



## AVANTATGES DELS FORMATS VECTORIALS DEL **MIRAMON** RESPECTE DELS SHAPEFILE

**Autor del document:** Xavier Pons i Joan Masó  
**Proposta inicial:** 12-03-2013  
**Darrera modificació i versió:** 14-03-2013. **2.0**

### 1. Antecedents i motivació.

El desenvolupament del **MiraMon** va iniciar-se el 1994. Les primeres versions van concentrar-se en la lectura i representació de formats ràster IMG i BMP i de formats vectorials no estructurats (VEC). Com a SIG professional, la necessitat de gestionar grans volums de dades vectorials de forma ràpida, de gestionar gran varietat d'atributs temàtics, així com d'efectuar operacions, tant de tipus simple com d'anàlisi, que requerissin una estructura topològica explícita va fer que ben aviat es definís una família de formats propis per a punts, arcs/nodes i polígons (PNT, ARC/NOD, POL). Cal fer notar que en aquell moment no hi havia cap iniciativa d'especificació oberta que permetés un accés eficient a les diferents parts del fitxer, enllaç amb bases de dades, etc (per exemple, la idea del GML va ser presentada a l'OGC el 1999).

Cada tipus de fitxer està format per una tríada de fitxers:

- La **base geogràfica** (per exemple fonts.pnt), que conté les coordenades, codificades binàriament en reals doble precisió, així com les relacions topològiques entre les entitats vectorials.
- La **taula amb els principals atributs alfanumèrics** de les entitats vectorials (per exemple fontsT.dbf), que segueix el format estàndard *de facto* DBF III plus/DBF IV establert per dBASE, i és ampliada a DBF estesa quan cal. Aquesta taula és enllaçable a un nombre il·limitat de taules en

diversos formats i de diferents orígens (grans bases de dades, fitxers MDB, etc).

- El **fitxer REL** (per exemple fontsT.rel), que segueix el format estàndard *de-facto* INI definit per Microsoft en el Windows (basat en seccions, claus i valors) i que conté:
  - Les **relacions de la taula principal amb altres taules** (grans bases de dades, fitxers MDB, etc) o altres fitxers REL distribuïts.
  - Les **metadades** de la capa o, en el cas de sèries i multisèries cartogràfiques, enllaç amb els fitxers REL que especifiquen metadades a altres nivells jeràrquics (de la capa-full, de la sèrie, etc).
  - La **simbolització per defecte** a aplicar en obrir la capa individualment.

El document d'ajuda del **MiraMon** amplia la informació donada en aquesta introducció.

Aquesta especificació va ser pensada força detingudament, tant des del punt de vista conceptual com de rendiment, i bona mostra d'això és que ha sofert molt poques modificacions, essent les més importants:

- L'abandonament del format DVx de metadades que existia en les primeres versions, per excessivament simple, que havia estat inspirat en el format DVC dels formats vectorials de l'Idrisi, i la inclusió de **metadades seguint els estàndards internacionals** que s'estaven configurant (FGDC als USA, CEN a Europa, ISO 19115, etc). Aquestes metadades es van incorporar en el fitxer REL que forma part de qualsevol capa.
- La incorporació de la **secció 3D**.
- L'ampliació de possibilitats de la taula principal (i, per extensió, de qualsevol taula) gràcies a la **DBF estesa**. Vegeu el document tècnic DBF\_estesa\_v1\_4.doc per a detalls.

D'altra banda, des de fa força anys el **MiraMon** suporta diversos formats, en primera instància en importació, però actualment en molts casos en lectura directa, si més no en l'entorn de visualització, consulta i anàlisis simples. Un dels formats més populars que suporta és el **ShapeFile** d'ESRI (**SHP** en endavant). El format SHP és en realitat també format per, com a mínim, 3 fitxers: SHP, SHX i DBF; amb els anys s'ha anat incorporant altres fitxers opcionals que permeten millorar alguns aspectes com proporcionar el sistema de referència espacial (PRJ), etc. Podeu trobar informació detallada sobre aquest format al document:

"*ESRI Shapefile Technical Description*", del juliol de **1998**.  
<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>

Actualment els Mapes del MiraMon (MMM i MMZX/MMZ) poden incloure pràcticament qualsevol dels formats suportats pel programa i, donada la facilitat amb què són llegits, doncs, alguns usuaris del MiraMon es pregunten quins avantatges pot suposar el treballa amb els formats propis del MiraMon. A requeriment, doncs, d'aquest lògic dubte, ens ha semblat oportú d'escriure el

present document, que recull els principals avantatges de la utilització dels formats vectorials del MiraMon.

Cal fer notar que en la següent comparativa no hem destacat aquelles bondats que siguin comunes als dos formats, com ara el fet que les coordenades siguin binàries, amb el que això significa de millor velocitat de lectura, menor volum de transferència en xarxa, etc en comparació amb altres formats com DXF o GML.

## 2. Avantatges dels formats MiraMon.

**NOTA:** En aquesta secció utilitzarem les sigles **MM** per referir-nos als formats vectorials del MiraMon, i **SHP** per referir-nos al format ShapeFile.

### Estructura de fitxers

#### ***Extensions dels fitxers:***

**MM:** L'extensió dels fitxers MiraMon permet que en un explorador de fitxers convencional els usuaris sàpiguen quina natura d'elements conté un cert fitxer (punts, línies, etc) i distingir-los per l'extensió o per la icona.

**SHP:** Cal fer un accés a l'interior dels fitxers per a saber si un fitxer és de punts, línies, etc, la qual cosa pot ser més lenta en un directori amb molts fitxers i no és reconeguda per un explorador estàndard del Windows.

### Model de dades

#### ***Riquesa d'elements gràfics en punts:***

**MM:** Punts tractats per unicitat: si hi ha diversos punts al mateix lloc amb diversos conjunts de propietats (diversos registres), es tracten a través d'una relació 1 a n (registre múltiple; vegeu *Relacions entre els elements gràfics i els atributs alfanumèrics*, més endavant). No se suporten, de moment, agrupacions de punts en ubicacions separades.

**SHP:** Suporta multipunts

#### ***Riquesa d'elements gràfics en polígons:***

**MM:**

- Un mateix fitxer d'arcs pot servir per a ciclar més d'un fitxer de polígons, augmentant la consistència topològica i geomètrica (un límit municipal pot ser alhora límit comarcal i frontera de país, per exemple).
- Se suporten polígons topològics i no topològics. En el primer cas poden definir-se grups de polígons (arxipèlags) que són considerats com un "superobjecte" i no s'admeten superposicions. En el segon cas podem tenir polígons que comparteixin arcs, així com polígons explícits (cada polígon repeteix totes les coordenades que necessita) i superposicions. Algunes de les variants admeses s'han adoptat per facilitar la importació ràpida de

models particulars, com el que usa l'ICC per als polígons als DGN, o per a la lectura/interpretació ràpida de formats com SHP.

**SHP:** Només se suporten polígons no topològics i explícits (coordenades repetides a totes les fronteres comunes entre polígons, una línia no es pot reaprofitar per a ciclar més d'un polígon, etc), que poden definir grups i presentar superposicions.

#### ***Barreja de diverses tipologies d'elements gràfics:***

**MM:** A través de l'element de jerarquia superior en MiraMon: el mapa (MMM o MMZX/MMZ).

**SHP:** També cal emprar un element de jerarquia superior encara que aparentment podria semblar que sí en una exploració superficial: (extret de <http://en.wikipedia.org/wiki/Shapefile>) *Because the shape type precedes each record, a shape file is physically capable of storing a mixture of different shape types. However, the specification states, "All the non-Null shapes in a shapefile are required to be of the same shape type." Therefore this ability to mix shape types must be limited to interspersing null shapes with the single shape type declared in the file's header. A shape file must not contain both polyline and polygon data, for example, and the descriptions for a well (point), a river (polyline), and a lake (polygon) would be stored in three separate files.*

#### ***Suport multiidiomàtic:***

**MM:** Integrat en el model, de forma que el suport no duplica els aspectes no idiomàtics (per exemple no hi ha un fitxer de metadades per a cada idioma), facilitant enormement el manteniment de les dades i metadades. El suport s'ha establert per títols, títols alternatius, descriptors de camps, etc, cosa que és útil si cal elevar a instàncies superiors (europees, per exemple) les metadades d'una conjunt d'informació geogràfica.

**SHP:** Sense suport multiidiomàtic.

#### ***Suport a sèries cartogràfiques:***

**MM:** Sí, a quatre nivells: capa-full, full de la sèrie, sèrie i multisèrie, distribuïts de forma relacional . Vegeu el document tècnic EspecificacionsMDSerie\_1\_1.doc per a detalls.

**SHP:** Sense suport.

#### ***Suport a multisèries cartogràfiques parcials:***

**MM:** Sí. Útil per a relacionar famílies de sèries en capes diferents (per exemple per agrupar i jerarquitzar capes de carreteres de diferents nivells).

Vegeu el document tècnic MultiseriesParcials\_v6\_01\_0.doc per a detalls.

**SHP:** Sense suport.

### **Metadades**

#### ***Model adoptat:***

**MM:** Seguint els estàndards internacionals (ISO 19115, 19115-2, etc) i ampliant allà on cal afinar en aspectes encara no recollits en aquells.

**SHP:** No nadiuament integrat en l'especificació. Els fitxers .shp.xml poden ser presents per a documentar metadades.

**Coherència amb les dades:**

**MM:** Tots aquells aspectes derivables de les dades (natura dels objectes, nombre d'objectes, envolupant geogràfica, data d'adquisició, resolució, etc) automàticament vinculats, sense necessitat de "picar-los".

**SHP:** Sense cap relació amb les dades (purament documentals) excepte si s'utilitza un editor que faci aquestes operacions accedint als fitxers corresponents.

**Facilitats amb la IDEC:**

**MM:** Suport específic al perfil IDEC.

**SHP:** Sense suport específic.

**Riquesa:**

**MM:** Metadades molt detallades, que permeten per exemple especificar la natura dels camps de les bases de dades (o les cel·les en els ràsters) (dades categòriques, quantitatives contínues), les unitats dels camps, la seva qualitat (amb indicadors estandarditzats com l'R<sup>2</sup> o definits per l'usuari), característiques de les imatges de teledetecció, etc.

**SHP:** Sense suport detallat (ISO 19115 no l'inclou). Per exemple, sense relació explícita amb els camps de la DBF.

**Especificació de sensedades (nodata):**

**MM:** Suport explícit i tractament coherent.

**SHP:** Encara que teòricament possible (i MiraMon així ho fa en llegir SHP), a la pràctica inexistent: (extret de <http://en.wikipedia.org/wiki/Shapefile>) *While the current dBase standard, and GDAL/OGR (the main open source software library for reading and writing shapefiles) support null values, ESRI software represents these values as zeros---a very serious issue for analyzing quantitative data, as it may skew representation and statistics if null quantities are represented as zero.*

**Geometria**

**Suport 3D:**

**MM:** Sí en coordenades XYZ. És possible definir tant **cota simple** com **més d'una coordenada Z per cada vèrtex**, cosa que pot arribar a ser útil en extrusions volumètriques d'edificis, per exemple, o en punts que donen la cota al llit d'un riu i, alhora, sobre el pont que el travessa.

**SHP:** Sí en coordenades XYZ, amb **cota simple**.

**Suport a sistemes de referència:**

**MM:** Els sistemes de referència s'especifiquen a través d'un identificador textual que permet relacionar tots els paràmetres del sistema (dàtum, el·lipsoide, projecció i els seus paràmetres com paral·lels de referència, etc) en base a taules d'especificacions comunes. Això evita problemes de redundància, errors de transcripció, etc.

**SHP:** En cada capa cal especificar un fitxer PRJ que especifica dades com ara ["Clarke\_1866",6378206.4,294.9786982], la qual cosa fa que el control de la coherència de tot plegat sigui molt difícil en bases amb milers de capes (sèries, etc), a part de l'enorme repetició que constitueix aquesta estratègia.

## **Topologia**

### ***General:***

**MM:** Suporta contenir dades amb topologia explícita, o dades sense topologia. Els avantatges de la primera opció són evidents per a un munt d'operacions SIG, des del control que els polígons corresponents a un àmbit geogràfic (parcel·les del cadastre d'un municipi, per exemple) omplen exhaustivament aquell àmbit, a l'anàlisi de xarxes, a la garantia d'unicitat de les entitats espacials en cada punt del territori (per exemple no hi ha un punt que representa un vèrtex geodèsic repetit 3 cops), etc.

**SHP:** No suporta topologia. Això vol dir, per exemple, que en obrir una capa de polígons no tenim cap garantia de si hi ha superposicions no desitjades, esclatxes entre polígons, etc, a part que això en limita les capacitats d'anàlisi directa.

### ***Punts:***

**MM:** Se'n garanteix la unicitat.

**SHP:** En no tenir topologia, no existeix garantia explícit d'unicitat.

### ***Nodes 1:***

**MM:** Els fitxers de nodes són fitxers enterament topològics i no contenen les seves coordenades ja que aquestes es deriven dels arcs a què són vinculats, augmentant la consistència.

**SHP:** En no tenir topologia, no existeixen els fitxers de nodes.

### ***Nodes 2:***

**MM:** Quatre tipus de fitxers de nodes, per a un control fi de les relacions entre arcs per topologia o per diferències atributives.

**SHP:** En no tenir topologia, no existeixen els fitxers de nodes.

### ***Nodes 3:***

**MM:** Els fitxers de nodes disposen de base de dades pròpia, útil per a definir propietats de semàfors, passos de vianants (si són adaptats per a persones amb mobilitat reduïda, per exemple), etc.

**SHP:** En no tenir topologia, no existeixen els fitxers de nodes.

### ***Polígons:***

Vegeu, dintre de **Model de dades**, el punt ***Riquesa d'elements gràfics en polígons***.

## **Simbolització per defecte**

**MM:** Integrada en el fitxer REL i distribuïda en un esquema relacional (per exemple la paleta de colors per a les construccions humanes d'una sèrie cartogràfica no està escrita a cada full, sinó definida un sol cop en un fitxer al qual apunten tots els fulls).

**SHP:** No en els 3 fitxers bàsics. Ignorem si hi ha una especificació detallada i àmpliament suportada per la indústria.

### **Atributs alfanumèrics**

#### ***Riscos en la taula principal:***

**MM:** En tenir un camp clau a la taula principal d'atributs (camp ID\_GRAFIC), aquesta no només es pot reordenar, sinó que admet més d'un registre per entitat (v. punt següent).

**SHP:** No existeix l'equivalent d'un camp ID\_GRAFIC, la qual cosa impossibilita saber a quina entitat gràfica correspon cada registre: Si es reordenen els registres per algun criteri es perd la relació i no és possible (llevat que haguéssim creat un camp clau per poder desfer l'operació) refer-la!

#### ***Relacions entre els elements gràfics i els atributs alfanumèrics:***

**MM:** Pot ser 1 a 1, o 1 a n. Això implica que una mateixa entitat pot tenir associats diversos registres a la base de dades. Per exemple una font en què fem periòdicament anàlisis d'aigües pot tenir tants registres (amb camps que indicaran els resultats de cada element físico-químic analitzat) com dates en què hem fet les analítiques.

**SHP:** No és possible associar una mateixa entitat a múltiples registres, ja que en no haver un identificador d'entitat en la taula principal, si el nombre de registres no fos igual al nombre d'entitats geomètriques no sabria quin o quins registres corresponen a cadascuna. En conseqüència, no podem per exemple tenir un polígon forestal associat a les diferents espècies d'arbre que hi ha al seu interior, o un polígon cadastral associat a diversos propietaris a la taula DBF del SHP.

#### ***Relacions entre taules:***

**MM:** Les taules d'atributs alfanumèrics poden ser enllaçades amb altres taules en relacions de cardinalitat flexible (1 a 1, 1 a molts, amb obligatorietat de relació [diccionari] o no, etc), amb gran potencialitat: Qualsevol camp es pot enllaçar a un nombre a la pràctica il·limitat de taules i així successivament, en un nombre a la pràctica il·limitat de nivells de relació. Les taules poden ser DBF, DBF estesa, taules i consultes dintre fitxers com ara MDB o XLS, dintre de grans gestors de i/o bases de dades (Oracle, SQLServer, etc). Poden, per tant, relacionar-se de forma transparent taules ubicades en sistemes diferents (per exemple una taula MS-Access pot relacionar-se amb un diccionari ubicat a un servidor Oracle i, alhora, un camp del diccionari Oracle, vincular-se contra un DBF en un ordinador al qual accedim a través d'una adreça UNC). Per a taules de les darreres versions dels sistemes operatius podeu consultar el document tècnic DSN\_and\_ODBC\_in\_Windows7\_v1\_0.doc

**SHP:** No existeixen relacions explícitament establertes més enllà de l'única taula DBF.

***Relacions alternatives entre taules:***

**MM:** Una mateixa base gràfica es pot vincular a diferents universos de relacions entre taules alfanumèriques, sense necessitat de duplicar cap dada gràfica. Per exemple la base municipal pot ser vinculada a dades sanitàries, a dades censals o a dades de gestors de recursos. Vegeu el document RelAlternatiu\_v0\_7.doc per a detalls.

**SHP:** En no definir relacions, és impossible.

***Suport a camps que especifiquen hipervincles:***

**MM:** Sí, tant a recursos locals com a Internet.

**SHP:** No.

***Suport a camps que especifiquen continguts incrustats:***

**MM:** Sí, tant a recursos locals com a Internet.

**SHP:** No.

***Suport a camps que especifiquen expressions:***

**MM:** Sí, però el tipus d'expressions suportades és parcial, a l'espera que es vagin ampliant.

**SHP:** No.

***Suport a separadors, textos de color, etc per a estructurar la informació donada en consultes, etc:***

**MM:** Sí (codi HTML).

**SHP:** No.

***Capacitat d'indicar si un camp és mostrable, si cal mostrar les seves unitats, etc.:***

**MM:** Sí.

**SHP:** No.

**Atributs alfanumèrics en usar DBF estesa en MM**

***Caràcters admesos per a escriure el nom d'un camp:***

**MM:** Pràcticament qualsevol caràcter, lletra accentuada, etc, excepte l'accent obert sol (sense accentuar cap lletra), els claudàtors, la dièresi i els caràcters no imprimibles (com ara el retorn de carro o el DEL) (DBF estesa).

**SHP:** Lletres majúscules de l'alfabet anglès, nombres i subratllat (però no en posició inicial).

***Màxima longitud del nom de cada camp:***

**MM:** 128 caràcters (DBF estesa).

**SHP:** 10 caràcters.

***Màxim nombre de camps per taula:***

**MM:** 13.4 milions de camps (DBF estesa).



**SHP:** 255 camps.

***Màxim nombre de caràcters en camps de text:***

**MM:** 4200 milions de caràcters (DBF estesa).

**SHP:** 254 caràcters.