



PRESERVACIÓ DE BASES VECTORIALS I RÀSTER EN EL MIRAMON

Autor del document: Abel Pau

Proposta inicial: 10-07-2018

Revisors: Xavier Pons

Darrera modificació i versió: 05-07-2023. 1.5

1. Antecedents i motivació

La necessitat de preservació de dades digitals és un fet indiscutible avui en dia, i la informació geogràfica no n'és una excepció. En efecte, existeix la necessitat de poder preservar les dades que generem digitalment i també la necessitat de poder assegurar-nos que, en un futur, serem capaços de recuperar-les, llegir-les, d'entendre'n el format, de poder consultar-les i, si convé, d'actualitzar-les.

És per això que el MiraMon ha implementat un sistema de preservació de bases vectorials, tant siguin en el format estructurat format propi del MiraMon (PNT, ARC, NOD, POL), com en format Shape, com procedents de bases de dades de l'Oracle Spatial (SDO), tot emmagatzemant aquestes dades originals i, alhora, convertint-les i emmagatzemant-les també en un format estàndard com el GML. De forma anàloga, el procés també s'ha implementat per a les bases ràster en format propi (IMG) i en format JPG/JPEG, convertint-les i emmagatzemant-les també en format TIF/TIFF. En tot els casos, el resultat s'empaqueta en un únic fitxer MMZX que segueix l'estàndard ISO 19165 de preservació de dades i metadades d'informació geogràfica.

2. Descripció del sistema de preservació de bases vectorials i ràster

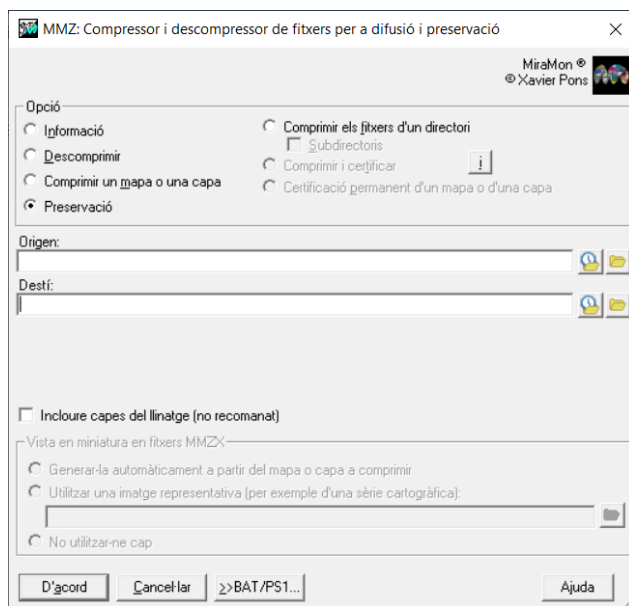
S'ha implementat un sistema de preservació de dades vectorials i ràster que permet preservar les dades gràfiques, les dades temàtiques, la simbolització i les metadades. El sistema suporta com a entrada un mapa (que pot contenir diverses bases) o una base individual i proporciona com a sortida un sol fitxer en format MMZX.

S'ha optat per incloure sempre que es pugui les dades en el format original i, a més, les transformades.

La preservació de la simbolització, en cas que se'n tingui, es realitzarà a través del format SLD, també estàndard, (SLD-OGC V1.0.0), a més d'en el format original quan aquest no sigui SLD.

La preservació de les metadades es farà a través d'un fitxer en format XML que segueix la ISO 19139, a més d'en el format original quan aquest no sigui l'ISO 19139.

Aquest sistema de preservació és accessible des de l'aplicació de suport del MiraMon (MSA) anomenada "MMZ" (Fitxer | Publicar mapa a Internet/intranet | Crear/Descomprimir un MMZX/MMZ...) on es troba l'opció de "Preservació" que permet empaquetar els mapes i les capes ràster i vectorials en un sol fitxer MMZX.



El sistema admet una base en format MiraMon (fitxers amb extensió *.pnt, *.arc, *.nod, *.pol, *.img i fitxers associats), en format Shape (fitxers amb extensió *.shp i fitxers associats), dades procedents de bases de dades de l'Oracle Spatial (fitxers amb extensió *.rel d'aquest tipus que permeten accedir aquestes bases de dades), en format JPEG (fitxers amb extensions *.jpg i *.jpeg) i TIFF (amb extensions *.tiff i *.tif). A més, també admet mapes del MiraMon (fitxers amb extensió *.mmm que continguin qualssevol dels tipus anteriors).

Les bases es poden preservar de forma individual o de forma conjunta utilitzant mapes del MiraMon (format MMM), podent, així, barrejar bases ràster i bases vectorials en un mateix MMZX.

Vegeu els apartats 2.1 i 2.2 per a més detalls sobre les bases vectorials i ràster que es preserven, respectivament.

El procés de preservació dóna com a resultat un MMZX que, un cop descomprimit (usant la mateixa MSA, "MMZ", o qualsevol aplicació que permeti llegir fitxers que segueixin l'especificació ISO 19165), conté la següent informació estructurada en carpetes i fitxers:

- Carpeta DOC: conté les especificacions dels formats involucrats en tot el procés: POL, ARC, PNT, IMG, SHP, DBF, MMZX, TIFF, JPEG, etc.
- Carpeta MM: conté els fitxers originals, ja siguin en format MiraMon, o els altres suportats.
- Carpeta SHP: conté els fitxers que en origen eren en format Shape i els fitxers que en origen eren en format MiraMon i s'han convertit a Shape.
- Carpeta GML: conté els fitxers que en origen eren en format Shape i els fitxers que en origen eren en format MiraMon/SDO i s'han convertit a GML.
- Carpeta TIFF: conté els fitxers que en origen eren en format TIFF/TIF i els fitxers que en origen eren en format MiraMon/JPG/JPEG i s'han convertit a TIFF.
- Fitxer logFiles: conté la llista de fitxers obribles que s'han preservat i dades sobre informació no preservada (donat que no tots els formats permeten preservar totes les característiques d'altres formats i s'ha optat per deixar-ne constància de cara a futurs usuaris).

Vegeu l'apartat 2.3, 2.4 i 2.5 per a detalls sobre l'estructura interna del fitxer resultat en format MMZX.

2.1 Descripció de les bases vectorials que es preserven

La descripció dels formats vectorials del MiraMon es poden trobar a l'apèndix 2 de l'ajuda del programa MiraMon (aquesta descripció s'inclou en la carpeta DOC dels fitxer MMZX generats).

Degut al fet que cada format té els seves avantatges i **limitacions**, i al fet que del que es tracta de preservar al màxim les dades i metadades, s'ha optat per incloure sempre que es pugui les dades en el format original i, a més, les mateixes dades transformades a GML (com a format estàndard àmpliament documentat) i a Shape (format no estàndard però àmpliament usat malgrat les seves limitacions).

El **GML**, acrònim anglès de *Geography Markup Language* (Llenguatge de Marcat Geogràfic) és un subllenguatge de l'XML descrit com una gramàtica en XML Schema per al modelatge, transport i emmagatzematge d'informació geogràfica. La seva importància radica en el fet que a nivell informàtic pretén constituir una llengua franca per al maneig i transvasament d'informació entre els diferents programaris que fan ús d'aquest tipus de dades, com els Sistemes d'Informació Geogràfica.

El GML es va dissenyar a partir de l'especificació abstracta produïda pel grup OpenGIS, ara Open Geospatial Consortium, i de la sèrie de documents ISO 19100. El GML no conté informació específica sobre com s'ha de fer la visualització de les dades representades.

Les **bases vectorials que es poden preservar** poden ser en diversos formats: bases vectorials estructurades en format MiraMon (PNT, ARC, NOD i POL), bases vectorials en format Shape (SHP) o bases vectorials procedents de bases de dades de l'Oracle Spatial (SDO).

En el cas de les bases de dades de l'Oracle Spatial (SDO), com que no es pot preservar la informació del servidor directament, aquesta es converteix a un fitxer en format MiraMon i es preserva com si l'origen fos aquest. Donat que el format vectorial MiraMon és públic i obert, que l'especificació no només existeix a l'ajuda del programa i a Internet sinó que també es copia a l'interior de cada MMZX, i que té unes característiques de gran flexibilitat (topologia, taules alfanumèriques sense

pràcticament límits en longituds de camps, nombre de camps, model relacional de taules, simbolització, metadades, etc) es considera que aquesta opció és una bona manera de preservar la informació SDO.

Les bases en MiraMon i Shape, cadascuna amb les seves peculiaritats, es converteixen a GML mantenint el màxim d'informació possible. Les bases originalment en MiraMon es converteixen, a més, a Shape. Com s'ha dit, les bases en Oracle Spatial es converteixen al format MiraMon abans de ser convertides a Shape i GML.

La **topologia** recull les relacions espacials entre els diferents elements gràfics (topologia de node/punt, topologia de xarxa/arc/línia, topologia de polígon, etc) que representen les característiques geogràfiques i la seva posició al mapa (prop de, entre, adjacent a, etc). Aquestes relacions, que per a l'ésser humà poden ser òbvies a primera vista, han de ser establertes, en el context d'un format informàtic, mitjançant un llenguatge i unes regles de geometria matemàtica. És la capacitat de crear topologia el que diferencia fortament els Sistema d'Informació Geogràfica (SIG) vectorials de més nivell d'altres sistemes de gestió de la informació.

El **GML** és un format **estàndard** que permet descriure bases **vectorials** amb i sense **topologia**. El **problema** d'aquest format és que és molt **complex** i **llarg** d'escriure, fins i tot per a exemples bastant simples. Així doncs, tot i ser estàndard es faria, en un futur, difícil de llegir; i, a més, crearia un fitxer tan gran que no es podria obrir amb els editors més simples. És per això que s'ha decidit fer una transformació a GML sense topologia però mantenint certa informació, en el GML i en altres fitxers, que permeti recuperar-la si un dia es desitja fer-ho des d'aquest format (des del format MiraMon no caldrà ja que la topologia és explícita).

La transformació d'una base a **GML sense topologia** només descriu els objectes finals, és a dir, els polígons es descriuen de forma explícita (atributs temàtics i coordenades) i no es descriuen els arcs ni els nodes que els conformen. Per la seva banda, els arcs es descriuen de forma explícita (atributs temàtics i coordenades) i no es descriuen els nodes que els conformen. Finalment, els punts també es descriuen de forma explícita (atributs temàtics i coordenades).

En el cas de la **preservació de bases MiraMon en GML** s'ha prestat especial atenció a la necessària preservació de la topologia i s'ha desat la informació topològica de la següent manera: una part en els atributs temàtics que presenta el GML (en la seva totalitat en arcs i punts) i una part (en el cas de polígons) en dues taules en format DBF.

En el cas de **preservació de bases Shape en GML**, com que no hi ha topologia es preserva tota la informació dels objectes gràfics i temàtica en el fitxer GML resultant, excepte certa informació que es perd i que es comenta més endavant en aquest document.

Taula que mostra les conversions entre formats i què es preserva en cada conversió.

Origen	Extensió	Conversió a...	Es preserva topologia	Es preserva simbolització	Es preserven metadades
MiraMon	POL/ARC/PNT	Shape	NO	SÍ	SÍ
Oracle Spatial	REL	Shape	NO	SÍ	SÍ
MiraMon	POL	GML	SÍ, amb fitxers adjunts	SÍ	SÍ

MiraMon	ARC/PNT	GML	SÍ	SÍ	SÍ
Oracle Spatial	REL	GML	NO	NO	SÍ
Shape	SHP	GML	NO	NO	SÍ
Mapa	MMM	Shape	NO	SÍ	SÍ
Mapa	MMM	GML	Si, dels fitxers MiraMon	SÍ	SÍ

2.2 Descripció de les bases ràster que es preserven

La descripció dels formats ràster del MiraMon es pot trobar a l'apèndix 1 de l'ajuda del programa MiraMon (aquesta descripció s'inclou en la carpeta DOC dels fitxer MMZX generats).

El format de fitxer **TIFF/TIF** s'utilitza per intercanviar fitxers entre aplicacions i plataformes d'ordinador. TIFF és un format d'imatge de mapa de bits flexible que és suportat pràcticament per totes les aplicacions de d'edició d'imatges i disseny de pàgines. A més, pràcticament tots els escàners d'escriptori poden produir imatges TIFF. Els documents TIFF tenen una mida màxima de 4 GByte; existeix una especificació posterior, BigTIFF, que pot depassar aquesta grandària i que caldrà incorporar si en algun moment es considera necessari.

El format TIFF admet imatges CMYK, RGB, Lab, Indexed Color i Grayscale amb canals alfa i imatges de mode Bitmap sense canals alfa.

Els fitxers d'imatge TIFF poden tenir una profunditat de bits d'1, 8, 16 o 32 bits per canal o banda espectral.

Les **bases ràster que es poden preservar** poden ser en diversos formats: bases ràster en format MiraMon (IMG) o bases ràster en format JPEG (d'extensió JPG o JPEG). Les bases en format TIFF (d'extensió TIFF o TIF) també s'empaqueten en el fitxer MMZX resultant.

2.3 Organització i estructura del fitxer en format MMZX

L'MMZX compleix l'especificació **ISO/FDIS 19165-1:2018(E)** i estructuralment segueix, com s'explica a l'apartat "**7.4 Open Packaging Convention (OPC)**" d'aquest estàndard, la norma ISO/IEC 29500-2. La informació continguda dins el fitxer MMZX creat en la preservació presenta aquesta estructura:

- Carpeta DOC (té tots els documents a l'arrel):
 - **shapefile.pdf**: especificació del format SHP
 - **FormatFitxersTopologicsMiraMon_v1_3.pdf**: especificació dels formats del MiraMon.
 - **DBF_File_structure.pdf**: especificació del format DBASE IV usat en el Shape i en el MiraMon (excepte quan les característiques de la taula precisa la utilització d'una DBF estesa, cas en què se segueix el document del següent punt).
 - **DBF_estesa_v1_5.pdf**: especificació del format DBASE IV estès usat en el MiraMon.
 - **MMZX_spec.pdf**: especificació del format MMZX del MiraMon.

- **tiff6.pdf**: especificació del format TIFF
 - **FormatFitxersRasterMiraMon_v1_0.pdf**: especificació dels formats ràster del MiraMon.
- Carpeta MM: conté els fitxers (MiraMon originals, Shape originals, SDO transformats a MiraMon, JPEG i TIFF originals) organitzats en la mateixa estructura de carpetes d'origen.
 - Carpeta SHP i GML: conté la transformació a Shape i GML, respectivament, dels fitxers a preservar i els manté organitzats com estaven a l'origen en un últim pis que és "POL", "ARC", "NOD", o bé "PNT", segons convingui tal com s'explica més endavant.
 - Carpeta TIFF: conté la transformació a TIFF dels fitxers a preservar.
 - Fitxer **logFiles.txt**: conté les adreces (de disc) dels fitxers obribles comprimits en l'MMZ. Són les adreces originals relatives a l'MMZ. Els fitxers en format TXT i acabats en "_unpreserved" contenen, per a cada fitxer no completament preservat, la llista de característiques no preservades.

2.4 Preservació dades gràfiques i temàtiques vectorials

La preservació de la part gràfica i temàtica de les dades es fa convertint les bases a GML creant també un esquema (XSD) que el descriu.

2.4.1 Preservació de bases vectorials en format MiraMon

Els fitxers en format MiraMon es conserven en el format original a la carpeta "MM" del MMZ, la seva conversió a Shape es conserva en la carpeta "SHP" i la seva conversió a GML es conserva en la carpeta "GML".

Per al cas de la conversió a GML:

En el cas que la base no tingui taules associades ni multiregistres (relacions 1 a molts entre la base gràfica i la taula principal d'atributs de les entitats) es creen, per a cada entitat (*feature*), tants nodes com atributs temàtics tingui la base (cada node amb el nom d'un dels atributs) i un altre node "geometry" per a la geometria.

En el cas que la base tingui taules associades o multiregistres, per a cada entitat es creen dos nodes, un de nom "registre" amb la informació dels atributs temàtics i un altre de nom "geometry" amb les coordenades de l'objecte, seguint l'estàndard GML.

Les taules associades es descriuen en l'esquema XSD, de manera que es pot recuperar l'estructura llegint primer l'esquema per obtenir l'estructura i després llegint el GML per saber els valors dels atributs temàtics i la geometria dels objectes.

En el cas de multiregistres es crearan tants nodes "registre" com registres múltiples tingui cada entitat.

2.4.2 Preservació de bases en format Shape

Els fitxers en format Shape es conserven en el format original a la carpeta "MM" del MMZ i la seva conversió a GML es conserva en la carpeta "GML".

La informació temàtica d'una base en format Shape no pot tenir multiregistre ni taules associades i, per tant, es crearan tants nodes com atributs temàtics tingui la base i un altre node per a la geometria.

2.4.3 Preservació de bases en format Spatial Oracle

La preservació de bases en l'Oracle es fa a través del pas pel format MiraMon. En aquesta conversió s'assegura que no es perdi informació. En el seu moment ja es va fer l'esforç de crear una DBF estesa que permet conservar tants camps com sigui necessari, de la longitud que calgui, amb els noms dels camps amb suport a caràcters especials, etc (característiques, totes elles, que no es poden seguir si s'adopta el format dBASE IV). El format de la DBF estesa queda descrit en el document:

https://www.miramon.cat/new_note/cat/notes/DBF_estesa_v1_5.pdf

Codi de camp canviat

2.4.4 Preservació de la topologia en el format MiraMon

En el format MiraMon, la preservació de la topologia es fa de la següent manera:

Per a bases d'arcs i punts, els propis atributs temàtics preservats en el fitxer GML permeten recuperar la topologia.

Pel cas de polígons, a més, es generen dos fitxers, en format DBF, que completen la informació que no s'ha preservat en els atributs de la base en el fitxer GML:

- Un fitxer en format DBF, que anomenarem fitxer PAL, i que dóna informació sobre els arcs que componen els polígons. El nom d'aquest fitxer és el nom del fitxer original seguit del literal "PAL", i amb extensió DBF.
- Un fitxer en format DBF que anomenarem fitxer APL i que dóna informació sobre els polígons que hi ha a cada costat d'un arc. El nom d'aquest fitxer és el nom del fitxer original seguit del literal "APL", i amb extensió DBF.

Organització de l'MMZ i fitxers que inclou

Transformacions de MM a GML

En una transformació de MiraMon a GML no hi ha gairebé cap pèrdua d'informació. El format GML exigeix que els noms dels camps, però, siguin en minúscula. En el MiraMon i perquè sempre que es pot se segueix el format dBASE IV, els noms dels camps són, per defecte, en majúscula (tot i que si l'usuari ho prefereix altrament, pot fer-ho a través de l'adopció del format DBF estesa).

El fitxer MMZX generat en la preservació, un cop descomprimit, conté una carpeta "GML" amb estructura de subcarpetes que contenen fitxers convertits i que permeten la reconstrucció de la base de la següent manera:

Per a cada base, s'ha conservat l'arbre de carpetes original i dins de l'última carpeta s'ha afegit diverses carpetes més segons el cas (tipus de fitxer amb o sense topologia):

La conversió de bases de polígons sense topologia garantida es deixa a:

- Una carpeta POL en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics dels polígons.

- Una carpeta ARC en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics del fitxer d'arcs sobre el qual se cicla el fitxer de polígons al format GML.
- Una carpeta NOD en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics del fitxer de nodes dels arcs dels polígons al format GML.

La conversió de bases d'arcs sense topologia garantida es deixa a:

- Una carpeta ARC en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics del fitxer d'arcs al format GML.
- Una carpeta NOD en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics del fitxer de nodes dels arcs al format GML.

La conversió de bases de punts sense topologia garantida es deixa a:

- Una carpeta PNT en l'últim pis de l'estructura original que conté la conversió a GML dels elements gràfics i temàtics del fitxer de punts al format GML.

Per a capes de polígons la preservació amb topologia garantida es fa igual que en l'anterior cas però, a més, s'afegeixen un parell de fitxers:

- El fitxer APL: per a cada arc de polígon indica quin polígon (expressat amb l'índex (id_grafic) del fitxer de polígons) queda a la banda dreta i quin a la banda esquerra (en el sentit del traçat dels vèrtexs de l'arc) a través del seu índex en el fitxer de polígons.
- El fitxer PAL: per a cada polígon hi ha un registre múltiple amb tantes repeticions com arcs té el polígon. Per a cada un dels arcs (expressat amb l'índex (id_grafic) del fitxer d'arcs) s'indica si aquest és part de l'exterior o de l'interior del polígon (vora), si és l'últim arc de l'anell (fi) i si cal invertir l'ordre de les coordenades llegides del fitxer d'arcs.

La topologia de polígons es pot recuperar gràcies als atributs temàtics dels següents fitxers:

- En el fitxer de polígons (desat a la carpeta POL): n_arc, n_polig
- En el fitxer d'arcs (desat a la carpeta ARC): node_ini, node_fi
- En el fitxer de nodes (desat a la carpeta NOD): arcs_a_nod, tipus_node

Els atributs temàtics dels arcs que permeten recuperar la topologia són:

- En el fitxer d'arcs: node_ini, node_fi
- En el fitxer de nodes: arcs_a_nod, tipus_node

No hi ha atributs temàtics en els punts que permetin recuperar la topologia. Simplement la capa amb topologia garantida no tindrà punts superposats.

El significat d'aquests atributs es pot trobar a les especificacions del format de bases vectorials del MiraMon (a la carpeta DOC del fitxer MMZX generat). Recordeu, però, que els noms dels camps en GML són en minúscula i en les especificacions els podreu trobar en majúscula.

Transformacions de l'Oracle Spatial a GML

Com ja s'ha dit, la preservació de bases en l'Oracle es fa a través del pas pel format MiraMon i, per tant, tot l'exposat en el punt anterior és aplicable a aquest punt. Les bases Oracle, però, no tenen topologia explícita, per la qual cosa la transformació generarà un fitxer sense topologia garantida.

Transformacions de MM a SHP

Part de la preservació passa per tenir les bases en diversos formats. Per tant, les bases en MiraMon es converteixen també a Shape sabent que això comporta la pèrdua d'informació: noms dels camps seguint l'especificació dBASE IV (camps en majúscules obligatòriament, de menys d'11 caràcters, nombre limitat de registres i camps,...) i pèrdua de topologia. A https://www.mirammon.cat/new_note/cat/notes/ComparativaFormats_MM_i_SHP_v_2_0.pdf es pot consultar una comparativa entre els dos formats.

En el fitxer MMZX trobarem la carpeta SHP que contindrà les bases en origen MiraMon convertides a Shape. Per a cada base, es conserva l'arbre de carpetes original i, dins de l'última carpeta, hi ha diverses carpetes més segons el cas (amb o sense topologia):

La conversió de bases de polígons amb o sense topologia garantida es deixa a:

- Una carpeta POL en l'últim pis de l'estructura original que conté la conversió a Shape de la base del polígon conservant així els atributs dels polígons.
- Una carpeta ARC en l'últim pis de l'estructura original que conté la conversió a Shape del fitxer d'arcs sobre el qual se cicla el fitxer de polígons al format SHP conservant així els atributs dels arcs.
- Una carpeta NOD en l'últim pis de l'estructura original que conté la conversió a Shape de punts del fitxer de nodes dels arcs dels polígons.

La conversió de bases d'arcs amb o sense topologia garantida es deixa a:

- Una carpeta ARC en l'últim pis de l'estructura original que conté la conversió a Shape de la base d'arcs conservant així els seus atributs.
- Una carpeta NOD en l'últim pis de l'estructura original que conté la conversió a Shape de punts del fitxer de nodes de la base d'arcs.

La conversió de bases de punts amb o sense topologia garantida es deixa a:

- Una carpeta PNT en l'últim pis de l'estructura original que conté la conversió a Shape de la base de punts en format MiraMon.

Aquesta base en Shape que en origen és en format MiraMon no es converteix a GML. La conversió es fa directament de MiraMon a GML, com ja s'ha comentat en l'apartat anterior.

La base originàriament en Shape no es converteix al format MiraMon i es deixa directament en una carpeta SHP (sense subcarpetes PNT, ARC, NOD, POL) i mantenint l'estructura original de subcarpetes.

Transformacions de SHP a GML

Finalment, el format originàriament en Shape es converteix a GML.

En una transformació de Shape a GML no hi ha gairebé cap pèrdua d'informació excepte que el format GML exigeix que els noms dels camps, però, siguin en minúscula. En Shape els noms dels camps són en majúscula (format dBASE IV).

2.5 Preservació dades gràfiques i temàtiques ràster

La preservació de la part gràfica i temàtica de les dades es fa convertint les bases a TIFF amb fitxers complementaris per guardar la part temàtica, si existeix.

2.5.1 Preservació de bases en format MiraMon

Els fitxers en format MiraMon es conserven en el format original a la carpeta "MM" de l'MMZ (sense compressió RLE ni similars ni indexat, per a màxima compatibilitat), i la seva conversió a TIFF es desa a la carpeta "TIFF".

2.5.2 Preservació de bases en format JPEG

Els fitxers en format JPEG/JPG es conserven en el format original a la carpeta "MM" i "JPEG" de l'MMZ i la seva conversió a TIFF es conserva en la carpeta "TIFF".

2.5.3 Preservació de bases en format TIFF

Els fitxers en format TIFF/TIF es conserven en el format original a la carpeta "MM" i "TIFF" de l'MMZ.

Organització de l'MMZ i fitxers que inclou

Transformacions de MM a TIFF

En una transformació de MiraMon a TIFF no hi ha pèrdua d'informació pel que fa al valor dels píxels.

El fitxer MMZ generat en la preservació, un cop descomprimit, conté una carpeta "TIFF" amb els fitxers convertits a TIFF i una còpia de les taules DBF que expliquen la informació temàtica associada als valors de la base original, si cal.

Transformacions de JPEG a TIFF

El format originàriament en JPEG es converteix a TIFF, sense pèrdua d'informació.

2.6 Característiques no preservades (*unpreserved features*)

Els fitxers que, per algun motiu, hagin perdut informació en la conversió, aniran acompanyats d'un fitxer de nom com segueix:

"Nom del fitxer original (sense extensió)" + "_" + "extensió del fitxer resultant (sense punt)" + "_unpreserved" + "." + "extensió TXT".

Per exemple, si el fitxer original es diu Municipis.pol i s'ha convertit a Municipis.shp i aquesta conversió s'ha fet amb pèrdua d'informació degut a la limitació del tipus destí, al costat del fitxer Municipis.shp apareixerà un fitxer que es dirà Municipis_shp_unpreserved.txt.

Aquest fitxer contindrà una llista de les característiques que no s'han pogut conservar en la transformació.

El fitxer logFiles.txt, a més de tenir la llista de fitxers obribles, tindrà la llista de fitxers d'aquest tipus per tal que sigui fàcil veure si en una acció de preservació s'ha perdut alguna informació.

2.7 Preservació de les metadades

Les metadades es preserven sempre al mateix directori que el fitxer GML resultant a través d'un fitxer XML que segueix l'estàndard ISO 19139. També es mantenen en aquells formats original que en tinguin (com el MiraMon).

2.8 Preservació de la simbolització

La simbolització es preserva sempre al mateix directori que el fitxer GML resultant a través d'un fitxer SLD que segueix l'estàndard SLD-OGC V1.0.0. També es mantenen en aquells formats original que en tinguin (com el MiraMon).

3 Exemples

3.1 Preservació d'una capa de polígons en format MiraMon

Les dades d'entrada poden ser un fitxer de polígons .POL o bé un mapa .MMM amb una capa de polígons dins (vegeu més endavant detalls sobre mapes).

La capa d'exemple es diu "municipis.pol" i conté els municipis de Catalunya amb els següents atributs temàtics: Codi del municipi, Nom del municipi, Codi Comarcal, Comarca, Cap de comarca, Superfície de la comarca, Codi provincial i Província. Se cicla sobre un fitxer d'arcs que es diu municipis.arc el qual conté nodes que estan en un fitxer que es diu municipis.nod. Cadascun dels fitxers té les seves metadades i simbolització per defecte en un fitxer REL, mentre que els seus atributs temàtics són en un fitxer DBF (que serà DBF estesa si és necessari).

La preservació d'aquest fitxer genera un MMZX que en ser descomprimit sobre un directori conté el fitxer logFiles.txt amb la següent informació:

```
MM\municipis\Municipis.pol;OK
GML\municipis\ARC\Municipis.gml;OK
GML\municipis\NOD\Municipis.gml;OK
GML\municipis\POL\Municipis.gml;OK
SHP\municipis\ARC\Municipis.shp;OK
SHP\municipis\NOD\Municipis.shp;OK
SHP\municipis\POL\Municipis.shp;OK
```

Això vol dir que la informació preservada està distribuïda de la següent manera:

A la carpeta DOC hi ha totes les especificacions (en format pdf) necessàries per a recuperar la informació de dins de l'MMZ.

A la carpeta MM hi ha els fitxers de la capa de Municipis: Municipis.pol, Municipis.arc, Municipis.nod, MunicipisP.dbf,... Si aquests fitxers apunten a algun recurs

(simbolització, tesaurus, ...) aquests estan continguts en la mateixa carpeta mantenint l'estructura original.

A la carpeta SHP hi ha tres carpetes: POL, ARC i NOD.

La carpeta POL conté un Shape de polígons Municipis.SHP resultat de la transformació de la capa de polígons en format MiraMon a format Shape de polígons (fitxers SHP, SHX, PRJ i DBF). A més, la DBF manté els atributs temàtics originals de la capa del MiraMon sempre que sigui possible. Aquests fitxers van acompanyats per un fitxer SLD amb la simbolització del fitxer original, i per un fitxer XML amb les metadades.

La carpeta ARC conté un Shape d'arcs Municipis.shp resultat de la transformació de la capa d'arcs sobre la qual se ciclen els polígons en format MiraMon a format Shape d'arcs (fitxers SHP, SHX, PRJ i DBF). A més, la DBF manté els atributs temàtics originals de la capa d'arcs del MiraMon. Aquests fitxers van acompanyats per un fitxer SLD amb la simbolització del fitxer original i un fitxer XML amb les metadades de la capa d'arcs.

La carpeta NOD conté un Shape de punts Municipis.shp resultat de la transformació de la capa de nodes dels arcs sobre la qual se ciclen els polígons en format MiraMon a format Shape de punts (fitxers SHP, SHX, PRJ i DBF). A més, la DBF manté els atributs temàtics originals de la capa d'arcs del MiraMon. Aquests fitxers NO van acompanyats per un fitxer SLD amb la simbolització del fitxer original NI per un fitxer XML amb les metadades de la capa d'arcs.

A la carpeta GML hi ha tres carpetes: POL, ARC i NOD.

La carpeta POL conté un fitxer GML Municipis.gml resultat de la transformació de la capa de polígons en format MiraMon a format GML (un fitxer GML i un XSD). Aquests fitxers van acompanyats per un fitxer XML amb les metadades. Si la capa original té topologia garantida, aquesta carpeta conté dos fitxers MunicipisAPL.dbf i MunicipisPAL.dbf que, juntament amb informació dels atributs temàtics del GML permetrien recuperar-la.

La carpeta ARC conté un fitxer GML Municipis.gml resultat de la transformació de la capa d'arcs sobre la qual se ciclen els polígons en format MiraMon a format GML (un fitxer GML i un XSD). Aquests fitxers van acompanyats per un fitxer XML amb les metadades.

La carpeta NOD conté un fitxer GML Municipis.gml resultat de la transformació de la capa de nodes dels arcs sobre la qual se ciclen els polígons en format MiraMon a format GML (un fitxer GML i un XSD).

3.2 Preservació d'una capa de polígons en format Shape

Les dades d'entrades poden ser un fitxer de polígons .SHP o bé un mapa .MMM amb una capa de polígons dins (vegeu més endavant detalls sobre mapes).

La capa d'exemple es diu "municipis.shp" i conté els municipis de Catalunya amb els atributs temàtics: Codi del municipi, Nom del municipi, Codi Comarcal, Comarca, Cap de comarca, Superfície de la comarca, Codi provincial i Província.

La preservació d'aquest fitxer genera un MMZX que en ser descomprimit sobre un directori conté el fitxer logFiles.txt amb la següent informació:

MM\Municipis.shp;OK
GML\ARC\Municipis.gml;OK
GML\NOD\Municipis.gml;OK
GML\POL\Municipis.gml;OK
SHP\Municipis.shp;OK

Això vol dir que la informació preservada està distribuïda de la següent manera:

A la carpeta DOC hi ha totes les especificacions (en format pdf) necessàries per a recuperar la informació de dins l'MMZ.

A la carpeta MM hi ha els fitxers de la capa de Municipis en format original (Shape): Municipis.shp, Municipis.shx, Municipis.prj, Municipis.dbf i els que estiguin associats al shape.

A la carpeta GML hi ha tres carpetes: POL, ARC i NOD.

La carpeta POL conté un fitxer GML Municipis.gml resultat de la transformació de la capa de polígons en format Shape a format GML (un fitxer GML i un XSD). Aquests fitxers van acompanyats per un fitxer XML amb les metadades.

La carpeta ARC conté un fitxer GML Municipis.gml resultat de la transformació de la capa d'arcs de polígons vista com una capa d'arcs a format GML (un fitxer GML i un XSD). Aquests fitxers van acompanyats per un fitxer XML amb les metadades.

La carpeta NOD conté un fitxer GML Municipis.gml resultat de la transformació dels vèrtexs dels polígons de la capa de polígons a format GML (un fitxer GML i un XSD).

La carpeta SHP, en aquest cas, conté la mateixa informació que la carpeta MM.

3.3 Preservació d'una capa ràster (multibanda) en format MiraMon

Les dades d'entrades poden ser un fitxer ràster .IMG o bé un mapa .MMM amb una capa ràster dins (vegeu més endavant detalls sobre mapes).

La capa d'exemple es diu "MDE200m_ICC_Aster.img" i conté un "Model de malla regular que conté altituds ortomètriques distribuïdes segons una quadrícula de 200 m de costat, generat a partir del Model d'Elevacions del Terreny de 15 m de l'ICGC (remostrejat a 200 m) per les zones interiors a Catalunya i de l'Aster Global Digital Elevation Model (remostrejat a 200 m) per les zones exteriors a Catalunya". Se simbolitza a través d'una paleta guardada en les metadades del ràster anomenada CMDE256_i_NODATA.dbf i, a més, té una altra banda que s'anomena "MDE200m_ICC_Aster_ombrejat_marblau.img".

La preservació d'aquest fitxer genera un MMZX que en ser descomprimit sobre un directori conté el fitxer logFiles.txt amb la següent informació:

MMMDE200m_ICC_Aster.img;OK
TIFF\MDE200m_ICC_Aster.tiff;OK
MMMDE200m_ICC_Aster_ombrejat_marblau.img;OK
TIFF\MDE200m_ICC_Aster_ombrejat_marblau.tiff;OK

Això vol dir que la informació preservada està distribuïda de la següent manera:

A la carpeta DOC hi ha totes les especificacions (en format pdf) necessàries per a recuperar la informació de dins l'MMZ.

A la carpeta MM hi ha els fitxers de la capa ràster multibanda: MDE200m_ICC_Aster.img i MDE200m_ICC_Aster_ombrejat_marblau.img. Si aquests fitxers apunten a algun recurs (simbolització, tesaurus, ...) aquests estan continguts en la mateixa carpeta mantenint l'estructura original. Tots els ràster propis del MiraMon (IMG) es presenten sense compressió ni indexació per a màxima compatibilitat.

A la carpeta TIFF hi ha la conversió a TIFF de les imatges anteriors (totes, malgrat només se'n visualitzi una). Conté també els fitxers SLD (de visualització) i XML (de metadades) per a cada imatge.

3.4 Preservació d'una capa ràster en format JPEG

Les dades d'entrada poden ser un fitxer de ràster .JPG/.JPEG o bé un mapa .MMM amb una capa JPEG dins (vegeu més endavant detalls sobre mapes).

La capa d'exemple es diu "T31TEG_20170719T103021_PVI.jpeg" i és una imatge de satèl·lit.

La preservació d'aquest fitxer genera un MMZX que en ser descomprimit sobre un directori conté el fitxer logFiles.txt amb la següent informació:

```
MMT31TEG_20170719T103021_PVI.jpeg;OK  
TIFF\T31TEG_20170719T103021_PVI.tif;OK
```

Això vol dir que la informació preservada està distribuïda de la següent manera:

A la carpeta DOC hi ha totes les especificacions (en format pdf) necessàries per a recuperar la informació de dins l'MMZ.

A la carpeta MM hi ha els fitxers de la imatge JPEG en format original (JPEG): T31TEG_20170719T103021_PVI.jpeg i amb les metadades si en tenia.

A la carpeta TIFF hi ha la conversió a TIFF de la imatge anterior. Conté també els fitxers SLD (de visualització) i XML (de metadades)

3.5 Preservació d'una capa ràster en format TIFF (unibanda o multibanda en un sol fitxer)

Les dades d'entrades poden ser un fitxer de ràster .TIFF o bé un mapa .MMM amb una capa TIFF dins (vegeu més endavant detalls sobre mapes).

La capa d'exemple es diu "UsosCobertes_1987.tif" i és una imatge d'usos de cobertes del 1987.

La preservació d'aquest fitxer genera un MMZX que en ser descomprimit sobre un directori conté el fitxer logFiles.txt amb la següent informació:

```
MM\UsosCobertes_1987.tif;OK
```

Això vol dir que la informació preservada està distribuïda de la següent manera:

A la carpeta DOC hi ha totes les especificacions (en format pdf) necessàries per a recuperar la informació de dins l'MMZ.

A la carpeta MM i TIFF hi ha els fitxers de la imatge TIFF en format original (TIFF). Conté també els fitxers SLD (de visualització) i XML (de metadades) per a cada imatge.

3.6 Preservació d'una capa ràster en format TIFF (multibanda)

Si el TIFF és multibanda en el sentit que existeix un fitxer de metadades que lliga totes les bandes (emmagatzemades en fitxers TIFF diferents), caldrà tenir un MMM amb totes bandes incloses per a poder preservar-les en un sol MMZX. Vegeu la secció "preservació d'un mapa".

3.7 Preservació d'un mapa

La diferència entre preservar una capa sola i un mapa amb la capa és que en aquest segon cas, es preserva l'estructura inicial de carpetes que contenen la capa (o les capes si n'hi ha més d'una) per evitar solapaments de capes amb el mateix nom, però òbviament en carpetes diferents. A més, el mapa en format MiraMon també es preserva a la carpeta MM amb el mateix nom.

4.- Referències

- Article de la revista Computers & Geosciences que descriu i discuteix la idea de l'MMZ i estableix la base de la norma ISO de preservació 19165 : **Pons X, Masó J** (2016) A comprehensive open package format for preservation and distribution of geospatial data and metadata. Computers & Geosciences 97: 89-97. DOI: 10.1016/j.cageo.2016.09.001. <http://dx.doi.org/10.1016/j.cageo.2016.09.001>.
- Fitxers públics referents als formats del MiraMon: <https://www.mirammon.cat/CAT/NotesTecniques.htm>
- Format TIFF extret de la revisió 6.0 del 3 de Juny de 1992 (format públic), ampliacions posteriors i estandarditzat a través de la norma ISO 12639, <https://www.iso.org/standard/34342.html>. Més informació a: <http://www.fileformat.info/format/tiff/corion.htm>
- Format GeoTIFF: <http://trac.osgeo.org/geotiff/>
- Norma ISO **19165:2018** de Preservació d'informació geogràfica (dades i metadades): <https://www.iso.org/obp/ui/#iso:std:iso:19165:-1:ed-1:v1:en>
- ESRI Shapefile Technical Description. An ESRI White Paper, July 1998: <https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>
- GML d'OGC: <http://www.opengeospatial.org/standards/gml>