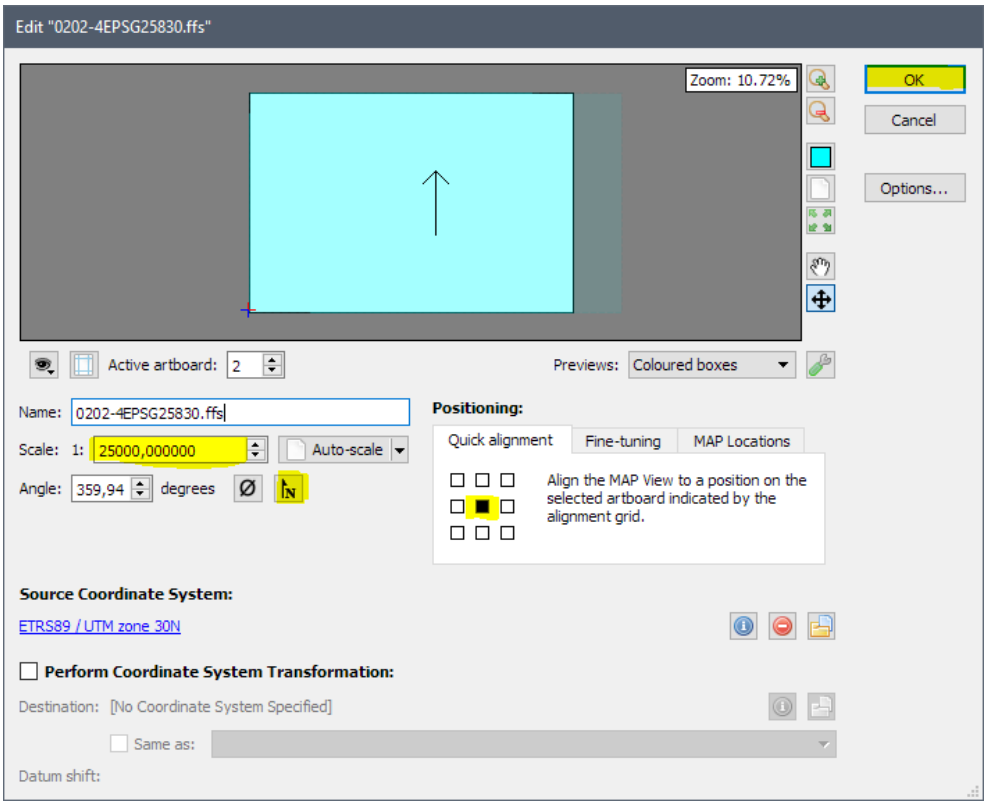
De esta manera se desactiva la visualización de todas las capas salvo la que deseamos mantener.

El aspecto de la ventana de capas quedaría como sigue:

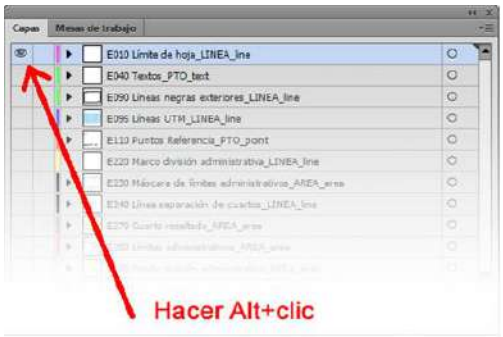


Acto seguido deberemos editar el Map View para **poner la escala correcta (25,000), quitar la convergencia y centrar el trabajo sobre el punto central**. Para ello pulsaremos dos veces el icono de Map Views.

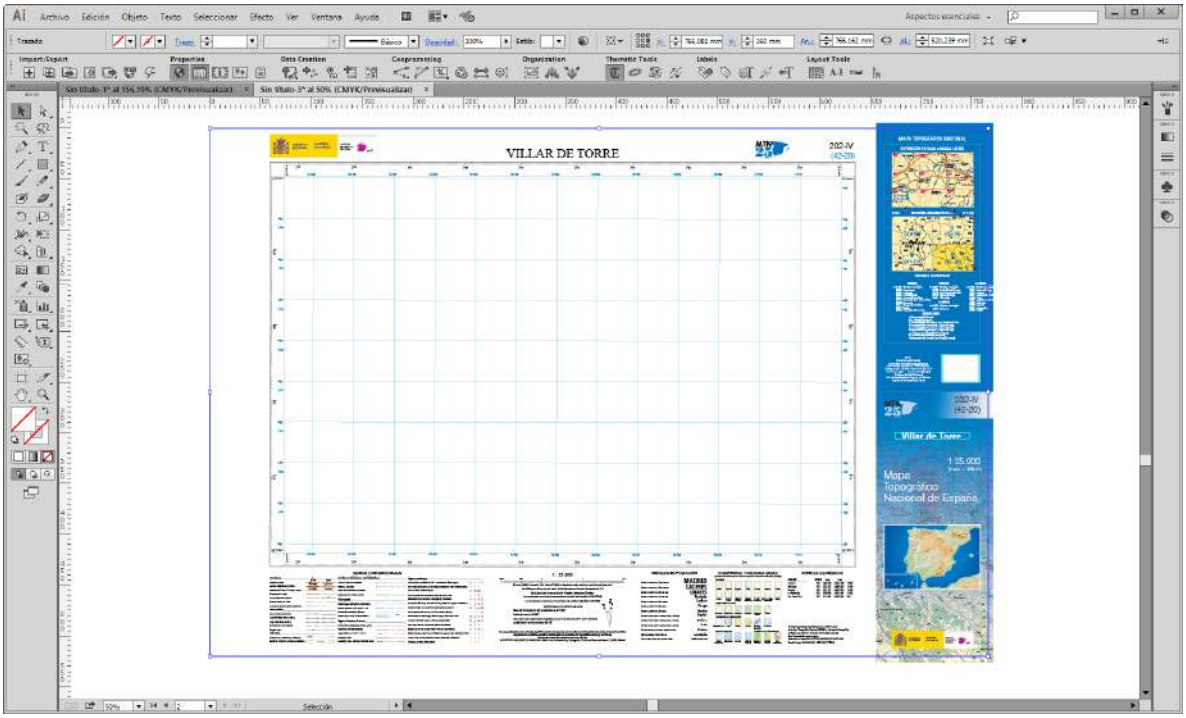
Y nos aparecerá la siguiente ventana, donde pondremos los datos en los lugares marcados de amarillo



Después de esto, hay que volver a visualizar el resto de capas que habíamos ocultado, para que así aparezcan todas.



Una vez activadas de nuevo todas las capas, el aspecto de la hoja será similar al siguiente:



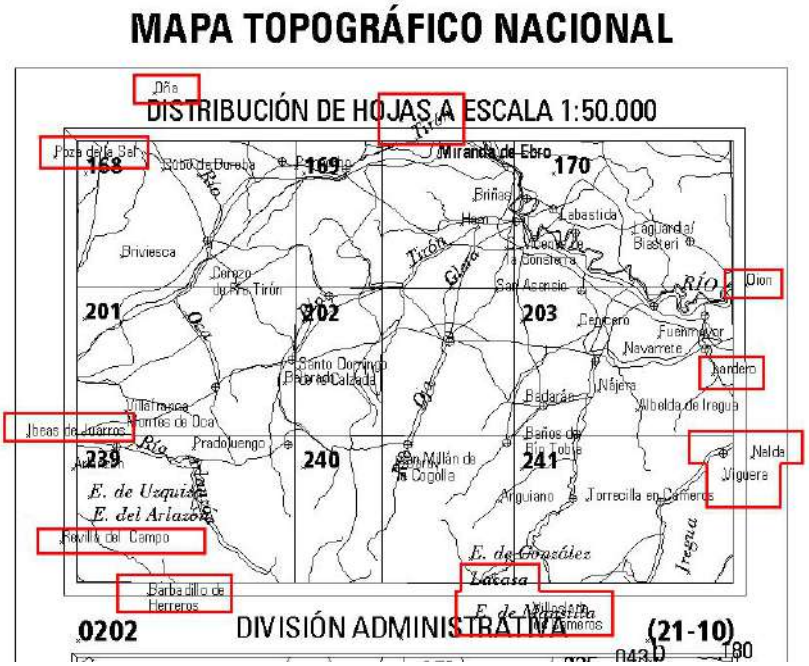
Retoques en la portada

Hay que fijarse si hay textos en la esquina superior derecha de la portada que se salgan del rectángulo del croquis. En su caso desplazar los textos para que queden en el interior si deben figurar o eliminarlos si no apuntan a ningún elemento del croquis.

Puede suceder que existan textos (generalmente azules) que se salgan del croquis y que a primera vista no localicemos por confundirse con el azul de la portada.

Para visualizar esos posibles textos que se confunden, deberemos cambiar el modo de visualización **a Contornear (Menú ver→Contornear)** al cual se puede acceder también pulsando **Ctrl+Y**.

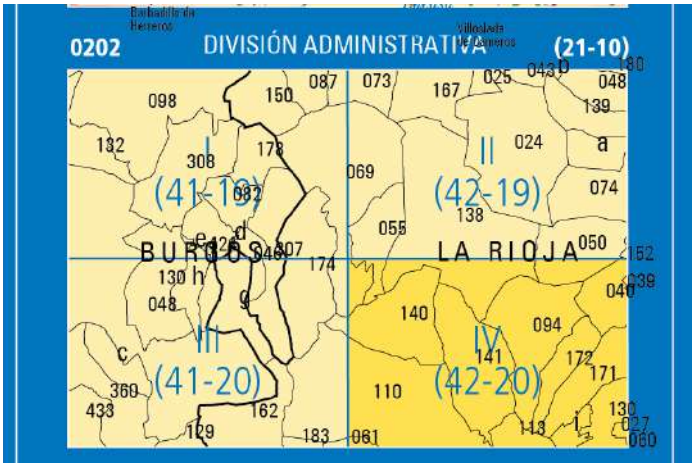
La portada se visualizará de la siguiente manera:



Para volver a ver el trabajo a todo a todo color deberemos activar el modo **Previsualizar (Menú ver→Previsualiza**r) al cual se puede acceder también pulsando Ctrl+Y.

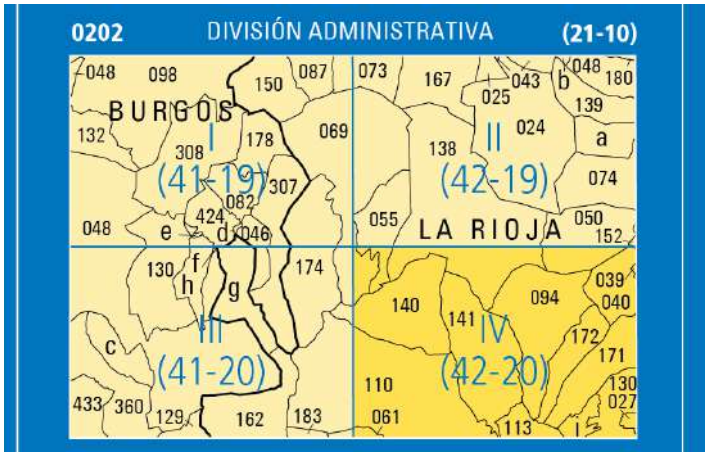
Después de una cuidadosa edición de los textos el resultado es el siguiente:

De igual forma que hemos hecho con los textos de la distribución de hojas, hay que recolocar los números de los términos municipales en el centro de las pequeñas áreas que representan su término.



En algunos casos se hará necesario dibujar una pequeña línea que ayude a un número a apuntar a su correspondiente área. Dicha línea se dibujará en negro con un grosor de 0,10 mm. Dicha línea se debe dibujar en la capa **E040 Textos_PTO_text** y se puede usar el estilo gráfico GENERAL Línea negra 0,10 incluido en la plantilla.

El resultado de la edición de los números de términos municipales de la división administrativa es el siguiente:



Hay que verificar que ningún término o jurisdicción queda sin su número correspondiente.

Puede que los nombres de los Términos municipales de la portada se salgan del área definida que debería contenerlos. Esta área es un rectángulo invisible dibujado al efecto (elemento con atributo “Marco términos municipales”). Para ayudarse en la visualización de los elementos invisibles y textos que deben ser recolocados y comprobar que no nos dejamos ninguno por el camino, se puede (y se debe) usar la opción de visualizar en modo Contornear (Ctrl+Y). Para volver al modo previsualizar hay que volver a pulsar las mismas teclas (Ctrl+Y).

En la siguiente imagen se puede comprobar la existencia de textos habrá que recolocar dentro del rectángulo invisible de ayuda en la medida de lo posible (esto depende de la cantidad de términos municipales que tenga la hoja).

En ese caso habrá que recolocarlos de forma que queden dentro del rectángulo invisible dibujado al efecto (elemento con atributo “Marco términos municipales”). Este es un área que no se imprime y que ayuda a saber la zona donde tenemos que colocar estos elementos textuales. En el caso del ejemplo, y por la gran cantidad de textos, hay que recurrir a cambiar el cuerpo de la letra, el interlineado y acomodar todo el conjunto de textos de una manera lo más armónica posible para que entren por lo menos dentro del rectángulo blanco de la portada

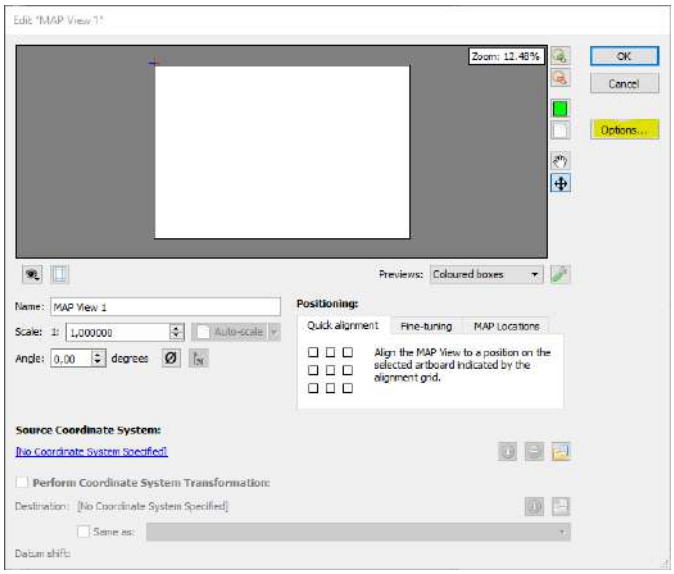
Hay que indicar que, por defecto, y si son relativamente pocos, no hará falta recolocar estos textos, pues el programa de exteriores ya los pone prácticamente en su sitio.

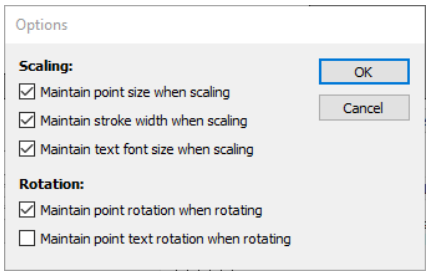
En este último caso, lo ideal es que el texto quede centrado sobre su cabecera de columna (MÁLAGA) quedando así:

PASO 2. Incorporación de interiores

Configuración importante

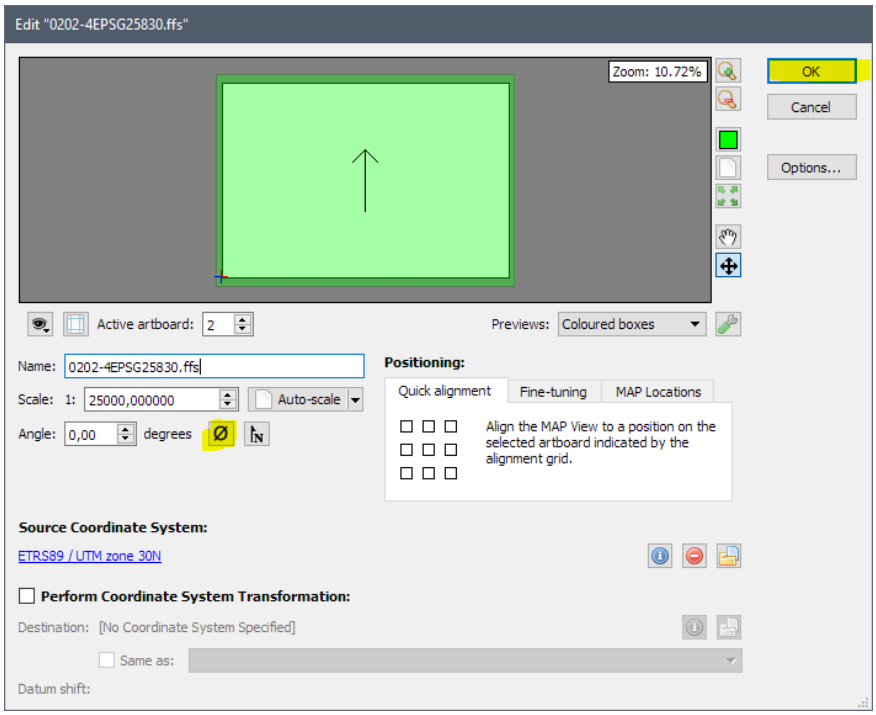
Antes de incorporar los elementos interiores, es muy importante comprobar de nuevo que la configuración de Map Views de Mapublisher tenga contemplado lo siguiente (accedemos desde el apartado Options):





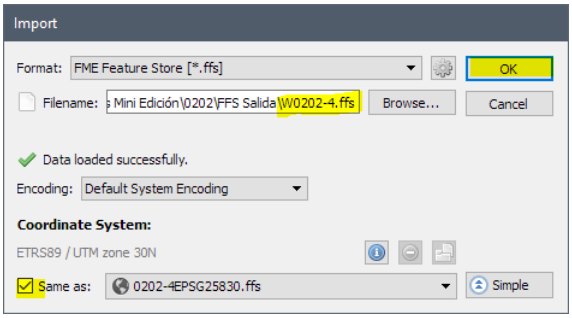
Hay que decir que una vez establecidos estos parámetros la configuración no cambia hasta que no se vuelva a editar.

Una vez comprobado que las opciones son las que se han indicado anteriormente, la primera acción a realizar es poner valor cero al ángulo de convergencia de la hoja, para que al incorporar los interiores los textos queden girados de igual manera que los exteriores.

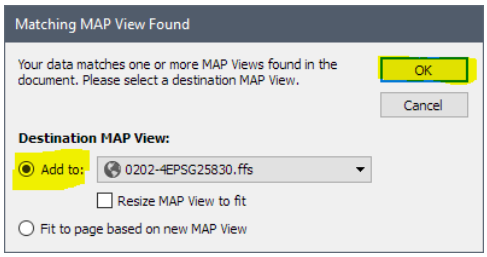


En este momento podremos Importar el archivo ffs de planimetría “W0202C4_H30.ffs” a través de la barra de herramientas de Mapublisher.

En la ventana que aparece deberemos indicar que el sistema de coordenadas es el mismo que el de los exteriores que habíamos importado anteriormente.

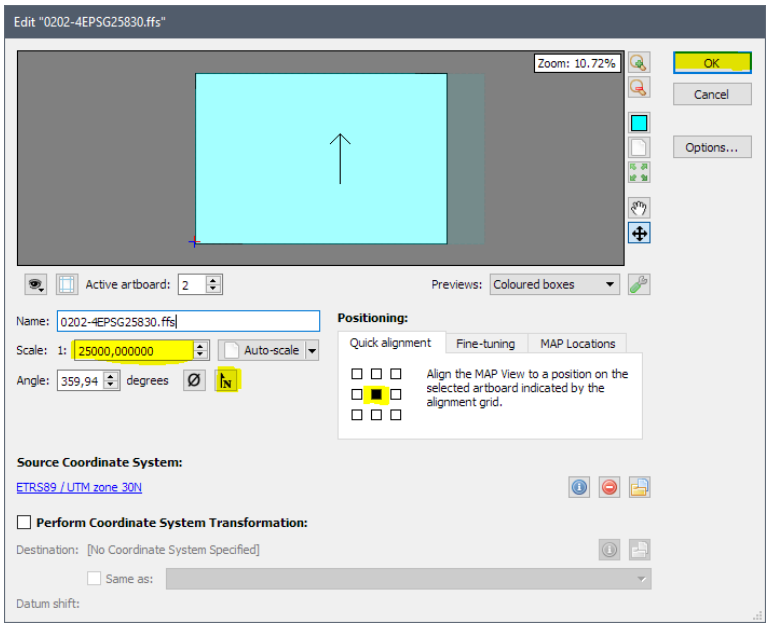


Posteriormente hay que indicar que el Map View de destino es el mismo que el de los exteriores. De esta manera nos aseguramos que todo tiene las mismas coordenadas y se encuentra en el mismo Map View.

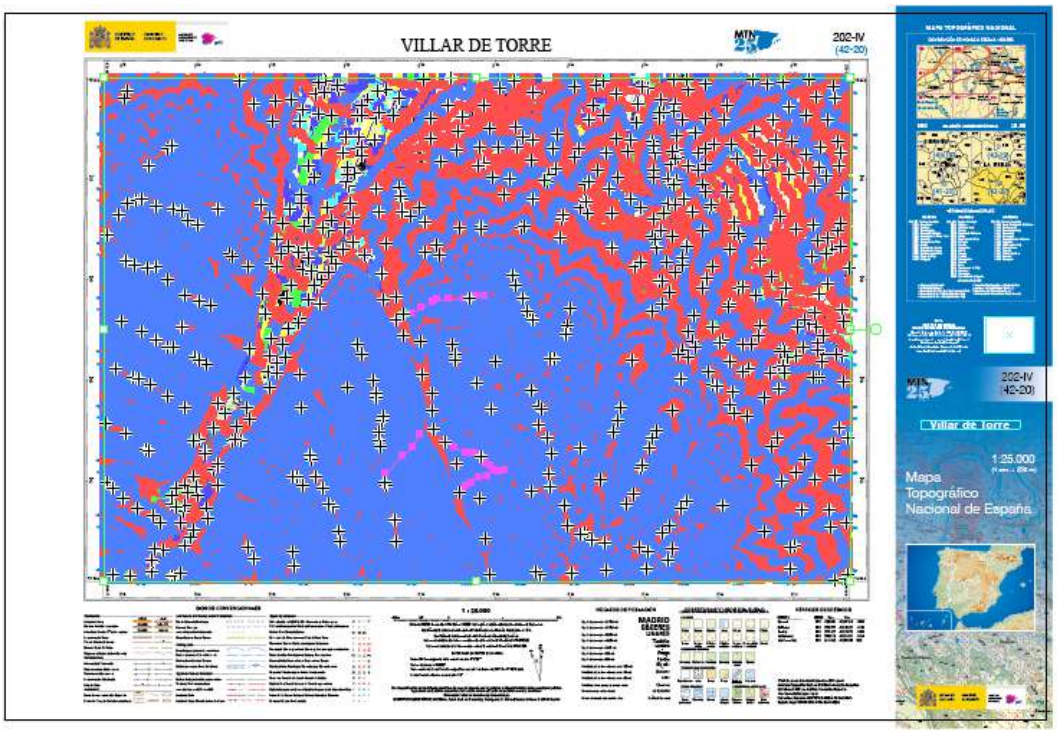


La importación tarda un rato, dependiendo de la cantidad de información que contenga el archivo de interiores (.ffs)

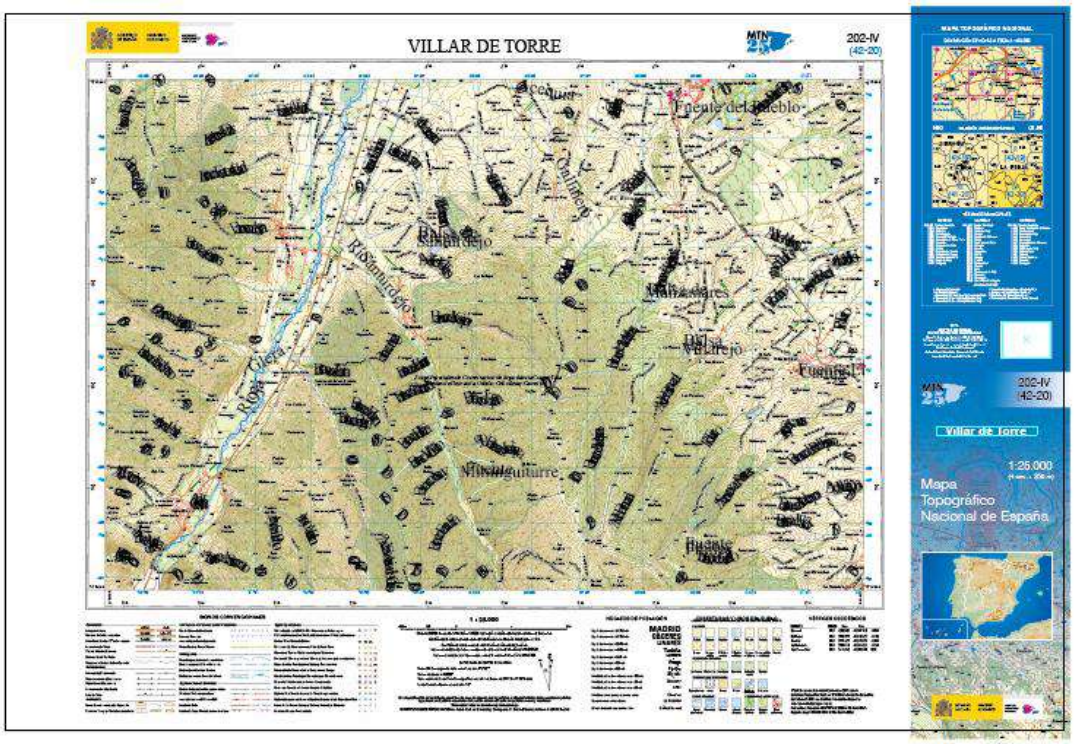
De nuevo hay que volver a poner la convergencia de la hoja que acabamos de quitar para importar los interiores.



Ahora el aspecto de la hoja será similar al de la siguiente imagen



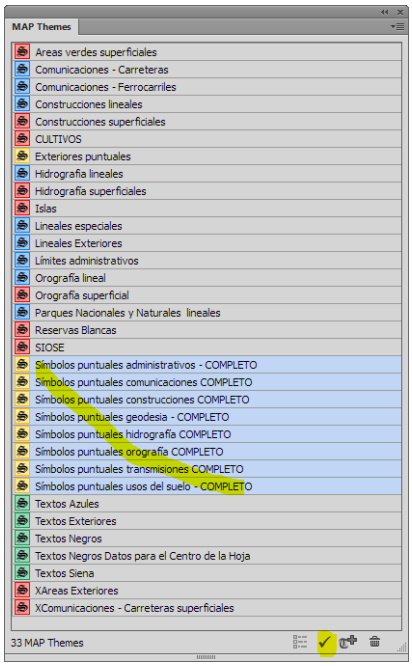
Para comprobar la simbolización hay que pinchar con el cursor negro en cualquier parte del mapa que no tenga trabajo (por ejemplo, en el exterior de la mesa de trabajo)

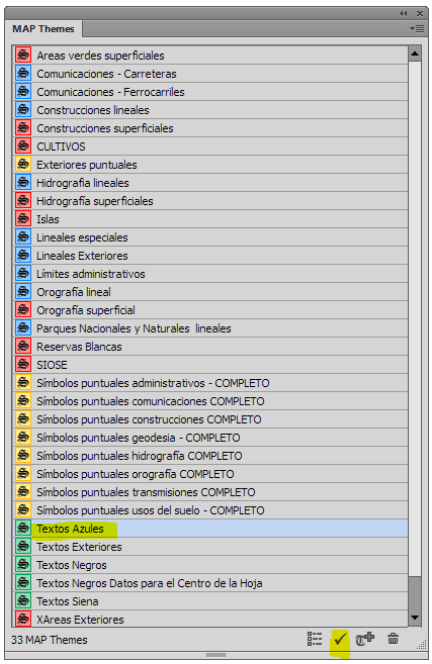


De esta manera comprobamos que hay elementos que quedan sin simbolizar (generalmente textos y símbolos puntuales).

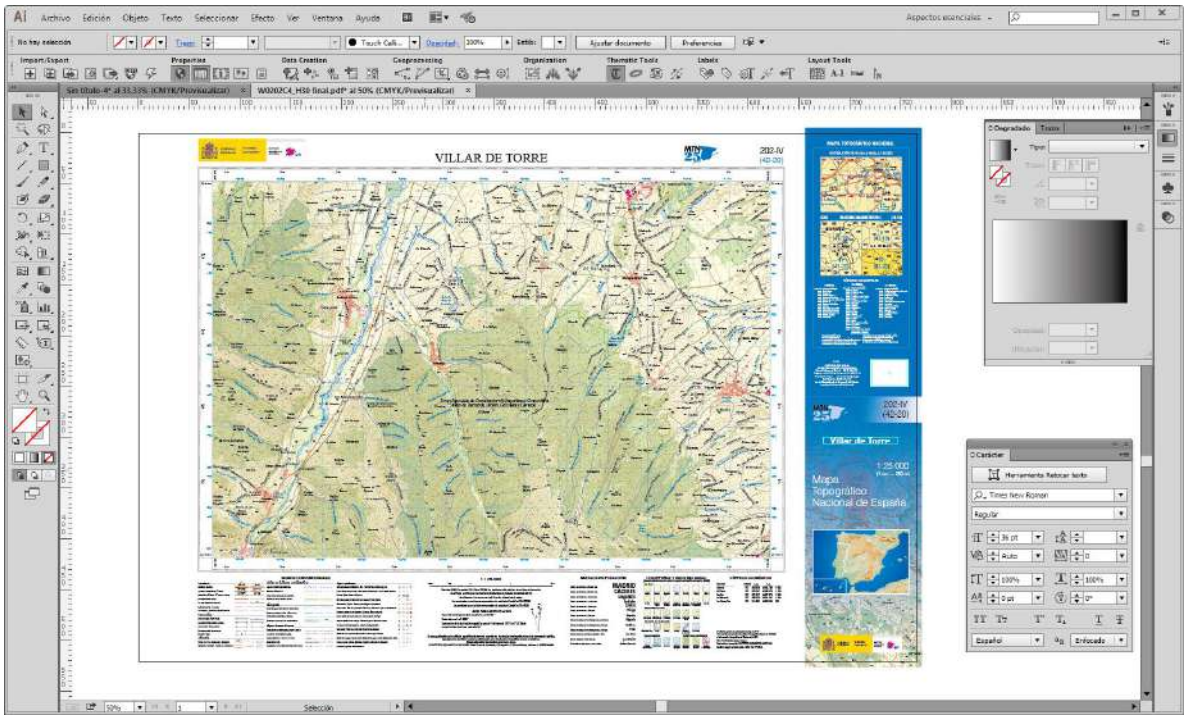
Resimbolización

Para continuar hay que resimbolizar los elementos que no aparecen bien simbolizados. En principio suelen ser todos los símbolos puntuales, así como los Textos azules, Textos negros y Textos siena





Después de resimbolizar, hay que ordenar las capas alfabéticamente. Para ello se puede utilizar un script de Illustrator, aunque se puede hacer de manera manual, quedando la hoja completamente formada a falta de pequeñas comprobaciones y de incorporar el sombreado.



PASO 3. Comprobación de encaje entre interiores y exteriores

Lo primero que deberemos hacer es revisar legibilidad de los textos de las coordenadas geográficas del marco y su posible intersección con las coordenadas UTM de la cuadrícula.

VÉRTICES GEODÉSICOS

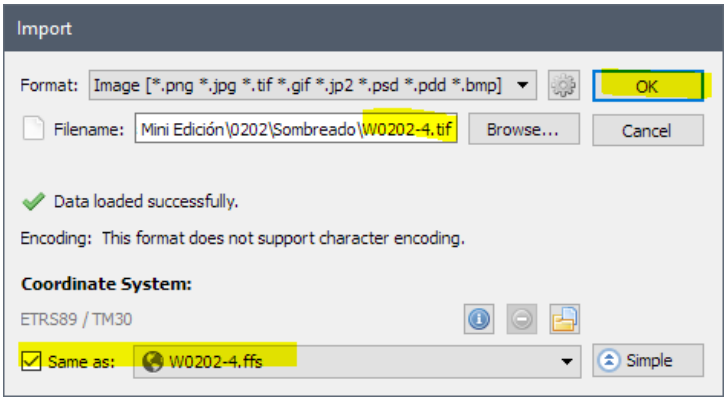
NOMBRE	ORDEN	X_{UTM}	Y_{UTM}	
Sampol	ROI	499.538	4.694.789	1.094
Bellicera	ROI	503.347	4.689.528	1.135
Fontría	ROI	507.127	4.687.675	1.304
La Retuerta	ROI	506.819	4.691.491	1.154
Las Coronillas	ROI	511.452	4.690.546	823

A continuación, debemos comprobar que los vértices geodésicos coinciden con el listado. Visualizando solo la capa “K205 Símbolos puntuales de geodesia” y cotejando el número con la lista escrita.

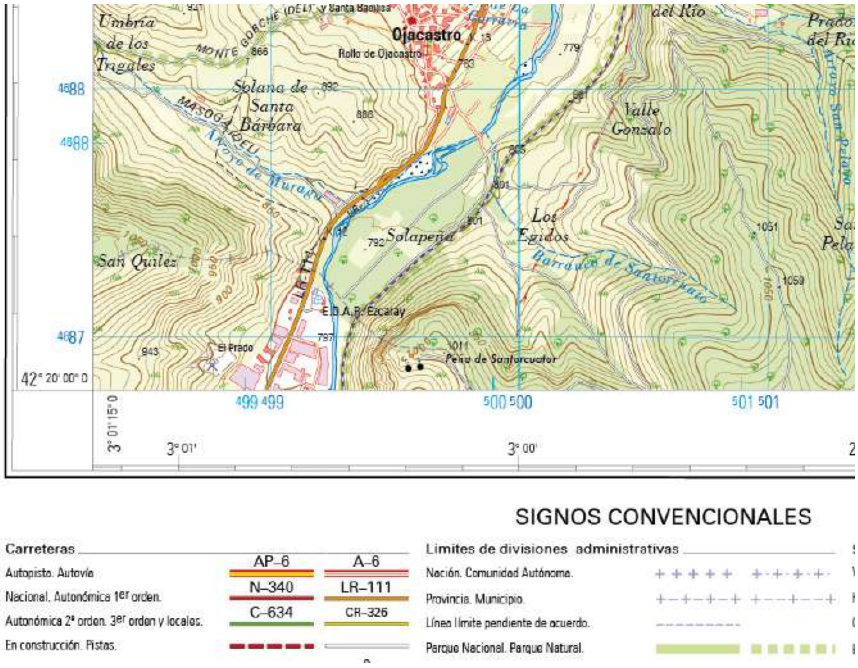
También hay que comprobar que no hay textos del interior que queden cortados en los bordes de la hoja (mares, ríos, poblaciones, sitios, parques naturales y nacionales, etc…)

PASO 4. Incorporación del sombreado

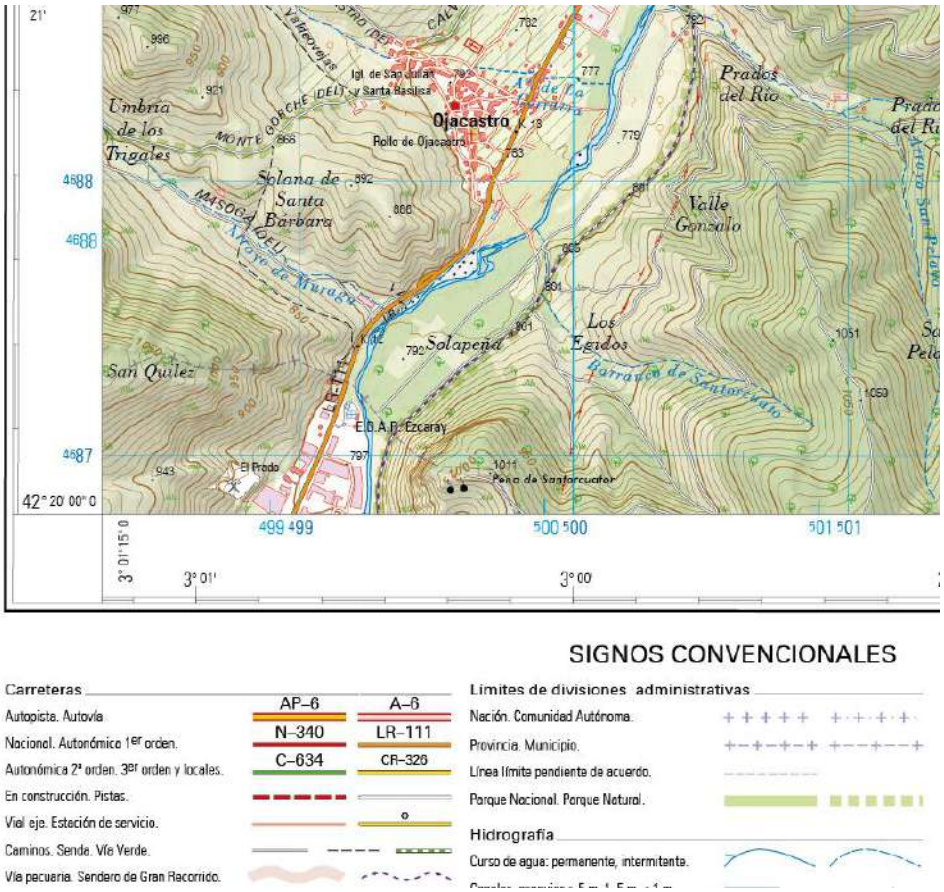
Como último paso del montaje de la hoja, y siempre que nuestra serie cartográfica nos lo demande, deberemos importar con Mapublisher el archivo de sombreado “w0202-4_30.tif”, que gracias a su georreferenciación debería entrar con las coordenadas correctas de la hoja que estamos montando.



Para comprobar la diferencia que existe entre ellos, se muestran en las siguientes imágenes un mapa sin sombreado:

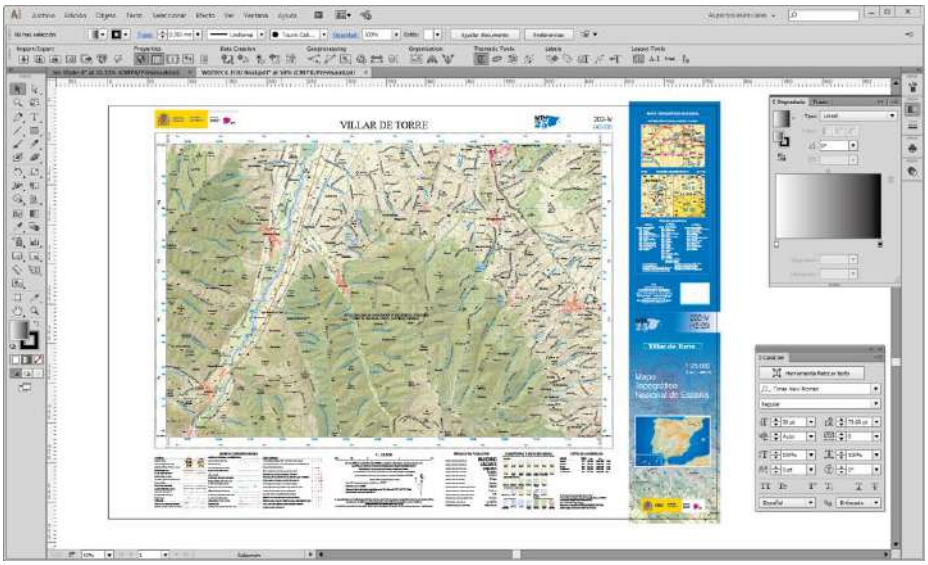


y un mapa con sombreado, en el cual se aprecia el valor gráfico añadido que se aporta:



PASO 5. Guardado del documento

Una vez completado el montaje y la simbolización de la hoja, queda únicamente guardar el documento para su archivo y posterior edición o su publicación impresa (bien sea en imprenta tradicional en máquinas offset o imprenta digital mediante plotters de alta producción)

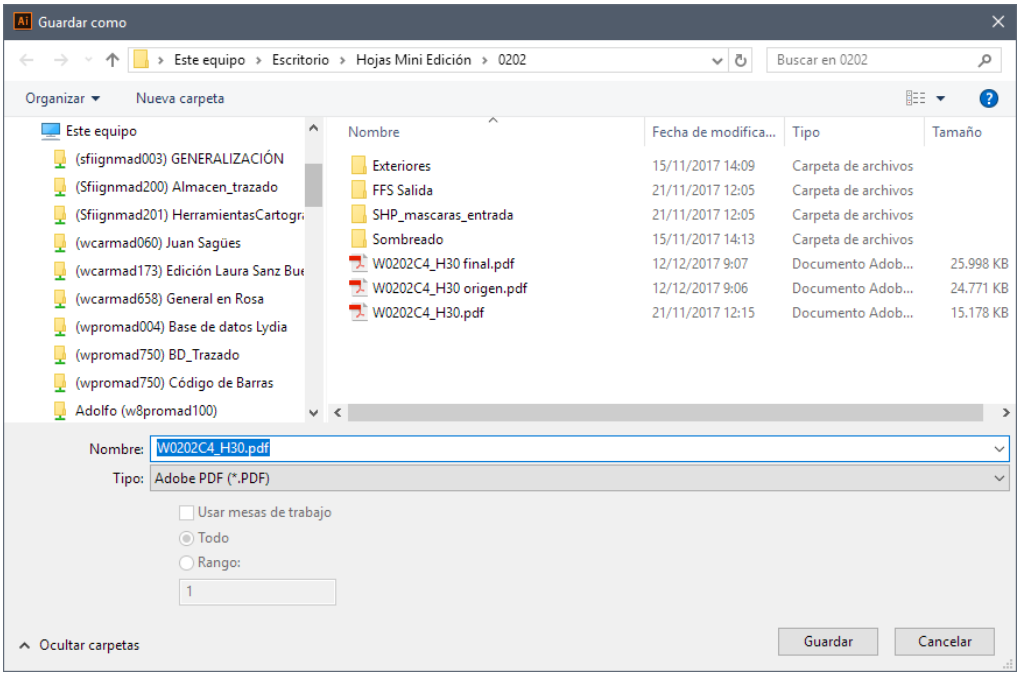


Para guardar el documento de una manera que nos sirva para diversos propósitos procederemos de la siguiente manera:

Iremos al **Menú Archivo→Guardar como**

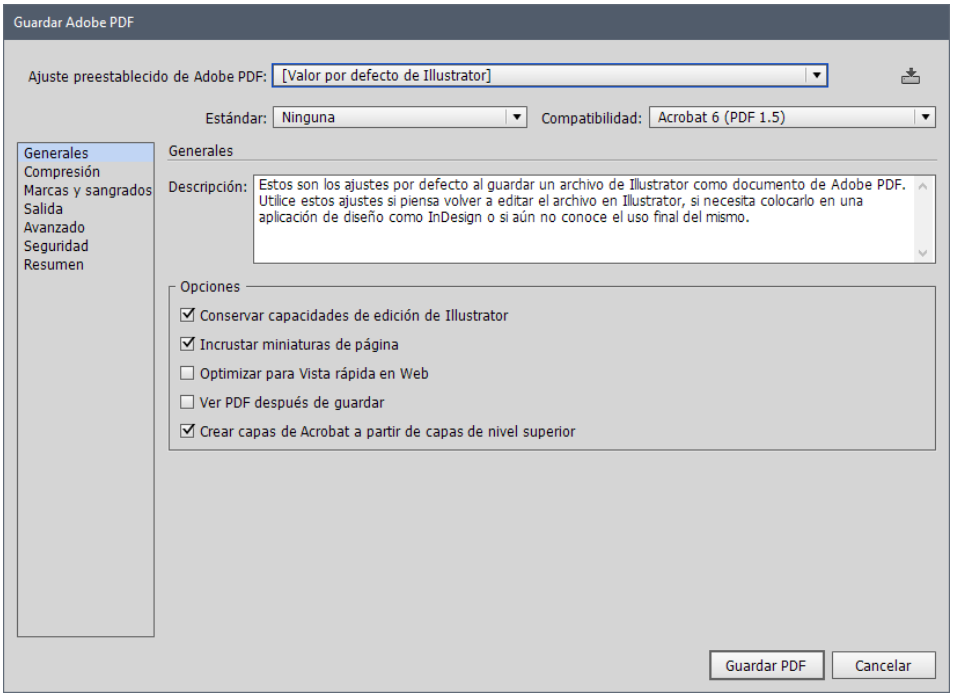
Tipo: Adobe PDF

Nombre: “W0202C4_H30.pdf” (igual que el ffs original de planimetría)



Cuando confirmemos, aparecerá otra ventana donde indicaremos lo siguiente:

Ajuste preestablecido de Adobe PDF: **Valor por defecto de Illustrator**

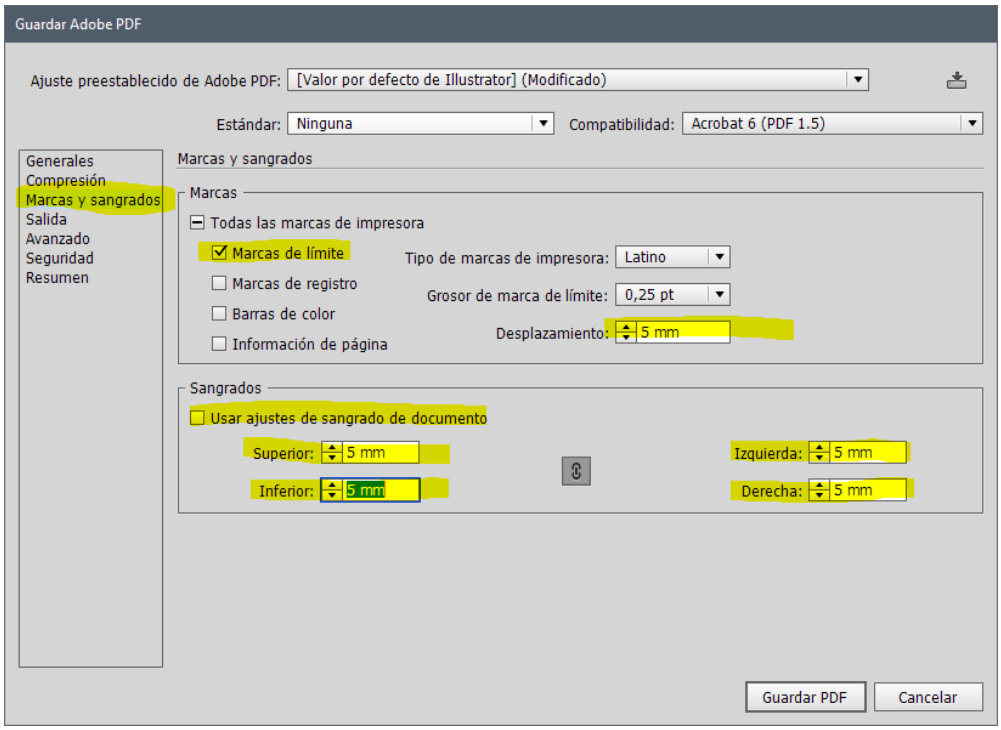


Posteriormente, en la pestaña de Marcas y Sangrado señalar:

Marcas de límite

Desplazamiento: 5mm

Deshabilitar los Ajustes de sangrado de documento y poner 5mm en los cuatro campos (Superior, Inferior, Izquierda, Derecha)



Una vez guardado el documento, concluye el proceso de simbolización de hojas individuales para publicación impresa.

6.3.2. Proceso automático de simbolización y rasterización

Software de producción. Consideraciones previas

PostGis: almacenamiento de los datos cartográficos

El proceso de simbolización y rasterización tiene como punto de partida los datos cartográficos generados en el proceso de generación automática del mapa (explicado en el Capítulo 5), almacenados en una base de datos PostGis.

Cada tabla está numerada desde 00 a 11 según su temática (comunicaciones, construcciones, cultivos, etc.) albergando en cada una de ellas los distintos tipos de geometrías. Su nomenclatura, y a modo de ejemplo es 00p (Temática_entidades puntuales), 00l (Temática_entidades lineales) y 00s (Temática_entidades superficiales). Dentro de cada tabla, cada tupla, que corresponde a una entidad, posee una serie de atributos, siendo estos fundamentales a la hora de la simbolización de la que hablaremos más adelante.

Por otra parte, y dentro de la Base de Datos en PostGIS, existe la **tabla hoja_c** en la que están almacenados cada uno de los marcos de hoja y que, a través de una clave primaria se relacionan con la clave externa de cada una de las tablas siendo, por tanto, una base de datos relacional.

Tabla hoja_c

	a	b	c	d	e	f	g	h	i	j	k
	id (PK) integer	num_hoja character varying (50)	fecha_alta timestamp with time zone	procesador character varying (100)	huso_marco integer	huso2_marco integer	solape_marco integer	nombre_marco character varying (250)	geom_marco geometry	fecha_act date	tipo_hoja character varying (20)
1	17439	0001-4	2021-09-18 18:26:20.095712+01	WCARMAD987	29	99	99	CARIÑO	0103000020A2100000...	1111-11-11	MTN
2	17441	0002-1	2021-09-18 19:01:24.572096+01	WCARMAD987	29	99	99	CELEIRO	0103000020A2100000...	1111-11-11	MTN
3	18479	0002-3	2021-09-21 09:56:57.209152+01	WCARMAD987	29	99	99	Ortigueira	0103000020A2100000...	1111-11-11	MTN
4	30994	0002-4	2023-05-29 11:30:19.036014+02	WCARMAD987	29	99	99	Celeiro	0103000020A2100000...	1111-11-11	MTN

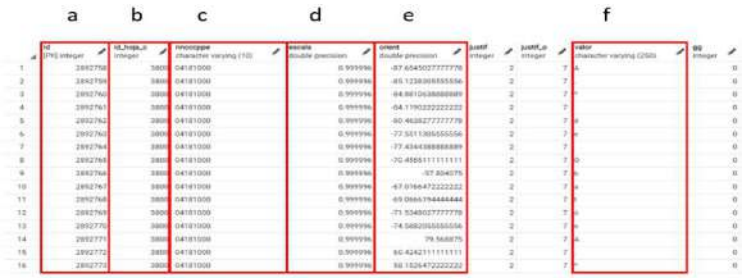
Ilustración 6 - 45: Tabla hoja_c en PostGIS

ATRIBUTOS IMPLICADOS EN LA TABLA "**hoja_c**"

a	Clave primaria de la tabla. Se relacionará con la clave externa del resto de tablas.
b	Número de hoja y cuarto que corresponde a la serie MTN25
c	Fecha en que ha sido generada la hoja
d	Equipo que lo ha procesado
e	Huso UTM de la hoja. Su dominio toma los valores 28, 29, 30 y 31 (en el caso de República Dominicana tomaría el valor 19)
f	Atributo que se utilizaba para reproyectar las hojas para usuarios sin posibilidad de reproyectar al vuelo.
g	No usado.
h	Nombre de la hoja.
i	Geometría que define el marco.
j	No usado.
k	No usado.

Ilustración 6 - 46: Atributos implicados

Tabla 08p (Tabla de textos)



	a	b	c	d	e	f
	id (PK) integer	id_hoja_c integer	rnocoppe character varying (10)	escala double precision	orient double precision	justif integer	justif_o integer	valor character varying (250)	dg integer
1	2892758	3800	04181000	0.999996	-87.6545027777778	2	7	A	0
2	2892759	3800	04181000	0.999996	-85.1238366555556	2	7	[illegible]	0
3	2892760	3800	04181000	0.999996	-84.8610638888889	2	7	[illegible]	0
4	2892761	3800	04181000	0.999996	-84.1190222222222	2	7	[illegible]	0
5	2892762	3800	04181000	0.999996	-80.4628277777778	2	7	[illegible]	0
6	2892763	3800	04181000	0.999996	-77.5311386555556	2	7	[illegible]	0
7	2892764	3800	04181000	0.999996	-77.4344388888889	2	7	[illegible]	0
8	2892765	3800	04181000	0.999996	-70.4868111111111	2	7	[illegible]	0
9	2892766	3800	04181000	0.999996	-67.804875	2	7	[illegible]	0
10	2892767	3800	04181000	0.999996	-67.0166472222222	2	7	[illegible]	0
11	2892768	3800	04181000	0.999996	-69.0866194444444	2	7	[illegible]	0
12	2892769	3800	04181000	0.999996	-71.5346027777778	2	7	[illegible]	0
13	2892770	3800	04181000	0.999996	-74.5802555555556	2	7	[illegible]	0
14	2892771	3800	04181000	0.999996	-79.568875	2	7	[illegible]	0
15	2892772	3800	04181000	0.999996	80.4242111111111	2	7	[illegible]	0
16	2892773	3800	04181000	0.999996	88.1826472222222	2	7	[illegible]	0

Ilustración 6 - 47: Tabla 08p de textos

ATRIBUTOS IMPLICADOS EN LA TABLA “**08p**” (aplicable al resto de tablas).

a	Valor correlativo de cada elemento
b	Clave externa de la tabla que se relaciona con el atributo **id** de la tabla hoja_c
c	Nivel-Color-Peso-Estilo de la entidad procedente de archivos dgn de cada hoja
d	Escala
e	Orientación (en este caso del texto)
f	Valor del elemento (letra de cada texto)

Ilustración 6 - 48: Atributos implicados

Como se puede ver, en la tabla correspondiente a la geometría, el atributo fundamental que es necesario a la hora de simbolizar es “nncccppe” (Nivel Color Peso y Estilo). Este atributo, procede de una concatenación de Nivel, Color, Peso y Estilo presentes en los archivos.dgn mediante el cual se clasifican todas y cada una de las entidades contenidas en dichos archivos.

Ventana de datos de un dgn en FME

Para obtener nncccppe se concatenan los atributos del archivo dgn:

- igds_level
- igds_color
- igds_weight
- igds_style

Feature Information

Property	Value
Number of Vertices	1
Min Extents	286911.5, 4233700.843670267
Max Extents	287115.9044742417, 4233743.18738111
Attributes (48)	
fme_color (string)	0.23529411764705882,0,0
fme_dataset (string)	\\sfiignmad162\mtnbtn$\MTNBTN\MAPA_AUTOMATICO\1.DGN\a0879-3_30.dgn
fme_geometry (string)	fme_point
fme_rotation (64 bit real)	-1.50825
fme_text_size (64 bit real)	37.0002
fme_text_string (string)	Almodóvares
fme_type (string)	fme_text
igds_basename (string)	a0879-3_30
igds_class (32 bit integer)	0
igds_color (32 bit integer)	219
igds_color.blue (32 bit integer)	0
igds_color.green (32 bit integer)	0
igds_color.red (32 bit integer)	60
igds_element_locked (string)	no
igds_element_modified (string)	no
igds_element_new (string)	yes
igds_element_type (32 bit integer)	17
igds_font (32 bit integer)	40
igds_graphic_group (32 bit integer)	0
igds_insertion_x (64 bit real)	287109.25
igds_insertion_y (64 bit real)	4233719.5
igds_insertion_z (64 bit real)	0
igds_justification (32 bit integer)	2
igds_level (32 bit integer)	3
igds_lower_x (64 bit real)	286911.5
igds_lower_y (64 bit real)	4233695.8
igds_original_justification (32 bit integer)	7
igds_rotation (64 bit real)	-1.50825
igds_snappable (string)	yes
igds_style (32 bit integer)	0
igds_symbology (32 bit integer)	56064
igds_text_horizontal_flip (boolean)	No
igds_text_size (64 bit real)	37.0002
igds_text_string (string)	Almodóvares
igds_text_vertical_flip (boolean)	No
igds_text_width_multiplier (64 bit real)	37.0002
igds_type (string)	igds_text
igds_upper_x (64 bit real)	287307
igds_upper_y (64 bit real)	4233743.2
igds_weight (32 bit integer)	0
igds_xhigh (64 bit real)	287307
igds_xlow (64 bit real)	286911.5
igds_yhigh (64 bit real)	4233743.2
igds_ylow (64 bit real)	4233695.8
igds_zhigh (64 bit real)	0

Ilustración 6 - 49: Datos de un archivo dgn en la ventana Inspector de FME

Atributos dgn	Descripción	Valor en el ejemplo
igds_level	Nivel en la entidad	03
igds_color	Color de la entidad	219
igds_weight	Peso de la entidad	00
igds_style	Estilo de la entidad	0
igds_rotation	Valor correspondiente al valor de la convergencia de la hoja	-1.50825
igds_text_string	Contenido del texto	Almodóvares
	Concatenación nncccppe	03219000

Illustrator y Mapublisher, ejemplo de entidad

Por otra parte, para que la simbolización se lleve a cabo, es imprescindible contar con una plantilla de Mapublisher en la que se alberguen cada una de las entidades correspondientes, su clasificación y su simbología.

Asimismo, además de la plantilla es necesario definir la mesa de trabajo que albergará a la hoja. Hay que tener en cuenta que las dimensiones de la mesa de trabajo han de estar acordes al tamaño máximo de las hojas pues, como hemos dicho anteriormente, puede haber hojas con dimensiones mayores a los estándares. El tamaño de la mesa de trabajo son 900 mm. X 550 mm. correspondiendo a 22,500 m. en sentido de longitud y 14,250 m. en sentido de latitud, a escala 1:25,000. En la parte de rasterización se explicará la especialidad de esta conversión.

Configuración Mesa de Trabajo

- Anchura 900 mm que a escala 1:25,000 son 22,500 metros.
- Altura 570 mm que a escala 1:25,000 son 14,250 metros.

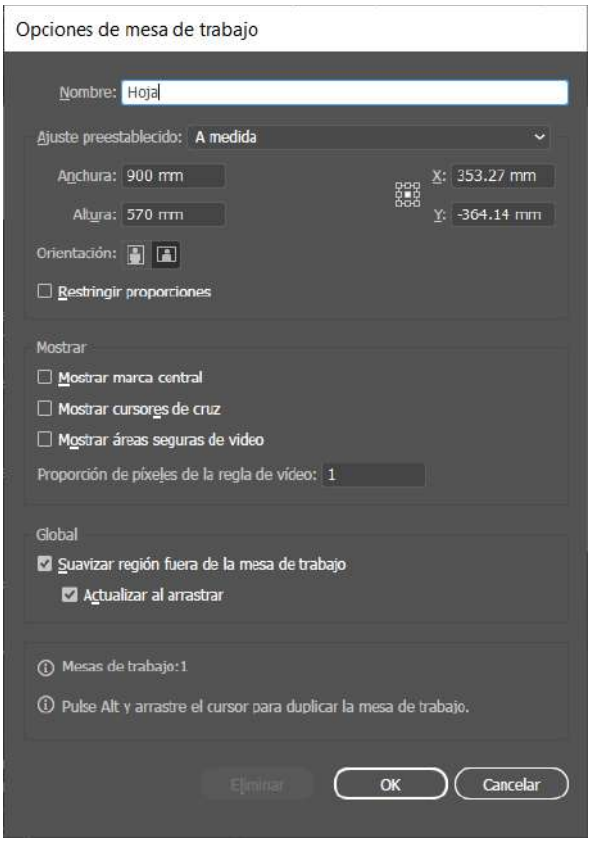


Ilustración 6 - 50: Configuración de la Mesa de Trabajo

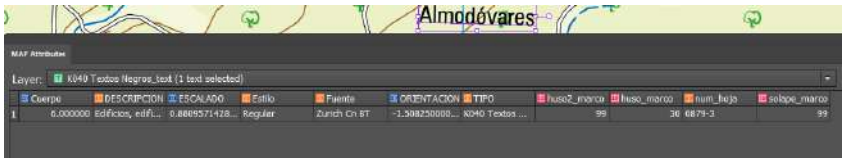


Ilustración 6 - 51: Entidad de texto con sus atributos en Mapublisher

A continuación, se expone el maptheme para las entidades (en este ejemplo de comunicaciones – carreteras).

En esta primera Ilustración 6 – 48 se señala el atributo (DESCRIPCION) y su valor (Fondo Autonómica 1ª) a través del cual se simboliza.

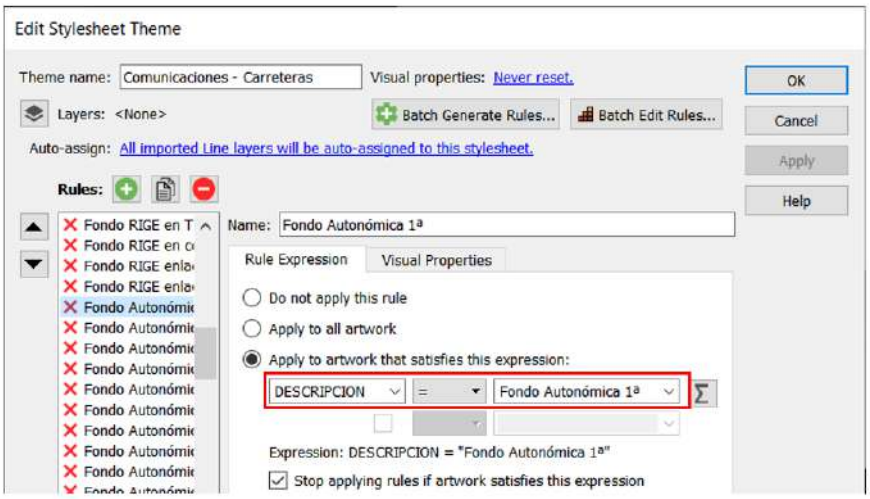


Ilustración 6 - 52: Regla para simbolización

En la segunda ilustración, se eligen las propiedades visuales del símbolo. En este caso la simbología es de un elemento lineal.

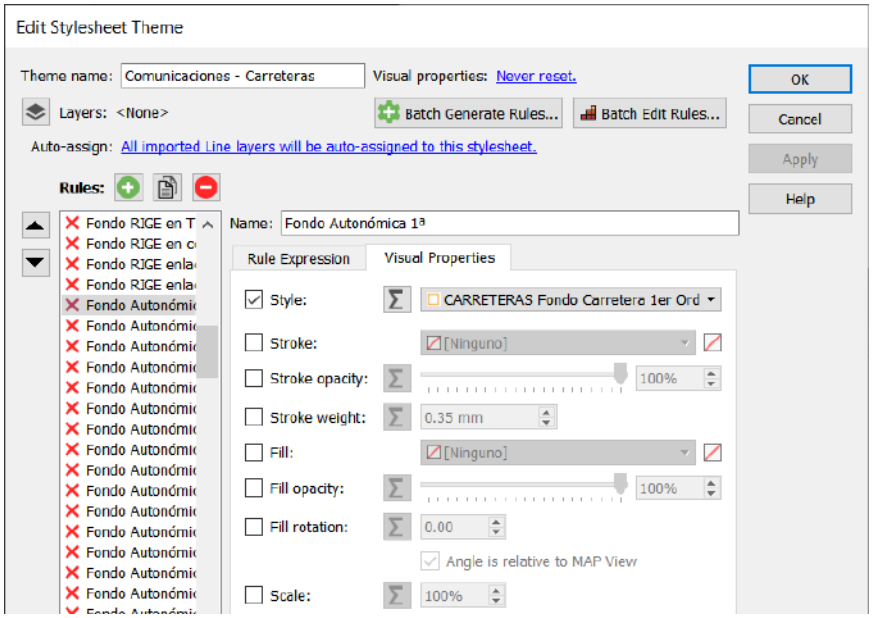


Ilustración 6 - 53: Estilo gráfico correspondiente a la simbolización del Fondo de Autonómica de 1º

Interacción entre de FME y la plantilla de Mapublisher (Maptheme)

Tablas de entidades y su tratamiento con FME

Dado que, para poder simbolizar cada una de las entidades es necesario "traducir" el nncccppe que se mencionó anteriormente a una descripción que Mapublisher "entienda", es imperativo contar con una serie de tablas en las que cada nncccppe corresponda a un atributo descripción que es a través del cual se simboliza.

	a	b	c	d	e	f	g
	B	D	E	F	G	H	L
1	descripcion_TablaPrioridad	Geometria	Letra	PRIORIDAD	NOMBRE_TABLA_MP	AGRUPA_CAPA	NNCCCPPE
2	Salidas carreteras	T	E	030	Salidas carreteras	Salidas Carreteras	03087000
3	Advocaciones de parroquias	T	K	040	Advocaciones de parroquias	Textos Negros	03129000
4	Aeródromos menores y bases aéreas militares	T	K	040	Aerodromos menores y bases aereas militares	Textos Negros	03089000
5	Aeropuertos	T	K	040	Aeropuertos	Textos Negros	03089010
6	Barrios en grandes ciudades	T	K	040	Barrios en grandes ciudades	Textos Negros	03249010
7	Calles en grandes ciudades	T	K	040	Calles en grandes ciudades	Textos Negros	03129010
8	Caminos y vías pecuarias	T	K	040	Caminos y vias pecuarias	Textos Negros	03028000
9	Campos de aviación menores	T	K	040	Campos de aviacion menores	Textos Negros	03089020
10	Capital de Comunidad Autónoma o Provincia me	T	K	040	Capital de Comunidad Autonoma o Provincia menor 200000 hab	Textos Negros	03119020
11	Capital de Estado, Comunidad Autónoma, Capita	T	K	040	Capital de Estado, Comunidad Autonoma, Capital de Provincia mayor 200000 hab	Textos Negros	03059020
12	Capital de municipio de 10000 a 50000 hab	T	K	040	Capital de municipio de 10000 a 50000 hab	Textos Negros	03009020
13	Capital de municipio de 2000 a 10000 hab	T	K	040	Capital de municipio de 2000 a 10000 hab	Textos Negros	03009010
14	Capital de municipio de 500 a 2000 hab	T	K	040	Capital de municipio de 500 a 2000 hab	Textos Negros	03009000
15	Capital de municipio de más de 50000 hab	T	K	040	Capital de municipio de mas de 50000 hab	Textos Negros	03149020
16	Capital de municipio de más de 50000 hab	T	K	040	Capital de municipio de mas de 50000 hab	Textos Negros	03069020
17	Capital de municipio de menos de 500 hab	T	K	040	Capital de municipio de menos de 500 hab	Textos Negros	03069010
18	Comarcas medianas	T	K	040	Comarcas medianas	Textos Negros	03117020
19	Comarcas menores	T	K	040	Comarcas menores	Textos Negros	03189020
20	Construcciones singulares	T	K	040	Construcciones singulares	Textos Negros	03159000
21	Cordilleras, macizos y sierras de 1er orden	T	K	040	Cordilleras, macizos y sierras de 1er orden	Textos Negros	03179020
22	Cordilleras, macizos y sierras de 2º orden	T	K	040	Cordilleras, macizos y sierras de 2º orden	Textos Negros	03209020
23	Cordilleras, macizos y sierras de 3er orden	T	K	040	Cordilleras, macizos y sierras de 3er orden	Textos Negros	03058020
24	Distritos en ciudades grandes	T	K	040	Distritos en ciudades grandes	Textos Negros	03249020
25	Edificios, edificios aislados	T	K	040	Edificios, edificios aislados	Textos Negros	03219000
26	Entidades colectivas: parroquia, concejo, etc	T	K	040	Entidades colectivas: parroquia, concejo, etc	Textos Negros	03099000
27	Entidades colectivas: parroquia, concejo, etc	T	K	040	Entidades colectivas: parroquia, concejo, etc	Textos Negros	03099000
28	Entidades singulares entre 500 y 50 hab: aldeas,	T	K	040	Entidades singulares entre 500 y 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	03039000
29	Entidades singulares mayores de 500 hab: aldea	T	K	040	Entidades singulares mayores de 500 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	03039010
30	Entidades singulares menor de 50 hab: aldeas, b	T	K	040	Entidades singulares menor de 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	
31	Entidades singulares menor de 50 hab: aldeas, b	T	K	040	Entidades singulares menor de 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	03028010
32	Entidades singulares menor de 50 hab: aldeas, b	T	K	040	Entidades singulares menor de 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	03058000
33	Entidades singulares menor de 50 hab: aldeas, b	T	K	040	Entidades singulares menor de 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Textos Negros	03039000
34	Espacios naturales grandes (parque nacional, na	T	K	040	Espacios naturales grandes (parque nacional, natural)	Textos Negros	03237020
35	Espacios naturales restantes (Reserva, monume	T	K	040	Espacios naturales restantes (Reserva, monumento, etc)	Textos Negros	03237010
36	Ferrocarriles	T	K	040	Ferrocarriles	Textos Negros	03209000
37	Hitos kilométricos de carreteras y ferrocarriles	T	K	040	Hitos kilometricos de carreteras y ferrocarriles	Textos Negros	03239000
38	Islas grandes mayores de 50 km2	T	K	040	Islas grandes mayores de 50 km2	Textos Negros	03118020
39	Islas medianas entre 10 y 50 km2	T	K	040	Islas medianas entre 10 y 50 km2	Textos Negros	03238020
40	Islas pequeñas entre 1 y 10 km2	T	K	040	Islas pequenas entre 1 y 10 km2	Textos Negros	03028020
41	Islotes, escollos, farallones, bajos, etc	T	K	040	Islotes, escollos, farallones, bajos, etc	Textos Negros	03238010
42	Islotes, escollos, farallones, bajos, etc	T	K	040	Islotes, escollos, farallones, bajos, etc	Textos Negros	03238000
43	Montañas y picos medianos	T	K	040	Montanas y picos medianos	Textos Negros	03088000
44	Montañas y picos pequeños	T	K	040	Montanas y picos pequenos	Textos Negros	03177010
45	Montañas y picos principales	T	K	040	Montanas y picos principales	Textos Negros	03088010
46	Monumentos y edificios de interés en núcleos de	T	K	040	Monumentos y edificios de interes en nucleos de poblacion	Textos Negros	03249000
47	Naciones limítrofes	T	K	040	Naciones limitrofes	Textos Negros	03059030
48	Nomenclatura de autopistas, autovías, carretera	T	K	040	Nomenclatura de autopistas, autovias, carreteras nacionales y autonomicas de 1er y 2º orde	Textos Negros	03119010
49	Nomenclatura de carreteras autonómicas de 3er	T	K	040	Nomenclatura de carreteras autonomicas de 3er orden, de Diputaciones y otras	Textos Negros	03119000
50	Parajes	T	K	040	Parajes	Textos Negros	03147000
51	Parajes pequeños Huerta, prado, cueva natural,	T	K	040	Parajes pequenos Huerta, prado, cueva natural, mirador natural, espacio, etc	Textos Negros	03147010
52	Playas medianas	T	K	040	Playas medianas	Textos Negros	03208010
53	Playas pequeñas	T	K	040	Playas pequenas	Textos Negros	03208000

Ilustración 6 - 54: Tabla de atributos y entidades.

a	Atributo a través del cual se simboliza.
b	Tipo de geometría. En este caso texto (T).
c	Letra para ordenación de entidades (prioridades).
d	Prioridad para ordenación de entidades.
e	Sin uso.
f	Nombre de la capa.
g	Nivel-Color-Peso-Estilo (nncccppe)

Ilustración 6 - 55: Atributos implicados

La concatenación de los valores de los atributos Letra+Prioridad+AGRUPA_CAPA nos permitirá no sólo nombrar cada uno de los dataset de entidades (en la Ilustración 6 - es el dataset correspondiente a textos) sino ordenar por orden de prioridad. Por ejemplo, K040 Textos Negros.

Por tanto, cuando se cruzan las entidades procedentes de cada una de las tablas contenidas en PostGIS a través del atributo nncccppe éste, corresponderá a una descripción.

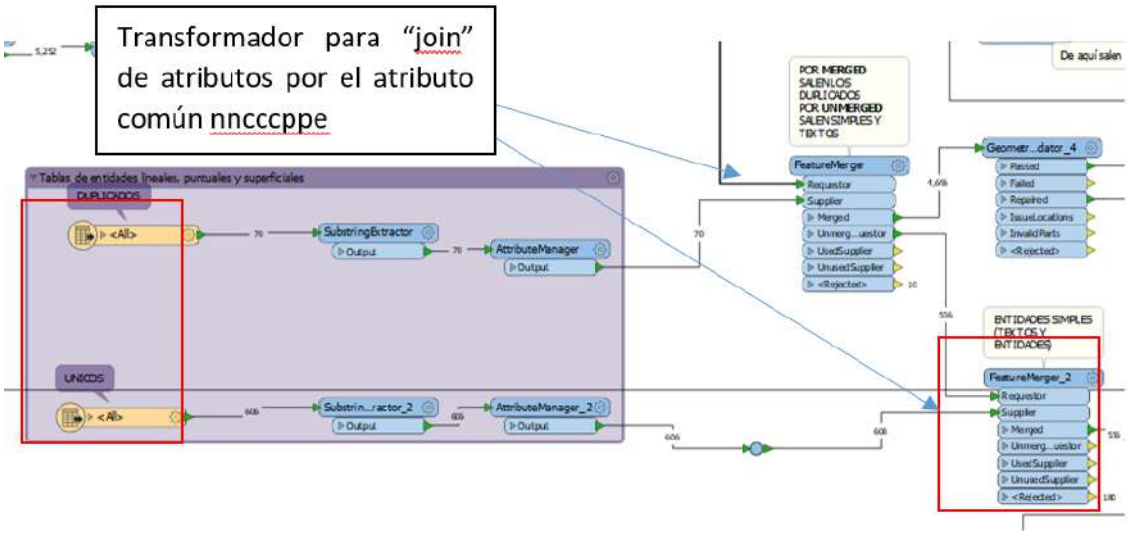


Ilustración 6 - 56: Cruce de entidades geográficas con la tabla en Excel

Dentro de la tabla de Excel, existen dos pestañas, **Duplicados** y **Únicos**. Los duplicados (también elementos triplicados), corresponden a vías de comunicación, concretamente vías de comunicación a excepción de ferrocarriles. En las vías de comunicación como carreteras, caminos, pistas, etc. es necesario hacer una duplicación de entidades para poder simbolizar la caja y el fondo (además el eje en el caso de autovías y autopistas).

Tablas de textos y su tratamiento con FME

Siguiendo con la explicación, existen una serie entidades que, además de la descripción, necesiten de otros tres atributos más, como es el caso de los textos.

Los textos que están en la tabla 08p, poseen una serie de atributos, como es **orientación** y **valor** que definen con qué orientación está el texto (orientación) y su texto correspondiente (valor). Se hace evidente que, para poder hacer la clasificación de los textos, pues no son iguales, es necesario que, de acuerdo con el nncccppe se clasifique cada uno de ellos. Para poder hacer esta clasificación, cada entidad corresponderá a un color (negro, azul o siena) y, además, una fuente y un cuerpo en el texto.

Para ello, también será necesario definir una tabla en la que la descripción se corresponda a una fuente y a un cuerpo dependiendo de la entidad que se esté clasificando.

	a	b	c	d	f
1	Simbología	DESCRIPCION	Fuente	Estilo	Cuerpo
2	(San Pedro)	Advocaciones de parroquias	CaslonOldFace BT	Roman	7.00
3	Aeródromo de la Cañada	Aeródromos menores y bases aéreas militares	Zurich BT	Roman	9.00
4	AEROPUERTO DE MADRID-BARAJAS	Aeropuertos	Zurich BT	Roman	10.00
5	Valleherm oso	Barrios en grandes ciudades	Caslon Opnface BT	Regular	12.50
6	El Pontilo	Barrios en otras ciudades. Pequeños núcleos no censados	Univers 55	Regular	6.50
7	Gran Vía	Calles en grandes ciudades	Zurich BT	Roman	5.50
8	Cañada Real Soperviela	Caminos y vías pecuarias	Zurich Italic BT	Italic	5.50
9	Pista de ultraligeros	Campos de aviación menores	Zurich BT	Roman	7.00
10	TERUEL	Capital de Comunidad Autónoma o Provincia menor 200000 hab	Zurich Cn BT	Bold	15.50
11	MADRID	Capital de Estado, Comunidad Autónoma, Capital de Provincia mayor 200000 hab	Zurich Cn BT	Bold	18.50
12	Linares	Capital de municipio de 10000 a 50000 hab	Zurich Cn BT	Bold	13.50
13	Colmenar de Oreja	Capital de municipio de 2000 a 10000 hab	Zurich Cn BT	Bold	12.00
14	Castropol	Capital de municipio de 500 a 2000 hab	Zurich Cn BT	Bold	10.00
15	CARTAGENA	Capital de municipio de más de 50000 hab	Zurich Cn BT	Bold	13.50
16	Espeluy	Capital de municipio de menos de 500 hab	Zurich Cn BT	Bold	8.50
17	LA ALCARRIA	Comarcas medianas	Bodoni Bk BT	Book Italic	13.50
18	La Vera	Comarcas menores	Bodoni Bk BT	Book Italic	11.00
19	Puerta del Diablo	Construcciones singulares	Zurich Cn BT	Regular	6.00
20	PIRINEOS	Cordilleras, macizos y sierras de 1er orden	Zurich Italic BT	Italic	16.50
21	SIERRA DE SEGURA	Cordilleras, macizos y sierras de 2º orden	Zurich Italic BT	Italic	13.00
22	Serranía de Cuenca	Cordilleras, macizos y sierras de 3er orden	Zurich Italic BT	Italic	11.00
23	CHAMBERÍ	Distritos en ciudades grandes	CaslonOpnface BT	Regular	11.00
24	Casa Pinoja	Edificios, edificios aislados	Zurich Cn BT	Regular	6.00
25	Cuteuro de Rabelos	Entidades colectivas: parroquia, concejo, etc	CaslonOldFace BT	Roman	9.50
26	Rodilana	Entidades singulares entre 500 y 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Univers Condensed	Medium	9.50
27	Algaida	Entidades singulares mayores de 500 hab: aldeas, barrios, lugares, urbanizaciones, etc	Univers Condensed	Medium	11.00
28	Pando	Entidades singulares menor de 50 hab: aldeas, barrios, lugares, urbanizaciones, etc	Univers Condensed	Medium	8.00
29	P. NAC. GARAJONAY	Espacios naturales grandes (parque nacional, natural)	Optima Medium Italic	Regular	15.00
30	Monum. Nat. Los Roques	Espacios naturales restantes (Reserva, monumento, etc)	Optima Medium Italic	Regular	13.00
31	Tren de alta velocidad	Ferrocarriles	Zurich BT	Roman	7.00
32	K 104	Hitos kilométricos de carreteras y ferrocarriles	Swis721 LtCn BT	Light	5.00
33	MALLORCA	Islas grandes mayores de 50 km2	Optima	Italic	15.00
34	ISLA ALEGRANZA	Islas medianas entre 10 y 50 km2	Optima	Italic	10.00
35	Isla del Cierro	Islas pequeñas entre 1 y 10 km2	Optima	Italic	7.00
36	Bajío de la Corva	Islotes, escollos, farallones, bajos, etc	Optima	Italic	6.00
37	Peña de los Gracias	Montañas y picos medianos	Futura Medium Italic	Italic	7.50
38	Pico de la Iglesia	Montañas y picos pequeños	Futura Medium Italic	Italic	6.00
39	Peña Ubiña	Montañas y picos principales	Futura Medium Italic	Italic	9.00
40	FRANCIA	Naciones limítrofes	CaslonOpnface BT	Normal	22.00

Ilustración 6 - 57: Tabla de textos negros

a	Muestra de la fuente y tamaño
b	Atributo DESCRIPCION por el que se simbolizará.
c	Fuente a utilizar.
d	Estilo.
e	Cuerpo (Tamaño de la fuente)

Ilustración 6 - 58: Atributos implicados.

La utilización de tablas para la simbolización de los textos es muy interesante, pues no es necesaria ninguna modificación en la plantilla de Mapublisher.

“Maptheme” de textos en Mapublisher

Para las entidades de texto, en la plantilla de Mapublisher será necesario definir aparte del atributo descripción unos atributos que “llamen” a la fuente, al cuerpo y que se orienten de acuerdo al atributo **orientación**.

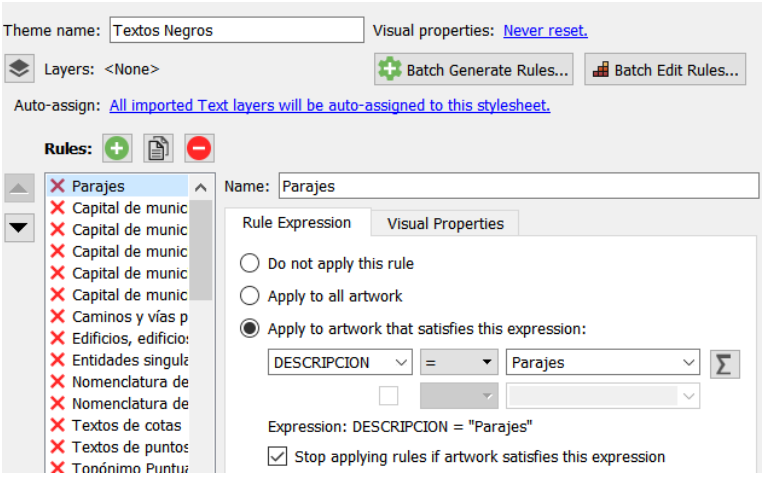


Ilustración 6 - 59: Regla para simbolización

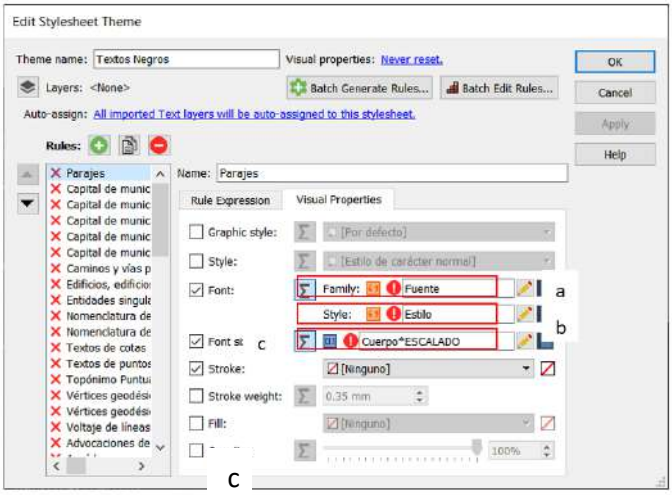


Ilustración 6 - 60: Estilo gráfico correspondiente a la simbolización de sitios

La fuente (*Font)* de texto con sus atributos Familia (a) y Estilo (b), así como el tamaño (c) se leen cuando se ha cruzado la tabla de textos junto con las entidades correspondientes de texto en FME como se ilustra en la siguiente figura:

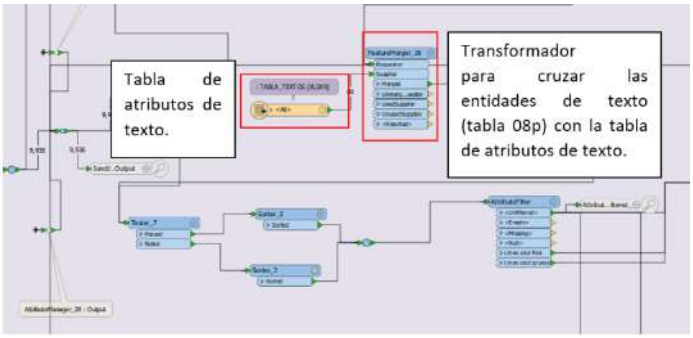


Ilustración 6 - 61: Cruce de entidades de texto con la tabla de atributos de texto

Por otra parte, y con el objetivo de que la georreferenciación en la generación del ráster se haga de forma correcta, se ha de realizar un proceso dentro del flujo de creación de textos que evite que los textos salgan fuera del marco se puede observar en la siguiente figura:

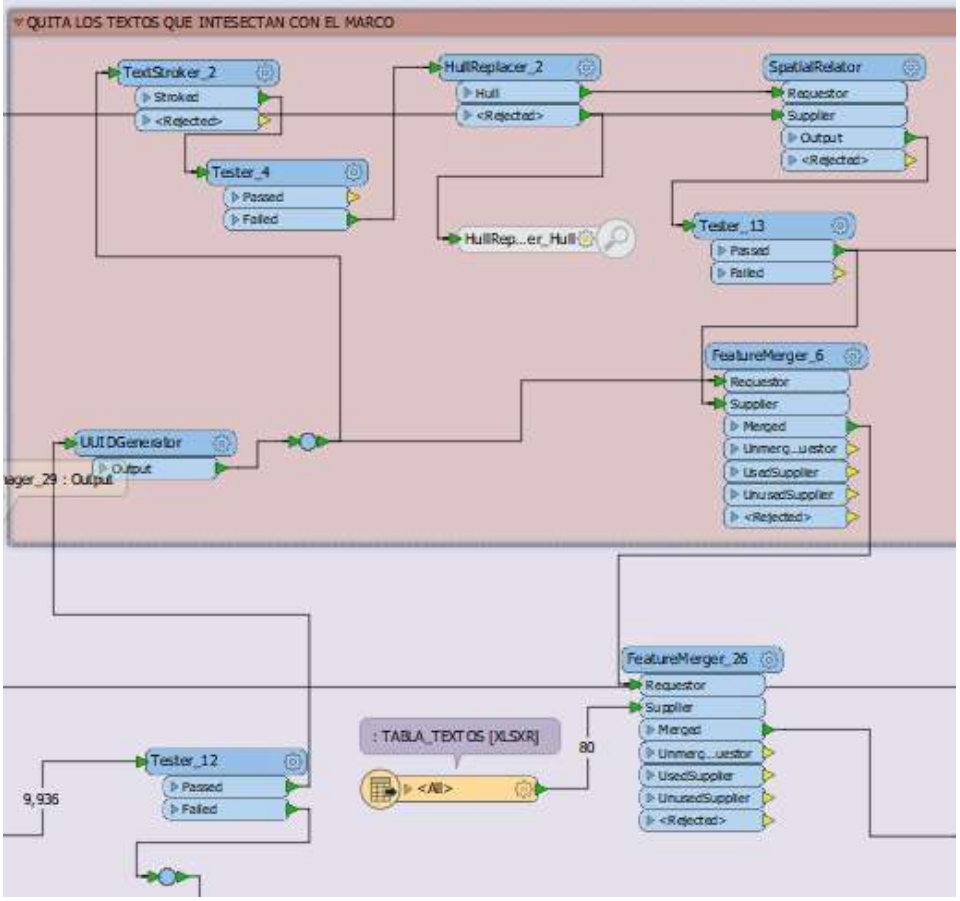


Ilustración 6 - 62: Selección de textos

Se trata de generar el polígono mínimo envolvente de cada texto y comprobar su relación espacial con el marco, de tal manera que, si intersecta con éste, será eliminado.

Una vez explicados todos los elementos intervinientes en el proceso, base de datos PostGIS, tablas y plantilla, pasamos a explicar el proceso automático de simbolización de mapas individuales, esto es, por hojas.

Proceso de simbolización y automatización (parte vectorial)

Esta fase de la producción del MTN25 se divide en dos procesos secuenciales, una primera parte vectorial en la que los mapas se generan en formato geoPDF (vectorial) y otra segunda parte en la que, a partir de estos mapas en geoPDF se generan diferentes salidas del tipo ráster.

En esta sección se describe el proceso automático de simbolización y generación de archivos GeoPDF.

Como el punto de partida son las entidades almacenadas en una Base de Datos Geográfica en PostGIS, lo primero es conectarnos al servidor donde se encuentra almacenada la Base de Datos a través de un “reader” de PostGIS previamente configurado para poder leer todas y cada una de las tablas.

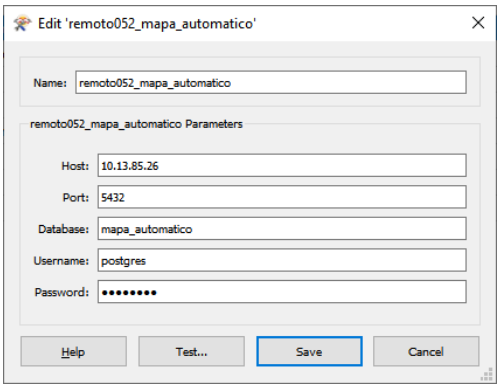


Ilustración 6 - 63: Parámetros de conexión a la Base de datos PostGIS

En la siguiente ilustración, se pueden observar los parámetros del “reader” que nos permite leer los marcos de cada una de las hojas según una sentencia de SQL creada como parámetro público, la cual nos permite elegir a priori, la hoja a procesar como más tarde veremos.

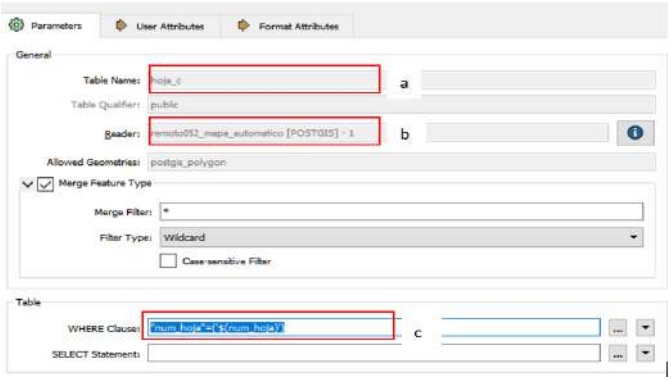


Ilustración 6 - 64: Parámetros del "reader" para leer los marcos de hoja.

a	Nombre de la tabla donde se almacenan las geometrías de los marcos (hoja_c)
b	Conexión predefinida (Ilustración 6 - 17)
c	Cláusula que nos permite elegir a priori la hoja a producir (a través de la creación de un parámetro público (Ilustración 6 - 19)

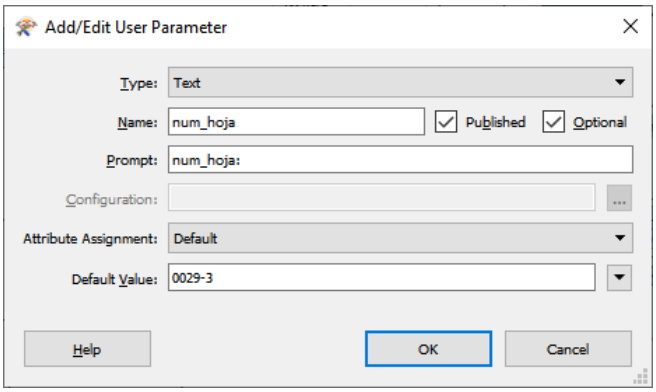


Ilustración 6 - 65: Cuadro de diálogo para la elección del número de hoja

A continuación, se exponen los parámetros del “reader” que lee las tablas de PostGIS:

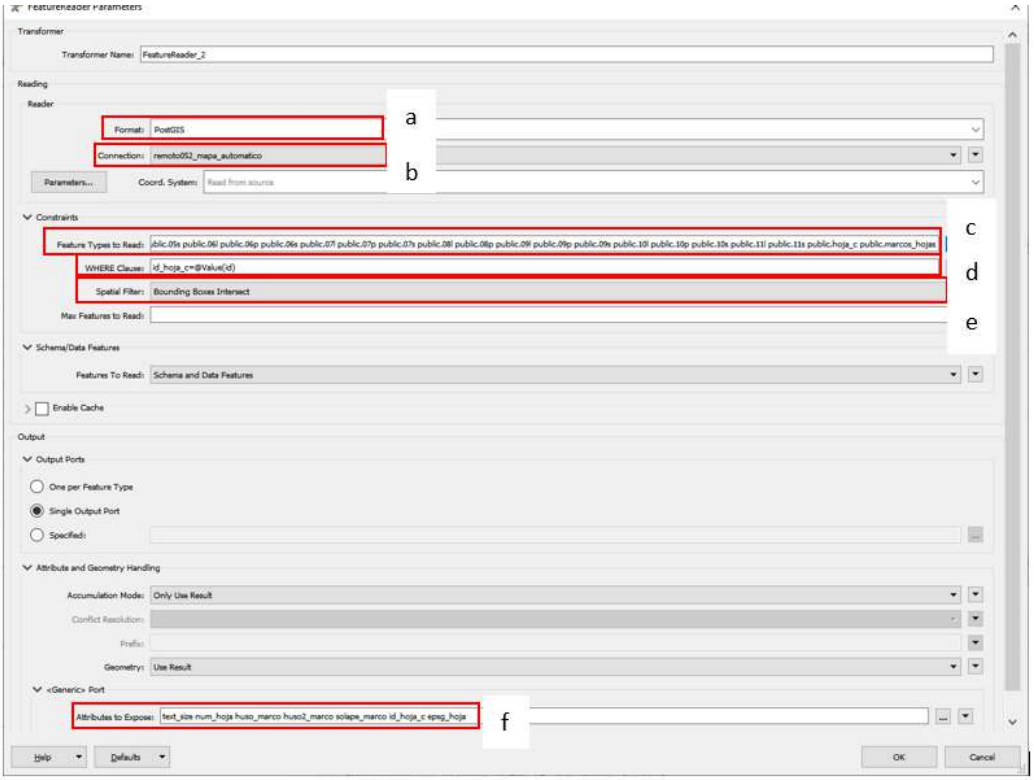


Ilustración 6 - 66: Parámetros del "reader" de tablas de entidades de PostGIS

a	Formato del “reader”. En nuestro caso, PostGIS.
b	Conexión a la Base de Datos con los parámetros de la Ilustración 6 - 17.
c	Tablas a leer de la Base de Datos.
d y e	Sentencia que permite que, a través del atributo id_hoja_c (D) y un filtro espacial (E), se lean las entidades que cumplen que intersectan dentro del rectángulo mínimo envolvente del marco de la hoja correspondiente al que fue elegido a través del parámetro público (Ilustración 6 - 19)
f	Atributos a exponer para cruce de tablas y cálculos como la reproyección si es necesaria.

Debido a que las entidades están almacenadas por hojas, en la ejecución del proceso se pregunta por el número de hoja a ejecutar. A través de una sentencia SQL dentro del “reader” de PostGIS en FME seleccionará la hoja pedida a través del atributo “num_hoja” y comenzará su ejecución llamando al marco correspondiente (Ilustración 6 - 18).

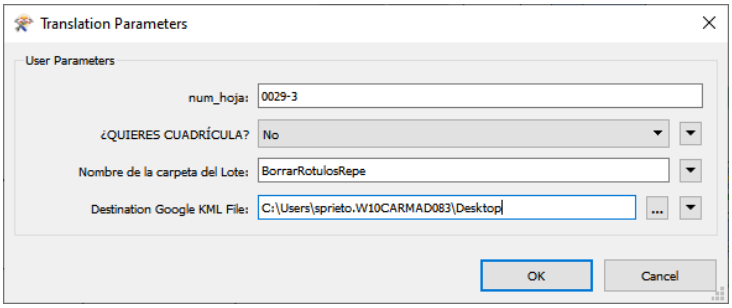


Ilustración 6 - 67: Ejemplo de inicio de ejecución del proceso de simbolización

En un paso inmediatamente posterior dividimos el proceso para que se lean las cuadrículas UTM del huso correspondiente, cortándolas con el marco correspondiente. Mencionar en este apartado que los marcos pueden no ser trapezoidales (en proyección) sino que pueden tener lados con una parte ampliada como se puede ver en la siguiente ilustración:

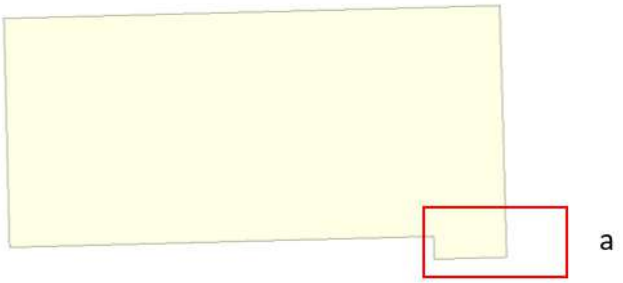


Ilustración 6 - 68: Marco con uno de los lados ampliados

Estas ampliaciones en los marcos se hicieron para evitar, sobre todo en zonas de costa, que la mayoría de la superficie de la hoja fuera "agua" como se puede ver en la siguiente figura:

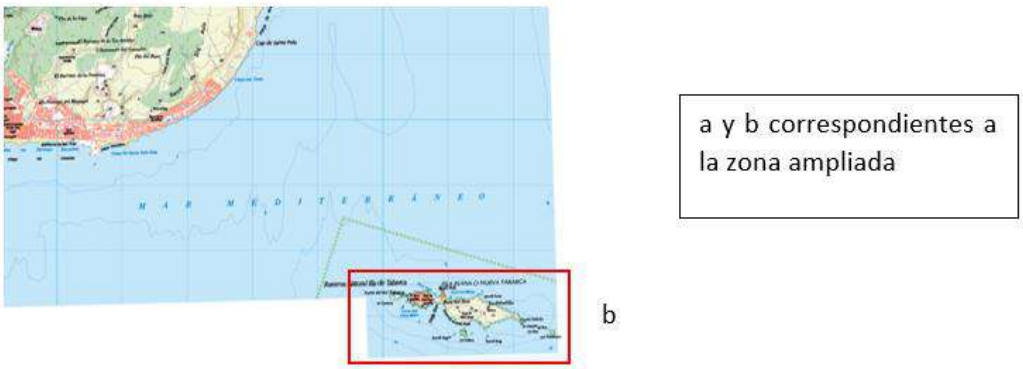


Ilustración 6 - 69: Ejemplo de mapa simbolizado con zona ampliada

A continuación, se ilustra la parte del proceso en la que se genera la cuadrícula que sólo se genera en el caso de que el contenido se destine a edición impresa, no generándose en el caso de que el mapa sea destinado a un servicio de visualización.

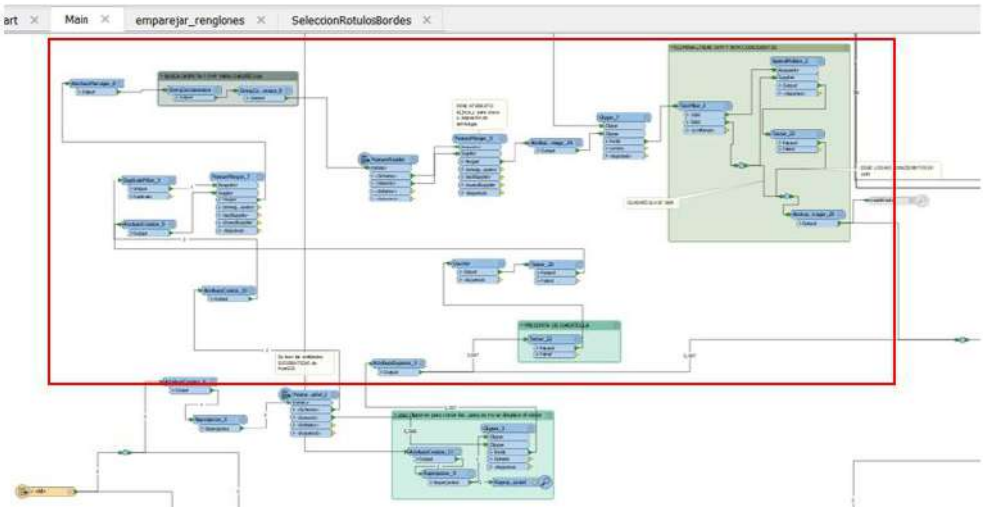
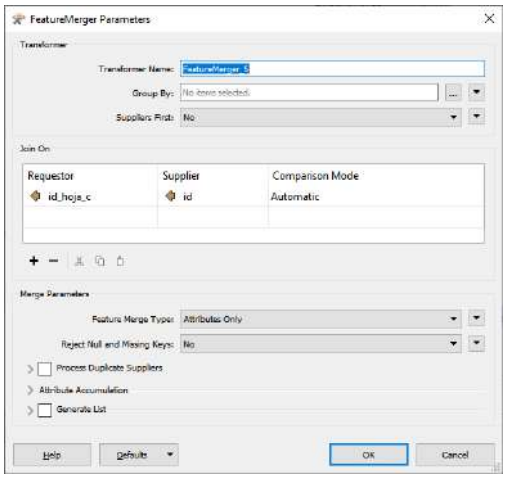


Ilustración 6 - 70: Proceso de generación de la cuadrícula de la proyección UTM

En este proceso se utilizan unos archivos en formato shape de la cuadrícula de toda España. Posteriormente y, a través del atributo correspondiente al número de hoja se extraerá la parte correspondiente de la cuadrícula.

Poco después se hace un cruce de cada una de las tablas con la tabla hoja_c para poder incluir el atributo de convergencia y orientación. Este cruce se hace a través de los atributos comunes de cada tabla (id) y de la tabla hoja_c (id_hoja_c). Cada id pertenece EXCLUSIVAMENTE a un id_hoja_c, es decir, a una hoja en concreto, de ahí que cada id sea clave externa de id_hoja_c como hemos mencionado en la introducción.



Cruce (join) de tablas en la que por el puerto Requestor entran las entidades correspondientes a los marcos y por Supplier los atributos de la tabla correspondiente de tal forma que se incorporen al flujo los atributos de convergencia (converg) y de orientación (orient)

Ilustración 6 - 71: Parámetros para el cruce de la tabla hojac con la tabla correspondiente

	id [PK] integer	id_hoja_c integer	nnccppe character varying (10)	escala double precision	orient double precision	cota double precision	geom geometry
1	142892	3926	33178020	1	0	[null]	0101000020E6640...

Ilustración 6 - 72: Entidad de PostGIS correspondiente a la tabla 05p

descripcion_TablaPrioridad	VALIDO_PRIO	Geometria	Letra	PRIORIDAD	NOMBRE_TABLA_M	AGRUPA_CAPA	NNCCCPPE
Cueva industrial	298	P	K	160	Cueva industrial	Simbolos puntuales construcciones COMPLETO	33179020

Ilustración 6 - 73: Tabla de excel que contiene los campos nnccppe y la descripción necesaria para la simbolización

Este cruce de tablas se realizará entidad a entidad a través del atributo común nnccppe. En el caso de la tabla Excel, es necesario hacer una depuración de nnccppe (atributo nuevo NCPE) debido a que, en algunos casos no cumple la estructura.

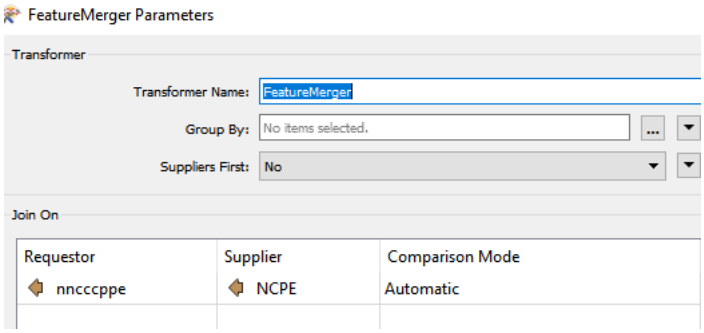


Ilustración 6 - 74: Parámetros para el cruce de tablas de PostGIS y Excel

Por otra parte, y siguiendo el mismo procedimiento de cruce de tablas, obtendremos las descripciones de los cultivos (entidades superficiales), y entidades puntuales, estando en este último caso, además, incluidos los textos ya que el punto es la geometría a partir de la cual se escribe el texto correspondiente.

Especialidad elementos de comunicaciones (Carreteras, pistas, caminos y vías verdes)

Las entidades correspondientes a vías de comunicación (autovías y autopistas, así como carreteras convencionales y de tipo autonómico y local), caminos, pistas y vías verdes, se simbolizan a través de dos (carreteras convencionales y de tipo autonómico y local) o 3 (autopistas y autovías) líneas coincidentes representando cada una de ellas el eje (sólo en autopistas y autovías), la caja y el fondo.

A continuación, se ejemplifica el resultado:

Ilustración 6 - 75: Eje, caja y fondo correspondiente a Autovías y su composición

En la siguiente Ilustración veremos cómo se duplican o, en su caso triplican las líneas coincidentes para obtener el resultado anterior:

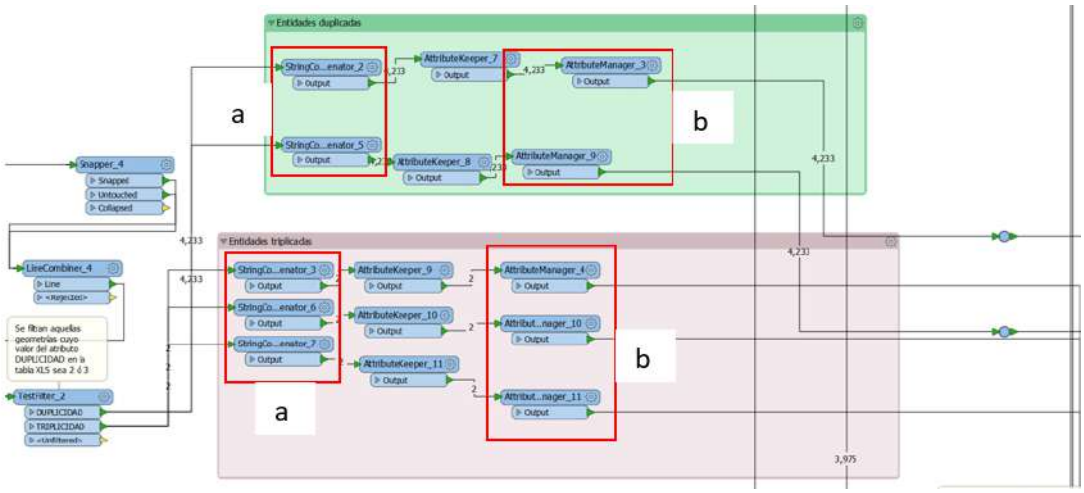


Ilustración 6 - 76: Flujo para generar entidades duplicadas y triplicadas

A	Concatenación de atributos que determinan la prioridad para la simbolización
B	Atribución de valores a los atributos

A continuación, se señalan los parámetros de los transformadores A y B:

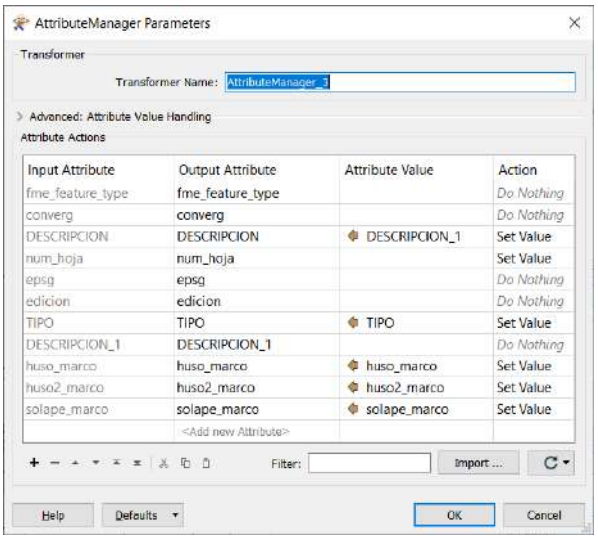


Ilustración 6 - 77: Parámetros Transformador AttributeManager

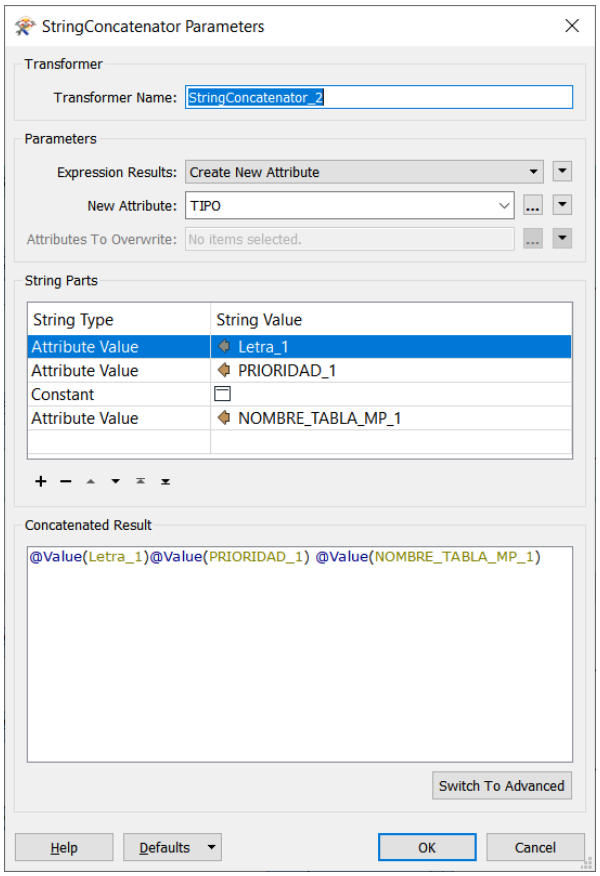


Ilustración 6 - 78: Parámetros Transformador StringConcatenator

En el transformador StringConcatenator se construye el atributo TIPO el cual nos proporcionará el orden de dibujado, como por ejemplo K261 Fondo Autopista peaje ramal en puente.

Manipulación de elementos para el case de hojas

Debido a que el proceso está enfocado, principalmente a suministrar las hojas a un servicio web de visualización de mapas (WMS) de tal manera que exista una continuidad entre hojas y no haya elementos de corte con los marcos, por ejemplo, a la hora de que se cierre un área contra el marco, se hace necesario incorporar un método para eliminar este inconveniente.

Primeramente, exponemos a continuación cómo son los elementos a los que nos referimos a través de una captura de Illustrator en la que se ha definido un estilo, por ejemplo, el de laguna permanente:

Ilustración 6 - 79: Estilo de estanque, piscina.

Ilustración 6 - 80:Trazo y relleno para el estilo de estanque, piscina.

Esta simbología, además de otras como edificios aislados, se compone de un trazo y un relleno, es decir, un área rodeada por una línea. Para poder hacer este proceso es necesario en Mapublisher crear un Maptheme con un estilo lineal que tenga los componentes de línea y área pero que venga de FME mediante una geometría del tipo línea como se puede ver a continuación:

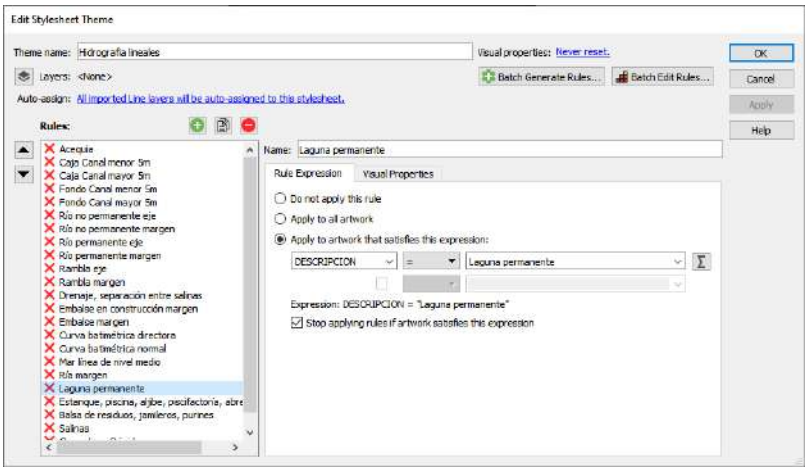


Ilustración 6 - 81: Maptheme de entidades lineales

A continuación, ilustramos el proceso general y pasamos a explicar los transformadores uno por uno:

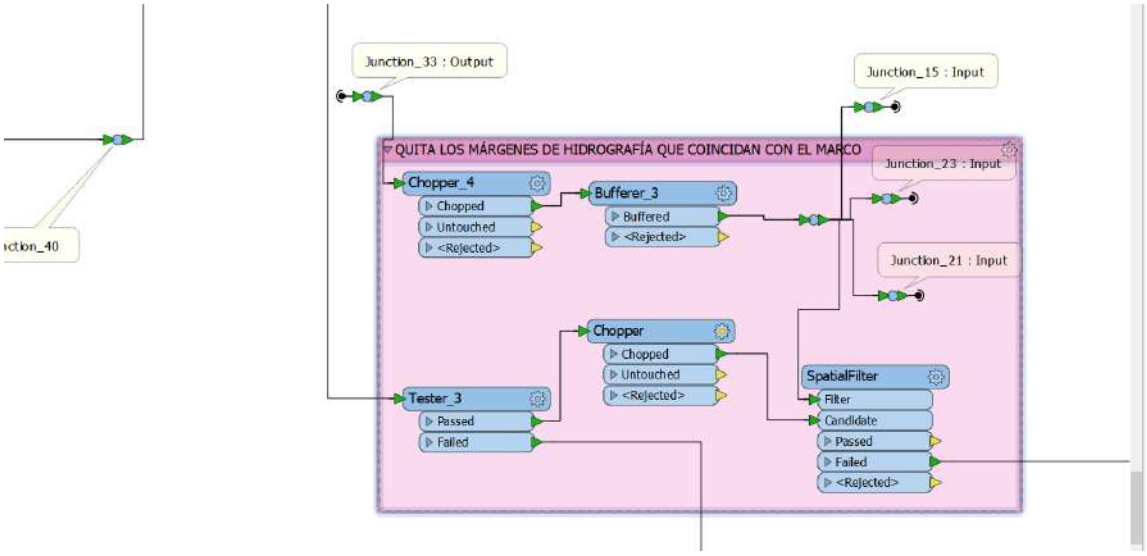


Ilustración 6 - 82: Vista general del proceso de extracción de entidades que tocan el marco.

En FME, hay que identificar, a través del atributo descripción, cuáles son las entidades a las que hay que aplicar este tratamiento según se ilustra en la siguiente imagen:

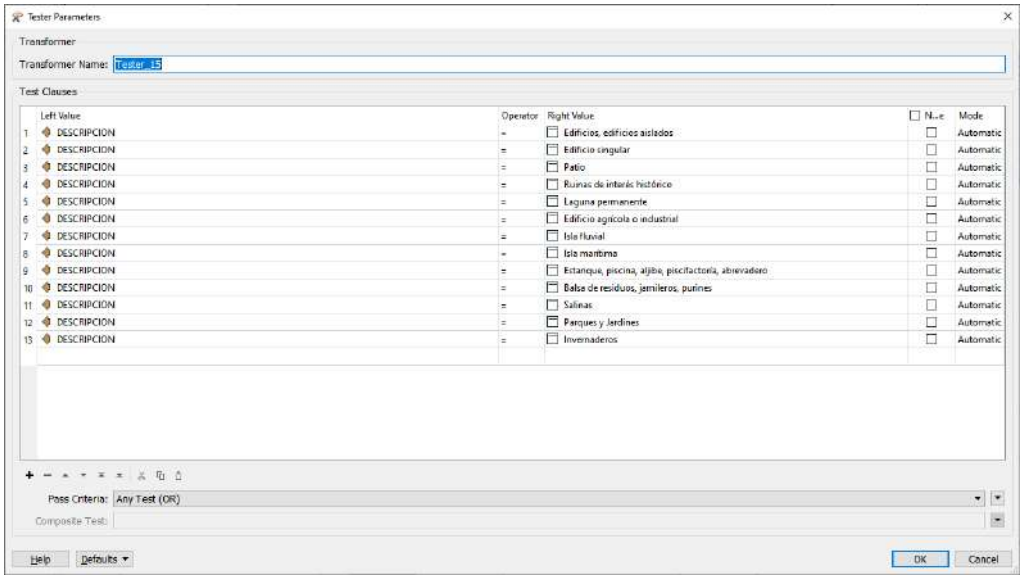


Ilustración 6 - 83: Filtro de entidades que tienen trazo y relleno

Como hemos comentado anteriormente, es necesario que a estas entidades de área se les obligue a ser entidades de línea para que se simbolicen como tales.

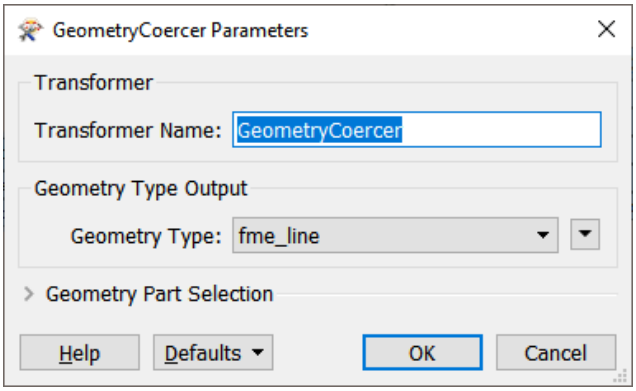


Ilustración 6 - 84: Transformador que obliga a las entidades superficiales a ser líneas.

El siguiente paso es que, debido a que era un polígono, la línea exterior está unida por lo que, para poder extraer la parte que intersecta con el marco, deberemos romperla en segmentos a través del siguiente transformador:

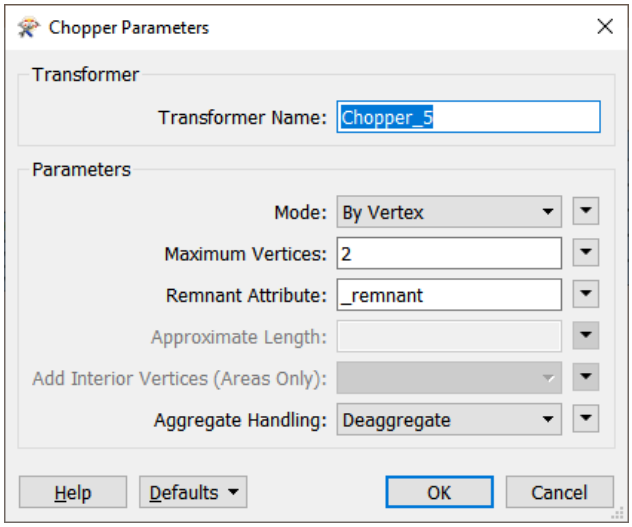


Ilustración 6 - 85: Transformador que "rompe" una línea cerrada en varios segmentos.

Una vez que se han obtenido los segmentos es necesario hacer una intersección con el marco mediante un buffer y, para filtrar los segmentos, ver la relación espacial existente entre los segmentos y el marco. Además, el marco ha de ser segmentado con el transformador chopper y posteriormente un buffer como se expone a continuación:

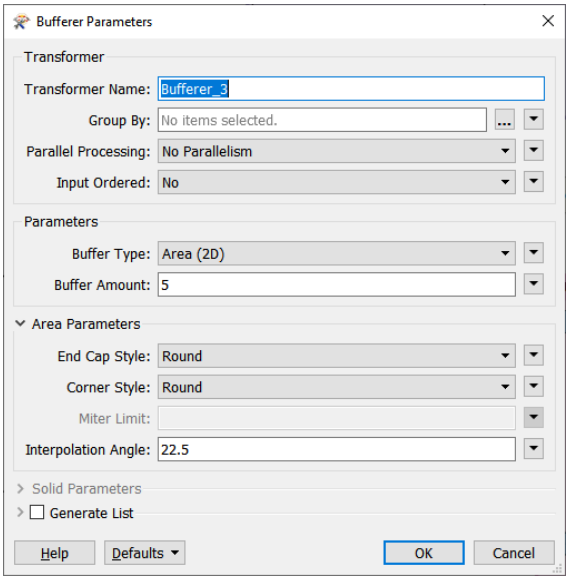


Ilustración 6 - 86: Parámetros buffer de los segmentos del marco.

Una vez que todos los parámetros están definidos, se analiza la relación espacial entre el marco y las líneas correspondientes a la entidad analizada de tal manera que sólo nos quedaremos con los que no intersecten con el marco, siendo los parámetros del transformador SpatialFilter, los siguientes:

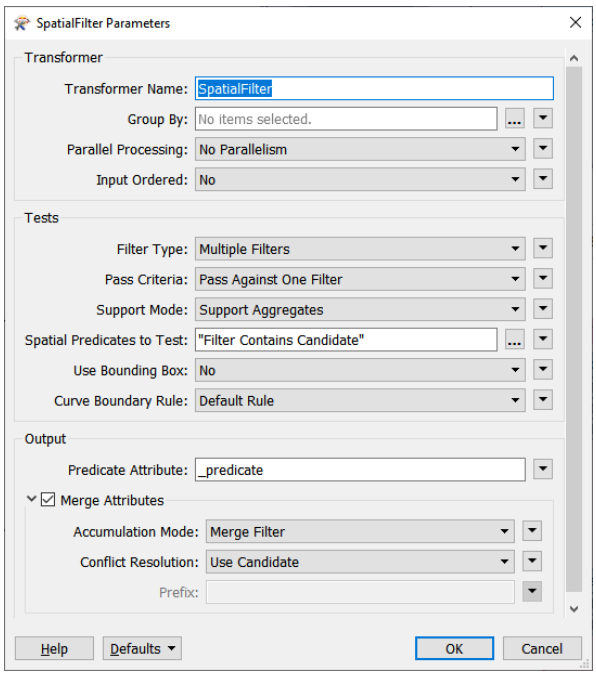


Ilustración 6 - 87: Parámetros para hacer el filtro espacial

Una vez que todos los elementos de la hoja poseen sus descripciones para poder simbolizar, entra en juego un “writer” de FME en el que interviene la plantilla, ejecutando automáticamente el simbolizado y guardando la hoja ejecutada.

EJEMPLO DE WRITER FFS DE FME

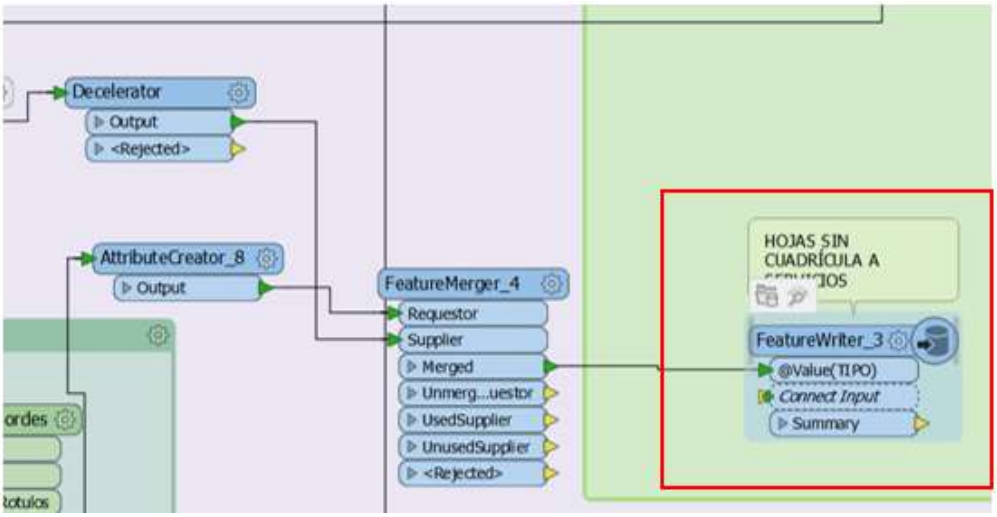


Ilustración 6 - 88: Writer de FME

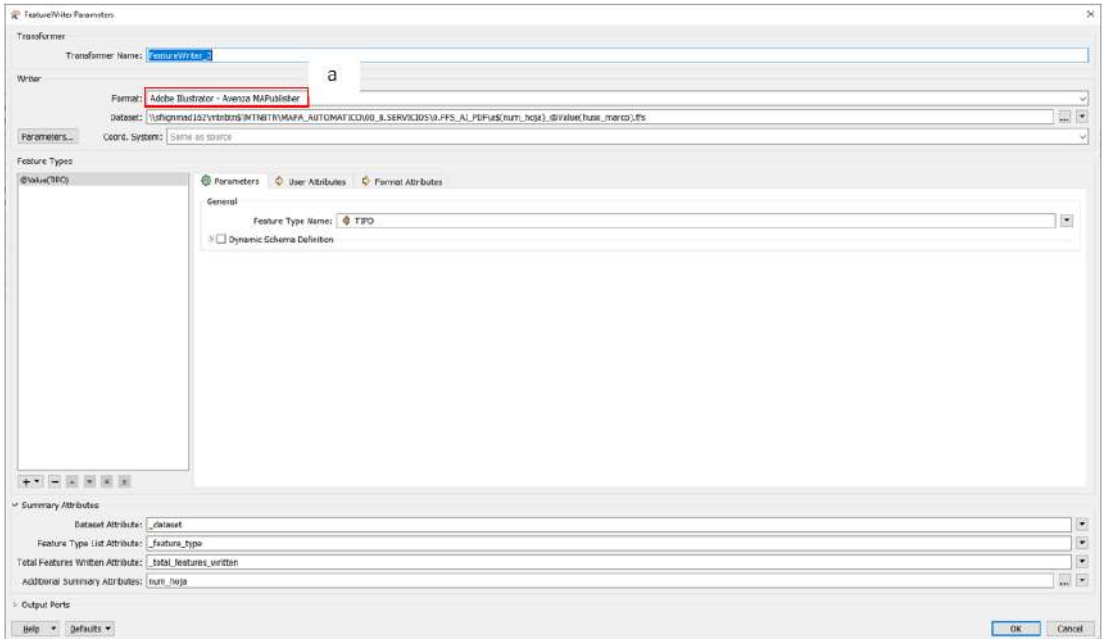


Ilustración 6 - 89: Configuración general del writer

Hay que destacar, como se desprende de la Ilustración 6 - 84, que el formato, a pesar de ser ffs, es necesario que sea específicamente el reseñado en (a).

Por otra parte, y dentro de la configuración general, es necesario definir una serie de parámetros (botón ***Parameters***), que nos permitirán tener éxito en el paso de la simbolización. Estos parámetros tienen la siguiente forma:

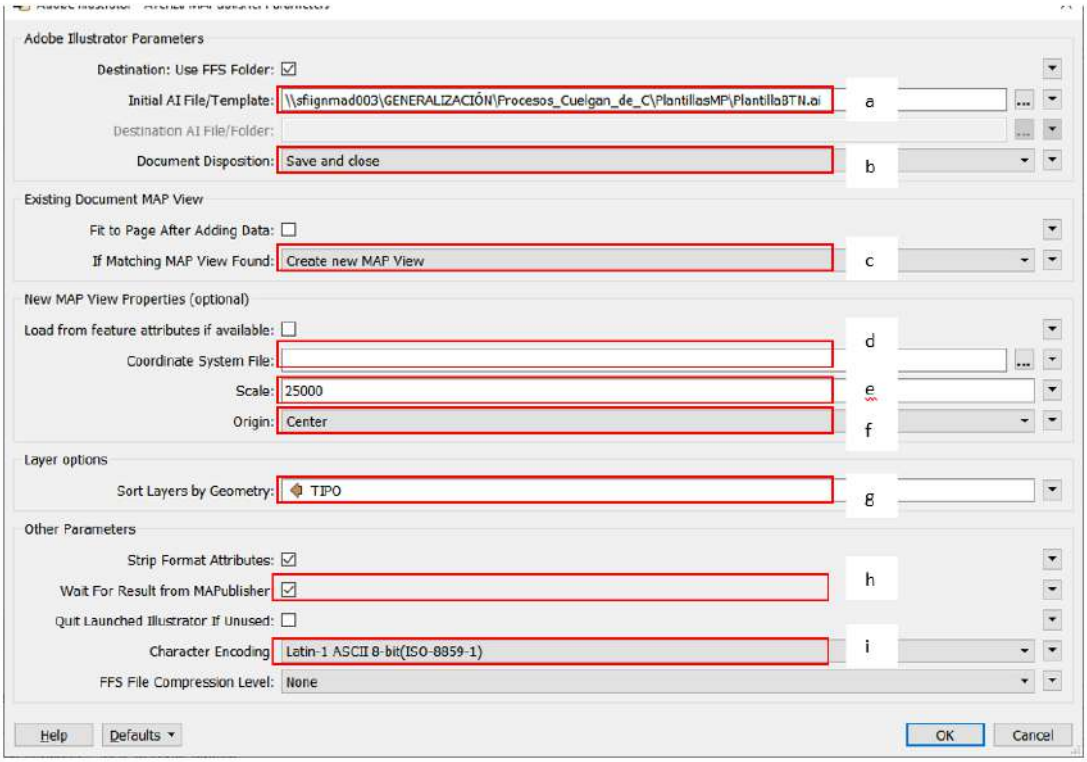


Ilustración 6 - 90: Ajustes del writer de Adobe Illustrator - Avenza MAPublisher

a	Ruta donde se encuentra la plantilla que sirve de base a la simbolización.
b	¿Cómo se ejecuta la simbolización? En este caso se trata de guardar y cerrar para que no existan problemas de caché ni bloqueos por falta de memoria.
c	Cada mapa lo crea como un nuevo MAP View.
d	En nuestro caso, todos los datos están georreferenciados por lo que no tiene sentido adjuntar ningún archivo. En todo caso los formatos admitidos son.wkt, .prj y .tab.
e	Denominador de la escala del mapa. En este caso 25000 (escala 1:25000).
f	Indica desde dónde van a situarse las entidades. En este caso, el centro es lo más adecuado.
g	Este parámetro es muy importante pues nos determina la prioridad en la visualización por el atributo tipo y en orden alfabético (K040 Textos Negros, K065 Textos Azules, etc.).
h	Este parámetro es, quizá, el más importante de todos pues permite que, en un proceso de producción continuo de hojas, FME se quede a la espera hasta que MaPublisher termine de simbolizar la hoja completamente. FME tiene el inconveniente de que cuando entran datos a un transformador, inmediatamente los procesa por lo que se va produciendo un desfase, que si la producción es grande, puede llegar a bloquear todo parándose la producción o no generándose adecuadamente las hojas.
i	Codificación de caracteres.

Para que se pueda llevar a cabo la simbolización, el “writer” de FME genera un archivo .ffs que, a través de la plantilla es simbolizado, guardando otros dos archivos más, uno con extensión .ai y otro con extensión .pdf siendo este último del tipo geoPDF, es decir, un pdf georreferenciado. Añadir que este geoPDF tiene incluida la mesa de trabajo que se utiliza y servirá como base para la generación del ráster como se expone en el siguiente apartado.

Proceso de simbolización y automatización (parte raster)

De acuerdo con las exigencias para servicios web, es necesario generar por cada hoja un archivo ráster en formato geoTIFF así como en formato ECW (Enhanced Compression Wavelet).

Para adecuarlo a la escala 1:25,000, el tamaño del píxel ha de ser de 1,1 metros lo que equivaldría a una resolución aproximada en papel de 600 dpi (dots per inch o puntos por pulgada).

Para ello, también se ha desarrollado un proceso automático que genera un archivo ráster por cada hoja añadiendo el sombreado para dar la sensación de relieve, así como otra serie de mejoras.

Lectura de archivo geoPDF

El proceso automático comienza con la lectura del geoPDF a través de un “reader” de FME en dicho formato. Una de las especialidades del archivo geoPDF es que, al generarse a través de Mapublisher, se encuentra en él un ráster embebido en él con coordenadas en el sistema papel, cuyo origen es la esquina inferior izquierda (coordenadas 0,0) y sin escalar.

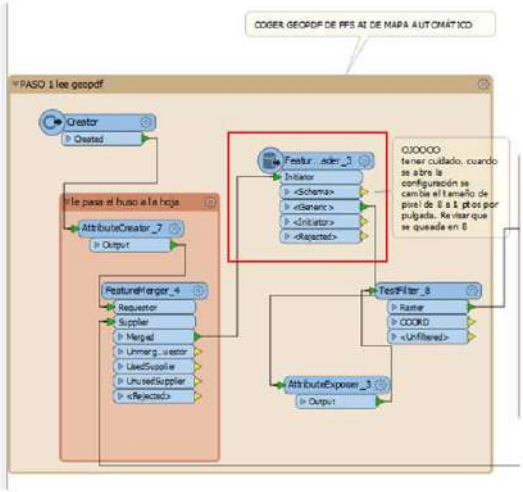


Ilustración 6 - 91: Lector de geoPDF

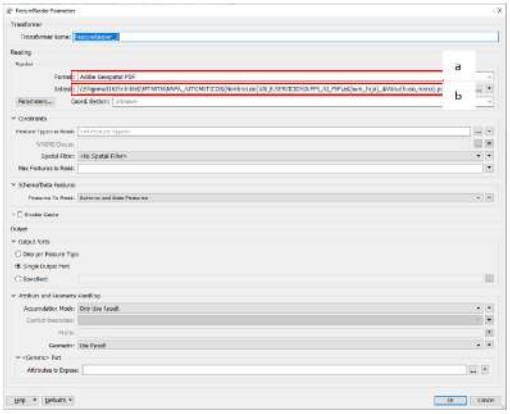


Ilustración 6 - 92: Parámetros del lector de geoPDF

a	Formato Adobe Geospatial PDF.
b	Ruta donde se encuentra el geoPDF generado en el proceso de simbolizado.

Como hemos dicho, el ráster embebido en el geoPDF tiene dimensiones y escala de papel además de no estar georreferenciado.

Georreferenciación del ráster embebido

Para llevar a cabo la georreferenciación del ráster es necesario realizar una serie de operaciones.

Como cada hoja tiene su marco correspondiente, se calculará a través del transformador de FME "Center Point Replacer" el centro del marco, el cual, con sus coordenadas correspondientes nos servirá para georreferenciar.

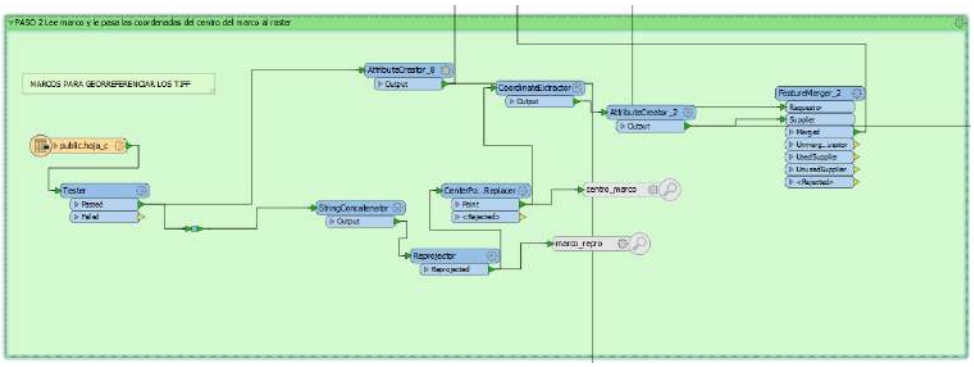


Ilustración 6 - 93: Cálculo del centro del marco que nos servirá para georreferenciar el ráster

Por otra parte, también es necesario calcular el centro del ráster embebido para que, por medio del transformador "offsetter" (a) nos podamos llevar las coordenadas del centro del ráster a las coordenadas del centro del marco las cuales están en su verdadera posición.

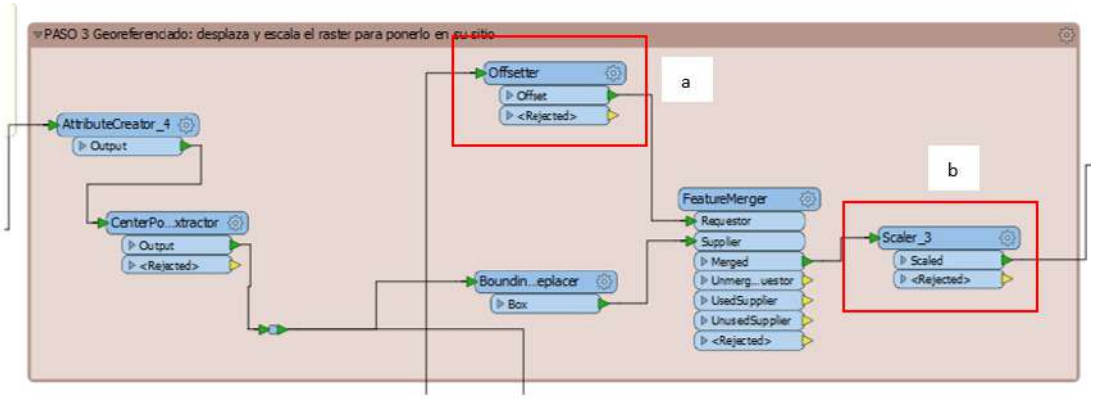


Ilustración 6 - 94: Escalado y desplazamiento del ráster embebido a su verdadera posición

EJEMPLO DE GEORREFERENCIACIÓN FME

Escalado del raster (b)

Como hemos dicho anteriormente, el ráster embebido no tiene las dimensiones reales de la hoja concretamente de la mesa de trabajo que también está incluida. Por tanto, necesitamos calcular el rectángulo envolvente del ráster embebido y que habrá que relacionar con las dimensiones de la mesa de trabajo a escala 1:25000. Hay que recordar que son 900 mm. X 550 mm. que a escala 1:25000 corresponden a 22.500 m. X 13.750 m.

Calcular el rectángulo envolvente del ráster embebido se realiza a través del transformador de FME "BoundingBoxReplacer" de tal manera que el largo y ancho de este rectángulo envolvente se almacene en dos atributos que son "CORTO_RASTER" y "LARGO_RASTER".

A través del transformador de FME "Scaler" daremos escala a través del cociente entre 22.500/"LARGO RASTER" y 13.750/"CORTO_RASTER" de la forma que se ilustra a continuación:

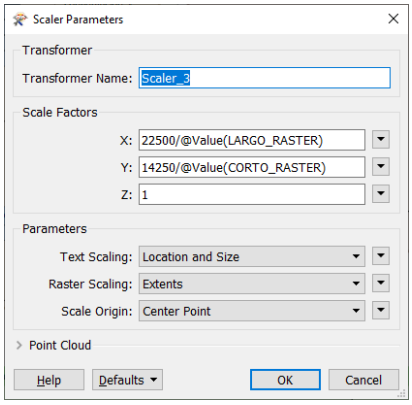


Ilustración 6 - 95: Relación de la dimensión de la mesa de trabajo en dimensiones papel y los lados de la mesa de trabajo en dimensiones terreno, es decir, a escala 1:25000

Es importante, asignar (no se georreferencia, sino que se le indica qué sistema de coordenadas tiene) mediante el transformador de FME "CoordinateSystemSetter", a través del atributo que está en la tabla hoja_c de PostGIS el sistema de coordenadas (ETRS89) así como en la proyección UTM el huso correspondiente a la hoja en cuestión (28, 29, 30 ó 31).

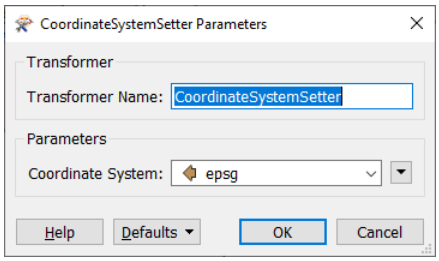


*Ilustración 6 - 96: Parámetro de Sistema de Coordenadas a través del valor del atributo **epsg.***

Ordenación de bandas para RGB

El ráster al que nos venimos refiriendo, está compuesto por 4 bandas cuyos índices son 0, 1, 2 y 3, correspondiendo al índice 0 el azul, el 1 correspondiente al verde, el 2 correspondiente al rojo y el 3 correspondiente a la banda Alpha que se utiliza como banda de transparencia de la que hablaremos un poco más adelante.

Como las bandas no entran ordenadas para que se pueda ver en color verdadero, esto es en RGB, tenemos que ordenarlas mediante el transformador de FME "RasterBandOrderer" cuya ordenación, por tanto es, 2 1 0 3.

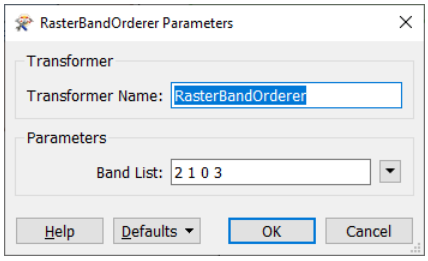


Ilustración 6 - 97: Transformador para ordenar las bandas y obtener el mapa en color verdadero (RGB).

Procedimiento de sombreado

Una vez que se ha completado este último paso, se pasa a sombrear la imagen que, todavía no se ha escrito como archivo, sino que está en memoria.

Este ráster que está en memoria, se va a “mezclar” con un Modelo Digital del Terreno (MDT). A partir de cada MDT de cada zona y, mediante un tratamiento digital en Photoshop y en Global Mapper, se crea un único archivo de sombreado para todo el territorio nacional. Hay que reseñar que a partir de este archivo se extraen todos los MDT ya sombreados, uno por cada hoja del Mapa 1:25,000. La resolución de cada MDT es, al igual que el ráster del mapa, de 1,1 m. ya que es necesario que ambos ráster (MDT y MapaRáster) tengan la misma resolución para poder mezclarse.

El ráster del MDT, debido al paso por Photoshop y Global Mapper, sus bandas vienen en forma de paleta, es decir, para cada píxel existen 3 componentes de color pero sin dividir en bandas, con lo que a través del transformador “RasterPaletteResolver” podemos descomponer cada píxel en las componentes para cada banda.

Este raster del MDE, por tanto, está compuesto por 3 bandas. Destacar que, en el caso del sombreado, cada píxel tendrá el mismo valor en sus bandas ya que el proceso de generación es en escala de grises.

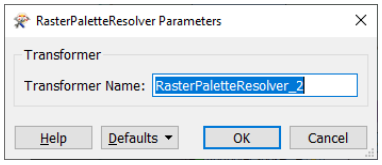


Ilustración 6 - 98: RasterPaletteResolver

Es necesario, en consecuencia, que una cuarta banda sea añadida se puedan mezclar el ráster del mapa con el ráster del MDT. Esta operación se hará a través del transformador “RasterBandAdder”. Además, puesto que cada banda puede adoptar 256 valores distintos (desde 0 a 255, esto es, 2^8 valores) es necesario “obligar” a ambos ráster a tener la misma interpretación de bandas, en este caso RGBA32.

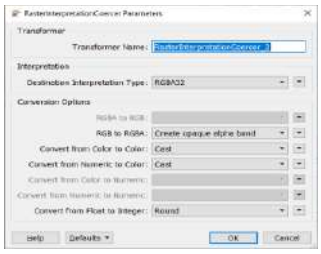


Ilustración 6 - 99: Parámetros que "obligan" a una cierta interpretación de las bandas del ráster.

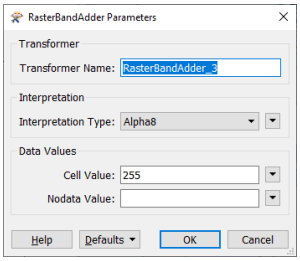


Ilustración 6 - 100: Transformador para añadir una banda con su valor para todos los píxeles.

Una vez que ambos ráster (Mapa y MDT) tienen las mismas características, esto es, misma resolución geométrica (tamaño de píxel), mismo número de bandas y el mismo número de posibles valores para cada píxel se está en condiciones para poder mezclar ambos.

El proceso de sombreado se va a realizar a través de álgebra ráster utilizando funciones de vecindad local, es decir, píxel a píxel; de ahí que tengan que estar armonizados los parámetros que hemos expuesto en el anterior párrafo. La operación del álgebra ráster a realizar es la de multiplicación de los valores de cada píxel para cada una de las bandas. El mapa resultante al aplicar el sombreado, tendrá los valores de cada píxel más bajos que antes de aplicar el sombreado.

Para poder realizar esta operación de álgebra ráster, es necesario utilizar el transformador "RasterExpressionEvaluator" el cual tiene como puertos de entrada el ráster del mapa y el sombreado.

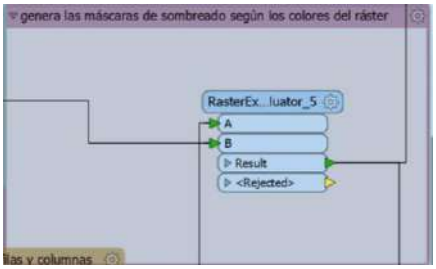


Ilustración 6 - 101: Transformador RasterExpressionEvaluator.

En la ilustración 6-97, por el **Puerto A** entra el ráster de sombreado y por el **Puerto B** el Mapa Ráster.

Generación de máscaras en entidades donde no se vaya a aplicar el sombreado

Por cuestiones de estética y de lógica (no hay relieve en edificios y por tanto sombreado, por ejemplo), hay que hacer una consideración importante a la hora de realizar el sombreado del mapa y es que ciertas entidades de las que hablaremos a continuación es necesario que sean opacas para que no se mezcle con el sombreado y esto se consigue a través del álgebra ráster.

¿Qué entidades vamos a considerar que no se tienen que sombrear?

Las entidades que no se van a mezclar con el sombreado son:

- Fondos de todas las vías de comunicación.
- Zonas de reserva planimétrica en zonas urbanas y algunos cultivos como arrozales, zonas pantanosas, marismas.
- Edificaciones de cualquier naturaleza.

Para poder discriminar qué entidades van a ser opacas, hay que extraer las componentes de color de cada banda (por ejemplo 255, 255, 255 que correspondería al color blanco) y para cada una de ellas establecer la condición correspondiente mediante operadores lógicos

La cuarta banda (Alpha) nos permitirá en el ráster del mapa y, mediante los operadores lógicos, establecer como opaco (valor de la banda alpha 0) si se cumple que las componentes de color han de ser opacas (por ejemplo, para vías de comunicación) o bien, si no tiene esas componentes predefinidas, establecer la transparencia (valor de la banda alpha 255).

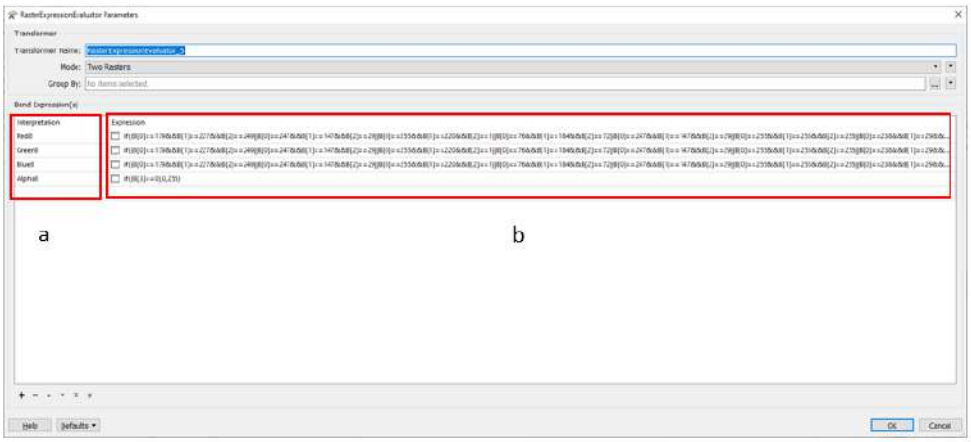


Ilustración 6 - 102: Configuración del transformador RasterExpressionEvaluator

a	Banda a evaluar.
b	Expresión a utilizar para hacer opaca las entidades correspondientes según sus componentes RGB de color.

La expresión se realiza a través de un condicional cuya forma es la siguiente:

If(B[0]==0&&B[1]==126&&B[2]==197||B[0]==0&&B[1]==174&&B[2]==239),B[0],B[0]*A[0]/255)

En la que los operadores lógicos "&&" significan AND y "||" significa OR, de tal manera que esta expresión se leería así: "Si las componentes 0, 1 y 2 OR 0, 1 y 2 (esto significa que pueden ser cualquiera de ellas) tienen unos determinados valores entonces que se mantengan cada una de las componentes opacas y si no se cumple quiere decir que son entidades que hay que mezclar, por lo que es cuando se multiplican ambos valores y se divide por 255".

Es importante destacar que algunas zonas que sí que han de ser sombreadas, poseen las mismas componentes de color que zonas que no han de ser sombreadas. La solución que se ha encontrado es variar mínimamente las componentes de color de tal manera que, visualmente no haya diferencia pero que en el transformador no haya coincidencia con las componentes de color impuestas para ciertas entidades. Esta variación de color se ha realizado en la plantilla de Mapublisher.

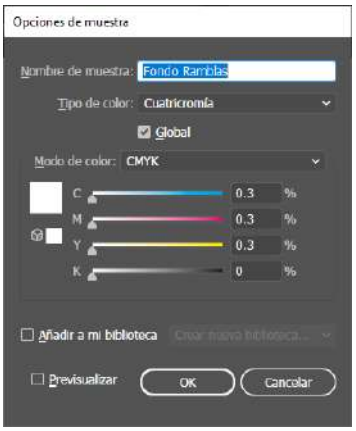


Ilustración 6 - 103: Cambio de las componentes de color del fondo de las ramblas.

El fondo de las ramblas debería mezclarse, pero al ser blanco, hay que variar mínimamente las componentes, de la forma que se puede apreciar en la ilustración 6-99, para que el sombreado se pueda mezclar con el Mapa Ráster.

Una vez que se ha hecho el tratamiento se pasa a almacenarlo en un archivo del tipo geoTIFF, COG y ECW, formatos cuya utilización corresponde, respectivamente a servicios web en los dos primeros casos y descargas en los tres casos (geoTIFF, COG y ECW).

Proceso integrado: geoPDF más ráster

Para que la producción de las hojas sea eficaz, se ha ideado un sistema que una, por una parte, la producción de las hojas en geoPDF y, por otra, la extracción y generación del archivo ráster a partir del primero.

Para integrar estos dos procesos en FME hay que llamar de forma independiente a los "workbench" que se encargan de cada una de las partes; esto se hace a través del transformador WorspaceRunner, como se ilustra a continuación:

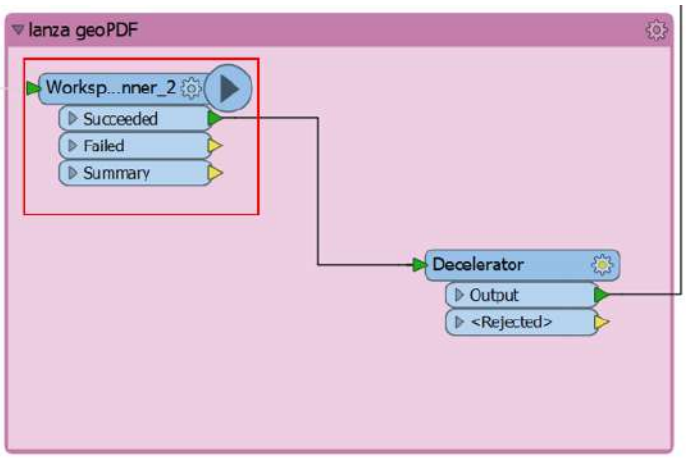


Ilustración 6 - 104: Transformador WorkspaceRunner.

En este caso, este transformador llamaría al workbench que simboliza, es decir, el que genera los geoPDF.

Ahora, veremos más en detalle los parámetros importantes que permiten la automatización de la producción:

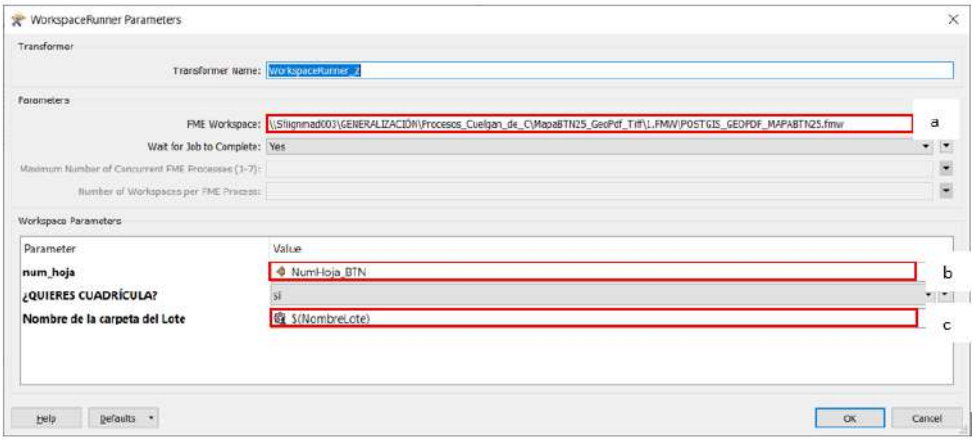


Ilustración 6 - 105: Parámetros del WorkspaceRunner para la simbolización y generación de archivos geoPDF.

a	Ruta donde se encuentra el "workbench" para simbolizar y generar el geoPDF.
b	Valor del atributo. num_hoja es el número de hoja que se va a ejecutar.
c	Parámetro público que nos permite nombrar el lote correspondiente a la producción.

Los números de hoja que entran a través del atributo “num_hoja”, se pueden leer desde una tabla de Excel, por ejemplo, de la siguiente forma:

	A	B	C	D	E	F	G	H	I	J	K	L
1		NUMERO										
2		0686-4	Toledo				987-2	0686-4	oct 17, 2020 11:33:02 AM	oct 17, 2020 11:47:52 AM	0:14:50	VERDADERO
3		0686-1	Toledo				987-2	0686-1	oct 17, 2020 10:15:48 AM	oct 17, 2020 10:29:06 AM	0:13:18	VERDADERO
4		0686-2	Toledo				987-2	0686-2	oct 17, 2020 10:02:50 AM	oct 17, 2020 10:15:01 AM	0:12:11	VERDADERO
5		0686-3	Toledo				987-2	0686-3	oct 17, 2020 9:49:55 AM	oct 17, 2020 10:01:59 AM	0:12:04	VERDADERO
6		0686-4	Toledo				987-2	0686-4	oct 17, 2020 9:37:43 AM	oct 17, 2020 9:49:11 AM	0:11:28	VERDADERO
7		0819-1	Albacete				987-2	0819-1	oct 19, 2020 7:16:01 PM	oct 19, 2020 7:28:19 PM	0:12:18	VERDADERO
8		0819-2	Albacete				987-2	0819-2	oct 19, 2020 6:57:15 PM	oct 19, 2020 7:14:37 PM	0:17:22	VERDADERO
9		0819-3	Albacete				987-2	0819-3	oct 19, 2020 6:43:11 PM	oct 19, 2020 6:56:16 PM	0:13:05	VERDADERO
10		0819-4	Albacete				987-2	0819-4	oct 19, 2020 6:22:01 PM	oct 19, 2020 6:41:38 PM	0:19:37	VERDADERO
11		0834-2	Ciudad Real				987-2	0834-2	oct 19, 2020 7:20:13 PM	oct 19, 2020 7:42:41 PM	0:13:28	VERDADERO
12		0835-1	Ciudad Real				987-2	0835-1	oct 19, 2020 8:28:09 PM	oct 19, 2020 8:41:51 PM	0:13:42	VERDADERO
13		0835-2	Ciudad Real				987-2	0835-2	oct 19, 2020 8:12:27 PM	oct 19, 2020 8:27:03 PM	0:14:36	VERDADERO
14		0835-3	Ciudad Real				987-2	0835-3	oct 19, 2020 7:57:38 PM	oct 19, 2020 8:11:30 PM	0:13:52	VERDADERO
15		0835-4	Ciudad Real				987-2	0835-4	oct 19, 2020 7:43:40 PM	oct 19, 2020 7:56:44 PM	0:13:04	VERDADERO

Ilustración 6 - 106: Tabla correspondiente a las hojas que van a entrar en un determinado lote.

Una vez que se ha completado la generación del archivo geoPDF, es importante tener en cuenta el tiempo de generación de cada una de las hojas puesto que, como habíamos visto anteriormente, es necesario que el check “esperar respuesta de Adobe Illustrator” esté activado. Aun teniendo esta precaución, es importante tener en cuenta que, a pesar de tener activado este check y que sólo afecta a la generación del archivo de Adobe Illustrator (.ai), de la creación del archivo geoPDF se encarga Mapublisher quedando fuera del control de FME y pudiendo ser una fuente de problemas en producciones de grandes cantidades de hojas. Esta parte es muy importante puesto que, si no se ha generado totalmente el archivo geoPDF, el rasterizado no tendrá éxito.

Para ello, acudimos a un transformador llamado “Decelerator” el cual permite demorar durante el tiempo que le digamos (en segundos) la entrada en el siguiente WorkspaceRunner que es el que llama al workbench del rasterizado y que exponemos en la siguiente ilustración:

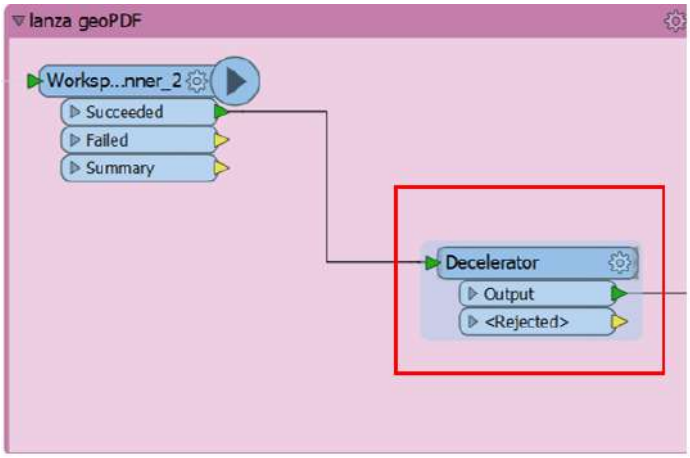


Ilustración 6 - 107: Transformador Decelarator.

Los parámetros de este transformador se visualizan en la siguiente ilustración:

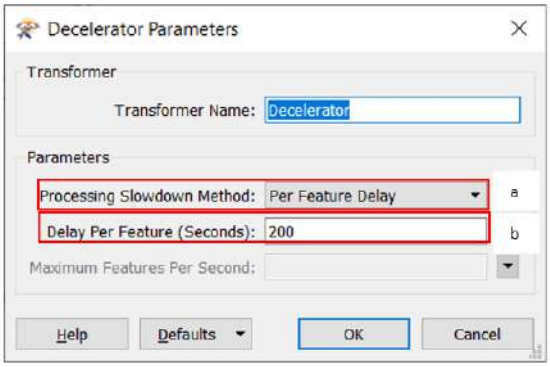


Ilustración 6 - 108: Parámetros del transformador Decelerator.

a	Forma de demorar la salida.
b	Demora en la salida en segundos, en este caso, 200 segundos.

Otra forma de demorar podría ser mediante el transformador "Sorter", el cual, hasta que no deja de entrarle órdenes no actúa:

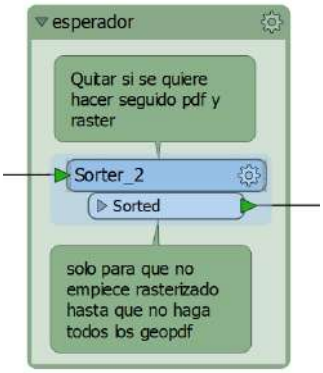


Ilustración 6 - 109: "Esperador" con un sorter.

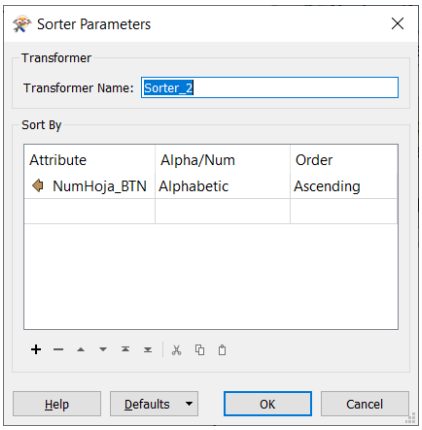


Ilustración 6 - 110: Parámetros del transformador "Sorter".

En este transformador, el orden es por número de hoja, de forma alfabética y en orden ascendente.

La salida de este transformador será la entrada en el WorkspaceRunner que llama al rasterizado de las hojas:

Ilustración 6 - 111:WorkspaceRunner que llama al workbench de rasterizado.

Los parámetros de este WorkspaceRunner se ilustran a continuación:

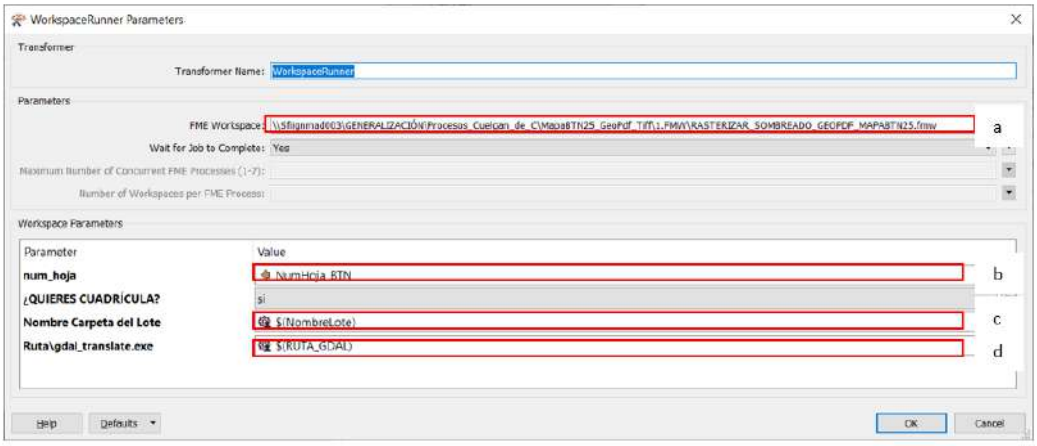


Ilustración 6 - 112: Parámetros del WorkspaceRunner de generación del rasterizado.

a	Ruta donde se encuentra el "workbench" para generar el archivo ráster a partir del geoPDF..
b	Valor del atributo. num_hoja es el número de hoja a través del cual va a buscar el archivo geoPDF para generar el ráster.
c	Parámetro público que nos permite nombrar el lote correspondiente a la producción.
d	Ruta donde se encuentra GDAL para la generación de los archivos COG (Cloud Optimized GeoTIFF).

A continuación, y para finalizar, se ilustra el proceso completo de producción automática:

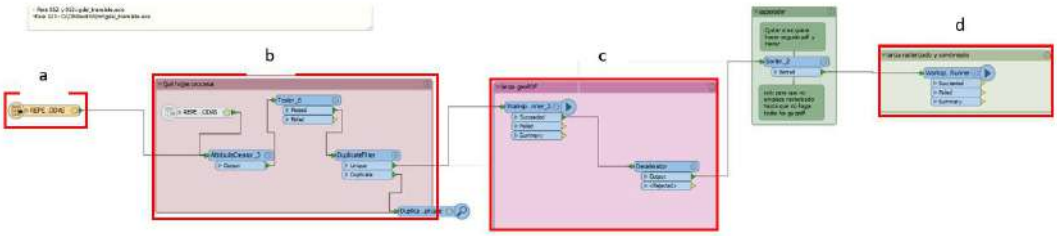


Ilustración 6 - 113: Proceso completo de automatización para la producción de geoPDF y su rasterizado.

a	Archivo Excel donde se encuentran los números de hoja a ejecutar.
b	Análisis de hojas duplicadas posibles.
c	WorkspaceRunner de generación de los archivos geoPDF.
d	WorkspaceRunner de generación de los archivos ráster a partir de los geoPDF.

6.4. Publicación de las diferentes salidas

6.4.1. Publicación impresa

Para propiciar la publicación impresa de las hojas de cartografía se utiliza el formato PDF/X, que permite conservar las características de formato, fuentes/tipos de letra, imágenes y gráficos de un archivo original independientemente del software con el que fue creado (en el caso del IGN de España, Adobe Illustrator).

El formato PDF/X es el formato estándar para la impresión sobre papel puesto que toda la información que contiene es la necesaria e imprescindible para ello. Se puede usar para imprimir en impresoras convencionales, plotters de alta producción y en sistemas profesionales consolidados como es la impresión offset.

6.4.2. Publicación archivos vectoriales a descarga: geoPDF

El formato GeoPDF ha sido el elegido por el IGN de España para la difusión de la serie del MTN a través de archivos a descarga, pues no requiere disponer de un software propietario para acceder a la información.

El funcionamiento de GeoPDF consiste en asociar un marco cartográfico (map frame) a una hoja de un documento PDF. Estos marcos consisten en la descripción del sistema de referencia por coordenadas asociado a la página de tal manera que esta asociación proporciona un conjunto de tuplas que permiten la aplicación desde el sistema de coordenadas de la página PDF al sistema de coordenadas del marco cartográfico. También se pueden definir extensiones o regiones de la página asociadas al marco cartográfico. Son características interesantes la posibilidad de definir varios marcos cartográficos sobre la página, así como que éstos se superpongan e incluso se inserte unos en otros.

Ajustes a realizar en la plantilla de Mapublisher en un caso concreto

Debido al tamaño importante de un archivo geoPDF pues es muy numeroso el número de capas y sus atributos aparte de la georreferenciación que también influye, es posible realizar una serie de ajustes en la plantilla de Mapublisher de tal manera que se pueda reducir considerablemente el tamaño y así facilitar la interacción con el archivo sin penalizar ninguna de sus funcionalidades.

Hemos de reseñar, que estos ajustes se realizarán **en caso de que no se vaya a producir cartografía en formato ráster sino sólo en vectorial (formato geoPDF)**.

A continuación, se señalan los pasos a realizar:

1º. Abrir la plantilla creada para la generación del mapa.

2º. Importar cualquier archivo ffs generado con anterioridad de la siguiente forma:

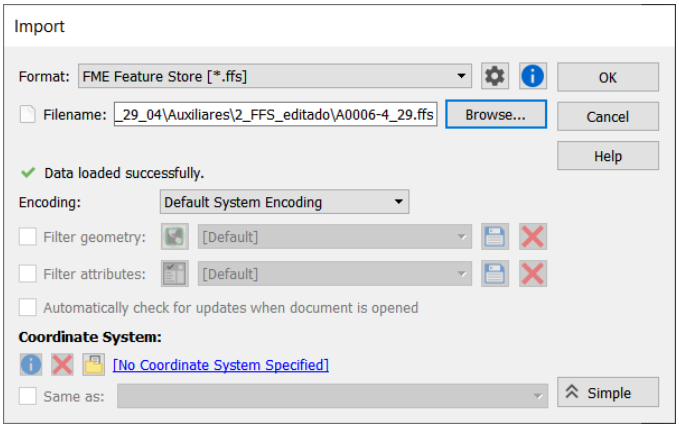


Ilustración 6 - 114: Ventana de importación del archivo ffs

3º. Después de que la importación se haya completado, abrir "Export Document to Geospatial PDF:

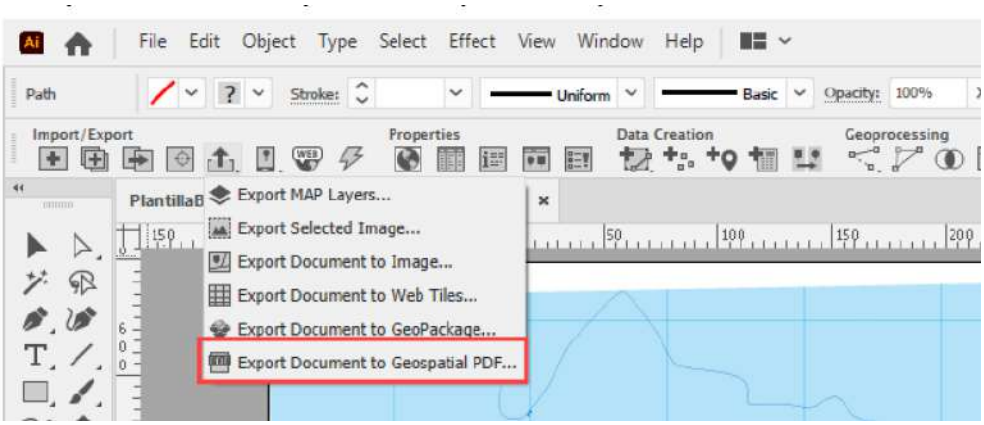


Ilustración 6 - 115: Exportación a geoPDF

4º. En la tabla de contenido, deseleccionar todas las capas que no se quieran incluir en el PDF, por ejemplo, cualquier capa ráster. Realmente es esta capa la que penaliza la fluidez de movimiento.

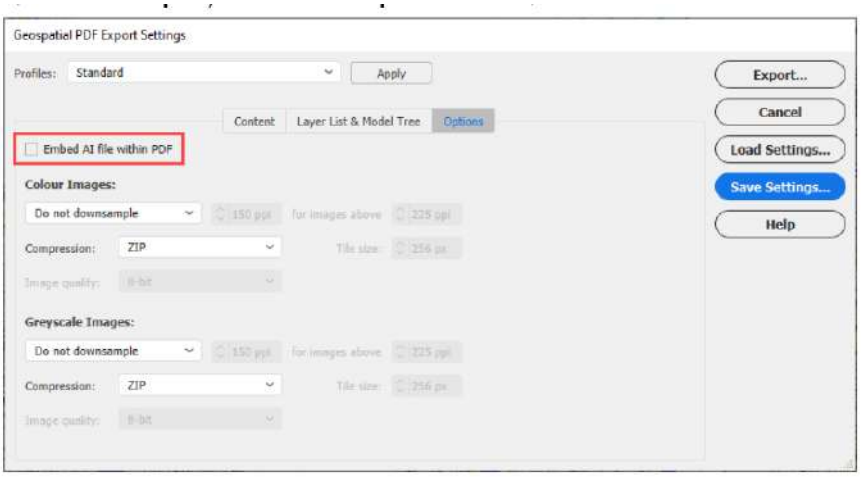


Ilustración 6 - 116: Cuadro de diálogo que permite quitar el archivo AI (responsable del ráster) del archivo geoPDF

5º. Una vez que estamos satisfechos con los ajustes del geoPDF, guardaremos los ajustes.

6º. Cerrar la plantilla Y NO GUARDAR LA MISMA.

7º. Volvemos a abrir la plantilla correspondiente

8º. En la barra de herramientas de MaPublisher hacemos click en el botón de automatización

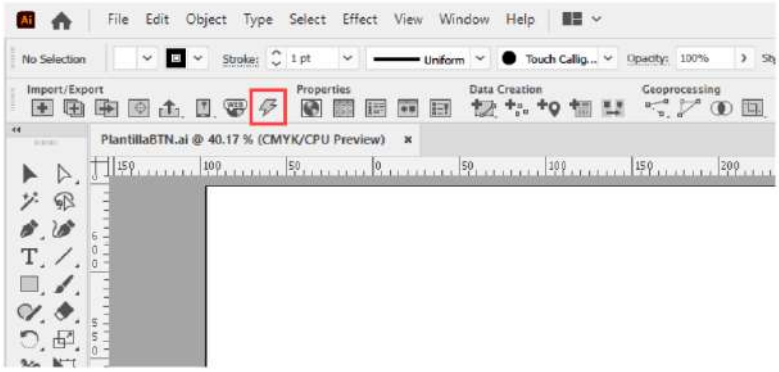


Ilustración 6 - 117: Ajuste para la automatización de cara a FME

9º. En la etiqueta de Exportar a pdf y en exportar a geospatial PDF, hay que seleccionar el archivo .pes (archivo de ajustes) que se ha generado y guardado en el paso 5º.

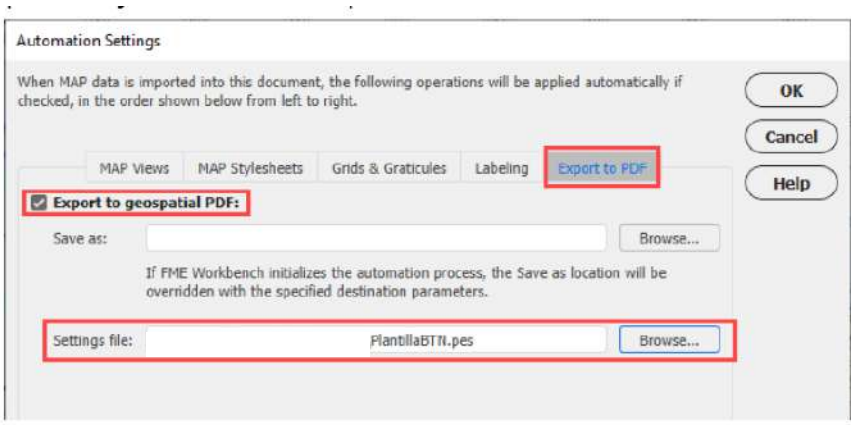


Ilustración 6 - 118: Introducción del archivo .pes de ajustes

10º. Hacer click en ok para guardar los ajustes de automatización

11º. Ahora, guardar la plantilla.

Debemos de destacar, por una parte, que esta plantilla con los nuevos ajustes (se recomienda guardarla con un nombre distinto), habrá que introducirla de nuevo en el writer de FME donde se genera el archivo geoPDF (Parte vectorial).

Por supuesto, este archivo .pes debido a que se ha abierto con una hoja para realizar estos ajustes, es necesario que contenga todas y cada una de las capas involucradas en el proceso para poder generalizar la producción.

La ventaja de este archivo es que es totalmente configurable pues está escrito en formato XML (eXtensible Markup Language) y, por tanto, editable con cualquier editor (por ejemplo, Notepad++)

```
805      <layer display="XX Portada" include_in_layer_list="false"
806    </display_trees>
807
808    <options>
809      <include_illustrator_file enabled="false"/>
810      <colour_images compression="zip" downsampling="none"/>
811      <greyscale_images compression="zip" downsampling="none"/>
812    </options>
813
814  </pdf_export_settings>
815
```

eXtensible Markup Language file

Ilustración 6 - 119: Parámetros para configurar si se quiere o no incluir el archivo ai dentro del geopdf

En esta última captura, podemos observar que, una vez creado el archivo .pes de configuración, podremos configurar que se incluya el archivo ai dentro del geoPDF por si se quiere volver a la producción de ráster.

6.4.3. Publicación archivos raster para servicios y descarga: COG, ECW

Cloud Optimized GeoTIFF" (COG) es un estándar basado en GeoTIFF, diseñado para simplificar el uso de GeoTIFF alojados en servidores web HTTP, de modo que los usuarios y el software puedan utilizar datos parciales dentro del archivo sin tener que descargar el archivo completo.

Debido a su versatilidad y a la facilidad en el manejo de transparencias, es el formato que últimamente se ha adoptado para la publicación del MTN en servicios de visualización por el IGN de España.

El formato Enhanced Compression Wavelet (ECW) es un formato de archivo privativo para imágenes raster desarrollado por la empresa Earth Resource Mapping (en la actualidad propiedad de Leica Geosystems) y que presenta unas ratios muy altas de compresión de forma que reduce considerablemente el tamaño de los archivos, manteniendo una alta calidad gráfica y permitiendo una rápida compresión y descompresión mediante un uso escaso de la memoria RAM.

Este formato (*.ecw) es ampliamente utilizado en geomática en software de procesamiento de imágenes digitales dado que, además de sus ventajas de compresión y rapidez de carga, preservar la georreferenciación de la imagen mediante un archivo de cabecera con extensión .ers.

El IGN de España genera archivos del MTN en este formato, durante el proceso de generación de los archivos COG, para su libre descarga por parte de la ciudadanía.

6.5. Evolución de las salidas cartográfica para visores

En este capítulo, junto con el anterior, se ha mostrado el proceso completo para la generación de cartografía por procedimientos automáticos. El presente apartado apunta la tendencia actual de producción de cartografía para explotación restringida a visores, es decir, sin generación de publicación impresa en paralelo.

La utilización de pantallas de visualización para consultar cartografía tiene una serie de ventajas e inconvenientes frente al uso del papel. Entre los inconvenientes, se puede destacar la menor resolución de las pantallas (frente a una impresión de alta calidad) y el tamaño de las mismas (si se compara con mapas de gran formato). Estos dos inconvenientes se suplen con las ventajas de poder hacer zoom (cambiar la escala de visualización) y desplazamientos de forma muy rápida.

De estas ventajas e inconvenientes se pueden extraer varias conclusiones:

- Se pueden generar múltiples salidas adaptadas para los diferentes niveles de zoom. Es cierto que en la cartografía en papel también se pueden generar múltiples escalas, pero el coste de impresión de cada una de ellas no justifica su existencia, por lo que los saltos de escala entre series suelen ser relativamente grandes (por ejemplo, 1:50,000, 1:200,000, 1:500,000, etc.).
- El diseño de estas múltiples salidas para diferentes niveles de zoom conlleva un esfuerzo de desarrollo, ya que implica todo el proceso de redacción cartográfica. Sin embargo, se debe aprovechar la facilidad de cambiar de zoom para no sobrecargar de información los mapas de escalas intermedias, ya que el usuario puede ampliar la escala en un momento dado para ver más detalles de una zona concreta. De este modo, ese proceso de redacción cartográfica queda más aligerado.
- En este sentido, se debe mencionar la importancia de introducir en las bases de datos origen información sobre la jerarquía o la prioridad de los elementos, con objeto de facilitar su selección para la generación de las distintas escalas.
- Si las bases de datos de partida contienen información con un nivel de detalle superior a la escala de referencia del mapa principal al que dan soporte, se puede generar una salida para escalas o niveles de zoom más grandes e introducir en ella toda la información de la base de datos sin necesidad de realizar prácticamente generalizaciones.

  Véase que, en estos casos, la utilización de simbolización de menor tamaño (que ocupa menos tamaño en la realidad), reduce también las operaciones de redacción cartográfica encaminadas a evitar solapes de simbologías, por lo que se mantienen los objetos geográficos en su verdadera posición. Además, la rotulación o el etiquetado de elementos ocupa también menos espacio, lo que permite la plasmación de un mayor número de nombres y una mejor ubicación de los mismos.

Centrando la atención en esta última conclusión, se muestra a continuación, a modo de ejemplo, la producción que se hace en el IGN de España de una salida cartográfica a una escala superior a 1:25,000 (aproximadamente al doble de esta escala).

6.5.1. Ejemplo de cartografía específica para visores: el mapa de Alta Resolución

Definición y características

A finales del año 2021, se comienza a producir en el IGN DE ESPAÑA un nuevo producto que complementa la visualización del mapa automático a escala 1:25,000 en los niveles más altos de zoom de los servicios de visualización. Se ha denominado mapa de Alta Resolución y es una cartografía obtenida también mediante procesos automáticos, producida en paralelo al mapa automático a escala 1:25,000, si bien, como se explicará más adelante, generada en una producción mucho menos laboriosa.

Se trata de un mapa de apariencia muy similar al MTN25, que tiene prácticamente el mismo contenido y la misma simbolización, pero que, sin embargo, difiere notablemente en el mayor nivel de detalle mostrado, convirtiéndose en un producto más apropiado para estos niveles más altos de zoom.

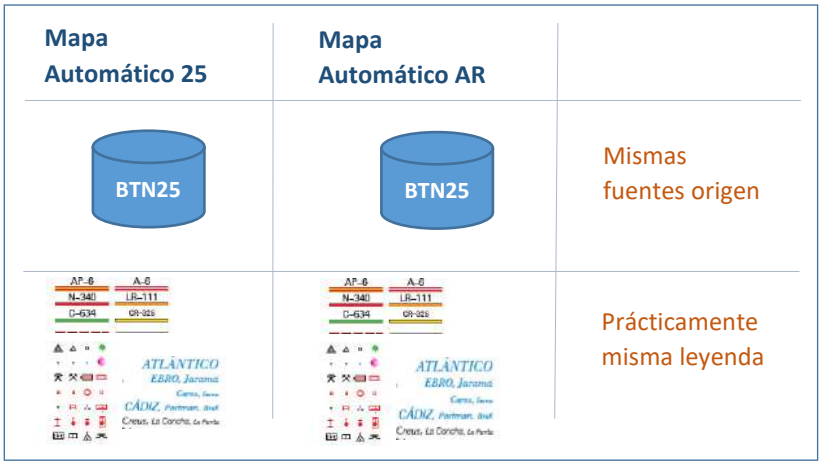


Ilustración 6 - 120: Analogías mapa de Alta Resolución con MTN25 automático

A diferencia del MTN25, utilizado para publicaciones tanto en papel como en visor, este nuevo mapa está concebido únicamente para su publicación en visualizadores. Permite prescindir de los procesos más complejos de generalización y redacción cartográfica automática que exige el MTN25, debido principalmente a los siguientes motivos:

- El nivel de detalle de la mayoría de los objetos geográficos de la base de datos es más acorde a la escala de este nuevo mapa.
- Los tipos de objetos geográficos habituales en bases topográficas de carácter territorial (edificios, carreteras, ríos, láminas de agua, etc.), suelen poder representarse con su dimensión real en escalas entre 1:5,000 y 1:15,000, sin quedar por debajo del límite de percepción visual y pudiendo interpretarse de forma adecuada sin necesidad de exagerar su tamaño. Se evita así la ocurrencia de conflictos entre elementos.

Otra de las ventajas es el menor espacio ocupado por los rótulos o etiquetas, lo que facilita y mejora el etiquetado de la cartografía.

Ilustración 6 - 121: Diferencias con MTN25 automático

Por lo tanto, este nuevo mapa permite visualizar los objetos geográficos con una gran resolución, tal cual proceden de las fuentes de referencia, permitiendo, por ejemplo, visualizar los contornos de edificios con un gran nivel de detalle, acompañando su visualización de una manera fiel a una posible superposición con la ortofoto o con las fuentes de origen.

Diseño de Implementación

Como se ha comentado, una de las ventajas del mapa de Alta Resolución se encuentra en su mayor sencillez de implementación respecto del mapa automático a escala 1:25,000.

Como muestra la figura 6-122, comparando el flujo de implementación con el descrito en el apartado 2.2.7, se aprecia la simplificación de implementación de este nuevo producto. Al no realizar ninguna generalización, se eliminan los subprocesos de tratamiento de lineales y de edificios, además de la generalización preliminar que se realiza en el proceso preparatorio. Se elimina también el subproceso de incorporación de SIOSE ya que para este producto se reutilizan los cultivos generados y almacenados en BBDD cartográfica del mapa automático 1:25,000. Por lo demás, el flujo queda compuesto de las mismas operaciones explicadas en el citado apartado de implementación del mapa automático 1:25,000 y utilizando el mismo software.

Ilustración 6 - 122: Flujo del Mapa de Alta Resolución a partir del de Mapa Automático a escala 1:25,000

BIBLIOGRAFÍA

[1] Sistemas de Información Geográfica. Víctor Olaya. 2014

[2] UNE-EN ISO 19101-1 (ISO 19101-1:2014). Modelo de referencia. Parte 1: Fundamentos

[3] UNE-EN ISO 19109 (ISO 19109:2015). Reglas para esquemas de aplicación.

[4] UNE-EN ISO 19110 (ISO 19110:2016). Metodología para la catalogación de objetos geográficos.

[5] UNE-EN ISO 19131 (ISO 19131:2007). Especificaciones de producto de datos.

[6] Especificaciones de la Base Topográfica Nacional de España. IGN España. 2021

[7] Servicios de geoprocesamiento de CartoCiudad

[8] SERVICIO DE GEOCODIFICACIÓN DE DIRECCIONES DEL SISTEMA CARTOGRÁFICO NACIONAL (ESPAÑA) (https://gzllpz.github.io/prueba2/)

[9] INSPIRE Data Specification on Addresses- Technical Guidelines

[10] Servicio Atom de datos de direcciones de catastro (España)

[11] https://es.wikipedia.org/wiki/JPEG_2000

[12] https://es.wikipedia.org/wiki/Portable_Network_Graphics

[13] Descripción del servicio WFS Parcelas Catastrales España

[14] Wiki principal de OSM https://wiki.openstreetmap.org/wiki/Main_Page

[15] ISO 19157:2013, Calidad de los datos de la información geográfica

[16] ISO 19157:2013/Amd 1:2018, Modificación 1 de la norma: describir la calidad de los datos utilizando coberturas.

[17] Especificaciones la Información Geográfica de Referencia de Redes de Transporte (IGR-RT).

[18] Sistema de indicadores para la mejora y el control integrado de la calidad de los procesos. Heredia Álvaro, José Antonio. 2001. Servicio de Comunicación y Publicaciones de la Universidad Jaume I.

[19] Implementación de procesos de control de calidad en la actualización de series cartográficas urbanas mediante combinación de CAD y SIG. Torrent Foz, Eulàlia; Muñoz-Nieto, Ángel Luis; González-Aguilera, Diego; Rodríguez-Gonzálvez, Pablo. Jun 2021 Revista Cartográfica. https://revistasipgh.org/index.php/rcar/article/view/884/1762

[20] Calidad en datos geográficos, geoservicios y productores de datos: análisis crítico. F. J. Ariza, A. F. Rodríguez. Revista Cartográfica: 105-143, 2018. http://dx.doi.org/10.35424/rcarto.i97.178

[21] Alba, M.V. (2009) Muestreos. En 2ª Edición del Curso de Experto Universitario en Evaluación de la Información Geográfica. Universidad de Jaén. Jaén.

[22] Ariza, F.J. (2009) Normas sobre muestreos (2859, 3951). En 2ª Edición del Curso de Experto Universitario en Evaluación de la Información Geográfica. Universidad de Jaén. Jaén

[23] Normas de Edición del MTN25. IGN. Año edición 2020.

[24] Normas de Toponimia del MTN25. IGN. 2005

[25] Software FME

[26] ArcPy