CO-431

NEW GIS BASED MAP OF SPAIN 1:500.000

NÚÑEZ-MADERAL E., GONZALEZ-MATESANZ J., GARCIA J., MERINO J.A., DAVILA F.J., ROMERO I., GULLÓN T., CALVO S., AGUIAR N., ALCAZAR A.

National Geographic Institute Spain, SPAIN, SPAIN

INTRODUCTION

The production and publication of the 1:500.000 scale map has been traditionally linked to the National Atlas project since 1960. The first version was composed of 15 sheets as a result from the generalization of the 1122 maps from National Topographic Map at scale 1:50.000. The Next version was produced in 1980 and following international specifications for this scale it was called “World Series 1404”, and it was compiled from Provincial Maps at scale 1:200.000.

The last version of this cartography was produced on 14 sheets based on the last update of provincial maps at scale 1:200.000 based on digital CAD technologies. This version was reedited changing the format from 14 sheets to 25 because of a new reduced format for the National Atlas of Spain.

Cartographic rules were written with the objective of generating a new updated version of the Spanish Map at scale 1:500.000. Outsourcing was used for the compilation, the edition and the sheet composition tasks, but after two years they project failed.

After decided to drive the project, it was included in the Strategic Plan of the National Geographic Institute as an innovation project led by the Support and Development Group created within the Cartographic Department. This group was created in parallel to workgroups that have been developing and producing maps on a CAD environment for many years with a clear purpose: To change the cartographic processes trough turning the well established CAD workflows into a new seamless cartography GIS based environment. This scale 1:500.000 was selected as a test workbench cartography because it is considered to be the most appropriated in medium-short terms to make conclusions based on the results and extend this experience to different and more complex scales as the 1:200.000 or the 1:25.000 scales which have longer production periods.

We define ME500 as a project to create a new seamless Map of Spain at scale 1:500.000 from BCN200 Geographic Information System which pretends to be the first whole experience in GIS based map production in the IGN Spain. ME500 integrates model and geometry generalization processes, cartographic debugging and validation procedures, hypsometry and shaded relief generation, symbology design and manage, automatic label placement and label edition on a GIS based environment.

The First two years of the project was considered as a pilot period to test all the processes and tools and we concluded with some results to consider, on one hand the experience gained in each step of the project and on the other hand the problems with the source data, (BCN200 was not ready to provide consistent geographic data).

We decided to star by developing technical specifications to document the project and provide an implementation reference manual to reach the objective, build a new ME500 and afterwards, conflate different valid sources into a common data base as a reference geographical data source for the project. A new database BCNd500 was created with the definition of a new data model to store all ME500 features considered.

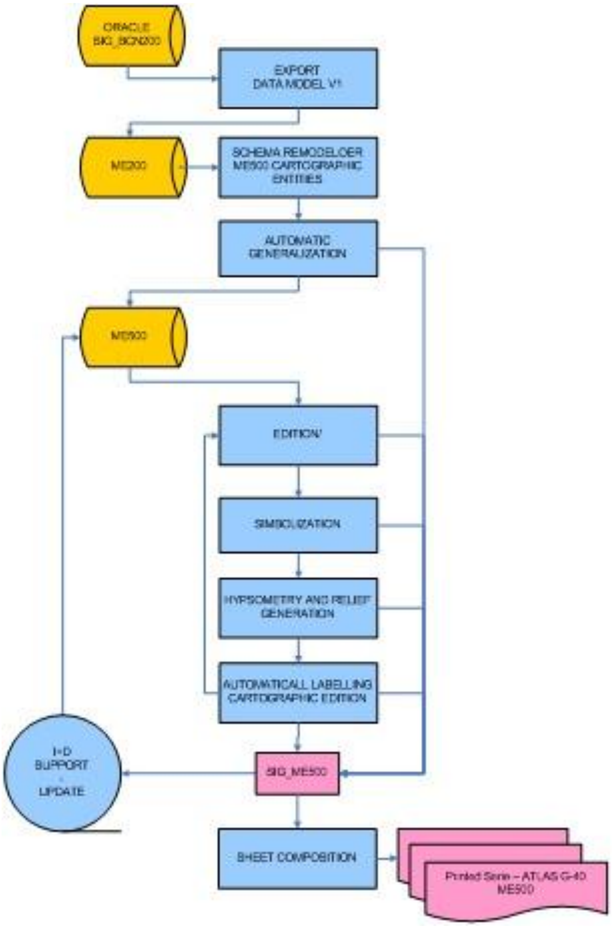
The team involved in this project was composed by Cartography Engineers, Surveyor Engineers, Geographers and Cartography Drafters graduates.

The project will be finished during this year (2011) with the result of BCNd500 as a seamless cartographic database to cover all Spain at scale 1:500.000 and a digital cartographic product which will give response to the visualization and analysis within multiscale projects such as the Spanish SDI (Idee) or Iberpix . Moreover, a printed version of the M500 in 25 sheets will serve the purpouse of publishing the new upcoming reduced format of the Spanish National Atlas.

PILOT PROJECT

The project was started from the scratch; engineers involved had not worked before with CAD processes and were introduced in GIS environment recently. We had the cartographic rules and the source data available in the IGN, (which was BCN200).

We have designed a logical workflow oriented to the different processes. In order to build the database we have imported data from derived cartography as a source. We have made a feature classification and selection which allowed us to set aside compound geometry based on the first model of BCN200 and start with point, line and area base geometry features. After that we applied lineal and areal geometry generalization with automatic routines on lines and areas features. The next picture shows the different processes and ME500 workflow:



Picture 1. Me500 pilot project workflow

From BCN200 data import and generalization processes we obtain a Cartographic Feature Classes Catalog, physically loaded in the database in 72 tables. Each table had BCN200 data model structure, its own ME500 class identifier and only one base geometry. Consequently, the same feature class could appear in two different tables depending on its base geometry.

The next step was the cartographic compilation starting with validation, updating and edition of geometries tasks. It was the most tedious process of the project. At this point we became aware of the project's non-viability due to the source data inconsistence and non updated features for example, area geometries that appear like lines as an inheritance from CAD files and automatic processes did not work as they should. These processes required many hours and a team that we did not have and really these were not tasks of this project but BCN200's.

Despite this, we continue working on the symbol design and legend, generating hypsometry and shaded relief from MTN25 (National Topographic Map scale 1:25.000) level contours and BCN200 level points. A great job was done with automatic labeling processes designing rules for each feature class with associated text reaching up to 70% of automatic labeling.

We also worked on a layout environment to generate printed maps to check results based on the original 1994 Atlas National format; 67 x 45 cm. This layout configuration had a map window template, sheet title, an overview map, the legend, scale, etc. Using a plotting tool that integrates the layout environment defined and the feature class containing the sheet boundaries we could print the sheet choosed.

Picture 2. ME500 pilot project printed test

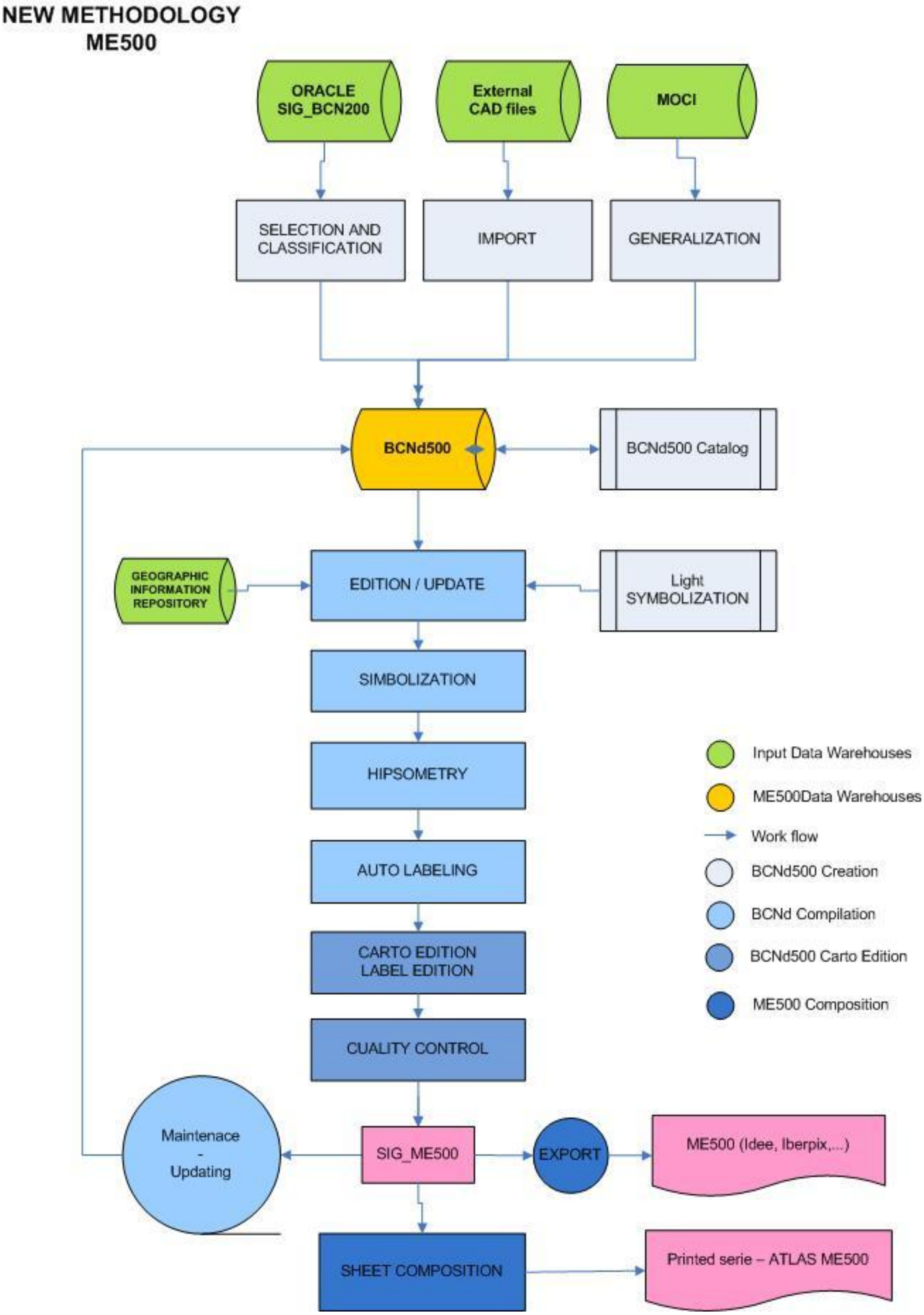
The platform used for this project was Geomedia Professional GIS software adding tools such as Fussion, Grid and Carto to validate, generate raster for relief and cartographic edition. We have also used FME software for generalization and format data change processes and Label EZ and Label Edit for automatic label placement.

REDEFINING THE PROJECT

As it was mentioned before, due to the inconsistence and outdated data source and added effects of generalization processes, we decided to redefine the project. At some point we decided to take a break in order to search for other different ways to carry out the correct realization of ME500. We checked the viability in order to integrate other sources as a reference data for this project and we wrote the specifications to give a formal definition to all processes to be addressed.

We analyzed the data given by one outsourcing company. This data was delivered on CAD files including geometries and texts. We also analyzed the vectorial data from Infraestructure and Roads Official Map database at scale 1:250.000 (called in Spanish MOCI) and we decided to use both of these in addition to BCN200 features that did not imply neither geometry validation nor geometric generalization processes, such as point features.

We redefined the project in this way, integrating three different data sources by designing a new database to support ME500, the BCNd500 with new feature class catalog following rules applied on the new BTN25 and BCN200 data models and obtaining a catalog based on 25 feature class with 28 tables loaded on the physical database. Next picture shows the new workflow for the ME500 compilation:



Picture 3. ME500 project workflow. New methodology.

It was not the ME500 formal workflow that should have been created out of the whole BCN200 product but we needed an updated version of it and we had to adopt this alternatives.

We were creating BCNd500 database in this way, by selection and classification from BCN200, importing CAD files from outsourcing ME500 and generalization from MOCI database.

ME500 INNOVATION PROJECT: CARTOGRAPHIC SPECIFICATIONS

The technical definition to elaborate a new GIS based Map of Spain at scale 1:500.000 from BCN200 GIS data source and other valid data sources are documented in these specifications. They detail the technical procedures that appear in the main workflow shown in the Picture 3. These specifications include appendix with documents such as ME500 feature class catalog that database BCNd500 was based on.

1. COORDINATE SYSTEM. PROJECTION

We have two different geographic areas in Spain, one area is the Peninsula (inside European Plate), Balearic Island and African territories Ceuta and Melilla and another area is the Canary Island. We loaded all the data in a continuous space, with the new adopted reference system ETRS89, in the same database in latitude and longitude coordinates.

It defines two different projections to use depending on the generated product. UTM projection will be used in digital products like Spanish SDI or Iberpix. The second one is the Lambert conformal conic with three different setups, two bands: north and south for the Peninsula, Balearic Island, Ceuta, Melilla and the third one for Canariy Island. This other projection will be used for National Atlas printed product

2. - SOURCES AND CARTOGRAPHIC ENTITIES OF THE MAP

It describes the data sources to create BCNd500: BCN200, ME500 CAD files and MOCI database. It details reference sources to use for compiling this map: neighbor country cartography, Official sources to check and update names, for example Local Entities Registry (REL) or IGN official names list and cartographic products to check and update geometries using SPOT images at scale 1:100.000 used in derived cartography.

In the second place it includes a list with all of the ME500 cartographic entities, detailing the description, selection rules and official reference sources used for checking and updating entities, for example, as in below list, Mountain elevations:

ID TYPE	CATEGORY, GEOGRAPHIC ENTITY, (THEME, SUBTHEME, Entity class Type)	COMMON GENERIC	DESCRIPTION AND SELECTION RULES	DATA SOURCES OFFICIAL NAMES
...				
02	**OROGRAPHY**			
021	**CONTINENTAL OROGRAPHY**			
0211	**Mountain range**			
021101	1st Order Range (Sierra 1)	Range, mountain range	Important Peninsular Orographic structures	Concise Nomenclator
021102	2dn Order Range (Sierra 2)	mountain range, mountains	Mountain ranges up to 30 km lenght	Concise Nomenclator
021103	3rd Order Range (Sierra 3)	mountain range, mountains	Mountain ranges between 15 and 30 km	Concise Nomenclator Catalonia Nomenclator
021104	4th Order Range (Sierra 4)	mountain range	Mountain ranges lower than 15 km	Concise Nomenclator Catalonia Nomenclator

Table 1. List and details of Mountain elevations cartographic entities

3. - BCNd500: Numerical Cartographic data Base derived at scale 1:500.000

It includes implementing rules to create database BCNd500 to support the logical and physical storage of the ME500 geographical information.

Main characteristics of stored information:

- Features are objects with homogeneous characteristics
- Base geometry is used to store information: point, line, area
- Explicit geometry is not considered:, data is structured at feature level and stored as they are
- Data are classified within six themes: Administrative boundaries, Relief, Hidrography, Transport networks, Populated Places and Buildings, Control points.
- Labels are extracted from LABEL attribute which is stored by its own feature.

Each feature class will be defined with different attributes but there is one generic feature class, common to all features with the same attributes (ID, DATE, ID_CODE, LABEL, GEOMETRY and GEOMETRY_SK) and TYPE_TTEE attribute (TT: theme code, EE: entity code) to define each feature type inside feature class.

An appendix of the specification includes a list of tables with all BCNd500 features with a complete description of name, theme, geometry type, feature types, type of map representation (geometry/text), short definition and content specifications with geometric rules. For example Natural parks:

FEATURE	2
FEATURE NAME	**BCNd500_0102S_ESPACIO_NAT**
BCN Theme	01: ADMINISTRATIVE BOUNDARIES
SubTHEME	
Geometry type	Area
TYPE	01:National Park 02: Natural Park
Label	YES
Map Representation	Virtual area geometry and Label
Descripción	Surface elements represented by extension and boundaries of National and Natural Parks.
Specifications	Store all National Parks and Natural Parks greater than 3.000 ha. When a limit of the park share geometry with other elements like administrative boundaries should be displaced toward exterior boundary an equidistant distance about double resolution of scale.

Table 2. BCNd500 catalogue: NATURAL PARK feature

This catalog includes a list of real features that describe the phenomena with its geometry considered at this scale and also other features that store virtual entities for the label placement of free text: this is related to the text without geometry in the map like the name of a range or a gulf. These virtual features are defined with line geometry which is used to place the text and store the feature TYPE and LABEL attribute to label the map.

Once the model and structure database is defined the data should be loaded. It has detailed the procedures to integrate each source: model generalization to integrate some BCN200 features, change format and importation from outsourcing CAD files and geometric generalization from MOCI database.

4. - CHARACTERISTICS AND QUALITY OF THE DATA

Data is organized in order to comply with the specific characteristics for the geographic information production of ME500 cartography. Geometric accuracy of data is validated with the difference between the position of elements on the map and the position of geographic phenomena in the source data used as a reference. BCNd500 is compiled to reach the 100 meters accuracy of the scale 1:500.000. We also considered the size of phenomena to be represented as the resolution of this scale. Geographic phenomena belonging to the same feature group has a separated representation with a minimum of 100 meters. The minimum area to be represented on the map should be greater than 6,25 ha (62.500 m2) in other case, should be deleted but sometimes, the feature has to be considered and converted to point geometry for example, many of populated places has been converted from areas to point features with the same feature class characteristics.

To ensure data quality, following recommendations by ISO:19113, it is considered the evaluation of the positional accuracy and thematic accuracy. The logic consistency and the completeness parameters are evaluated within compilation procedures.

Technical specifications also includes a feature class procedure manual which describes connection and sharing spatial relations between elements in the same feature class or between different classes for example, sharing relation between water reservoir and river area features or connection between all the road transport network.

5. - CARTOGRAPHIC COMPILATION

The processes in the cartographic compilation consists on preparing the geographic information stored in BCNd500. We have to debug, validate and update all this information to be coherent with its definition finishing with schedule controls to ensure quality of the data before edition tasks.

Geometric validation and updating are implemented through well described procedures organized by each feature class taking into account map priorities. Each feature procedure includes controls to be executed over geometries, also includes detected errors to edit and solve if an update is necessary and reference sources to validate and update data. For example, next table shows the validate procedure for the airport feature:

ValidationProcedure	11
BCN THEME	06:TRANSPORT NETWORK
FEATURE	BCNd500_0606P_AEROP_AEROD
Control	1- Complete data with reference source 2- Validate Classification for Airports and Aerodromes
Located Errors	1- Edit/insert/delete point if necessary 2- Update TYPE feature class attribute
Update	NO
Reference Sources	BTN25-BCN200 Map200, Atlas 1994 AENA (National Aerospatial Agency)

Table 3. Cartographic compilation: Validation procedure for Airports

A great effort has been done in the definition of the map content and in the selection of references sources to check and update BCNd500 geographic information. For example, the database at scale 1:1.250.000 is the first reference source because all data stored in this scale should be considered in the BCNd500.

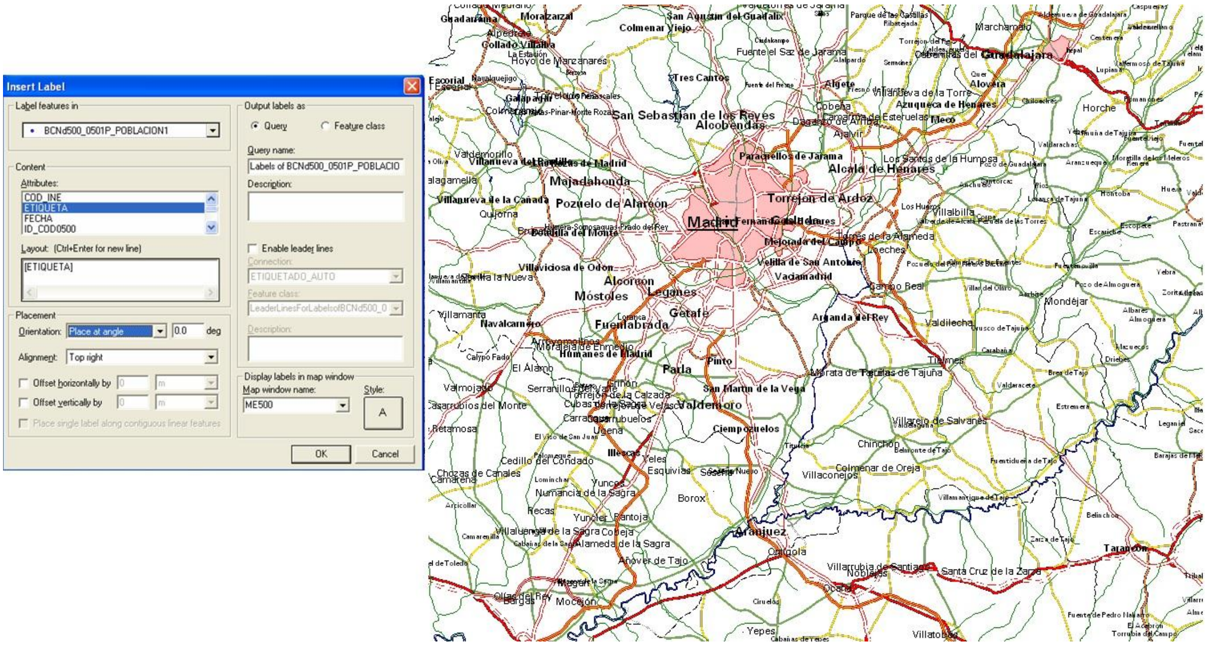
6. - CARTOGRAPHIC EDITION

It details procedures and tasks to do over a map workspace. In this phase of the project we have prepared the final symbols and the legend with priority rules definition for cartographic edition tasks. The text layer has been generated from automatic labeling software and from incorporating the generated raster hypsometry. Once we have prepared the map workspace, some tasks are defined and done, such a conflict resolution and updating semi automatic procedures with specific tools for cartographic edition.

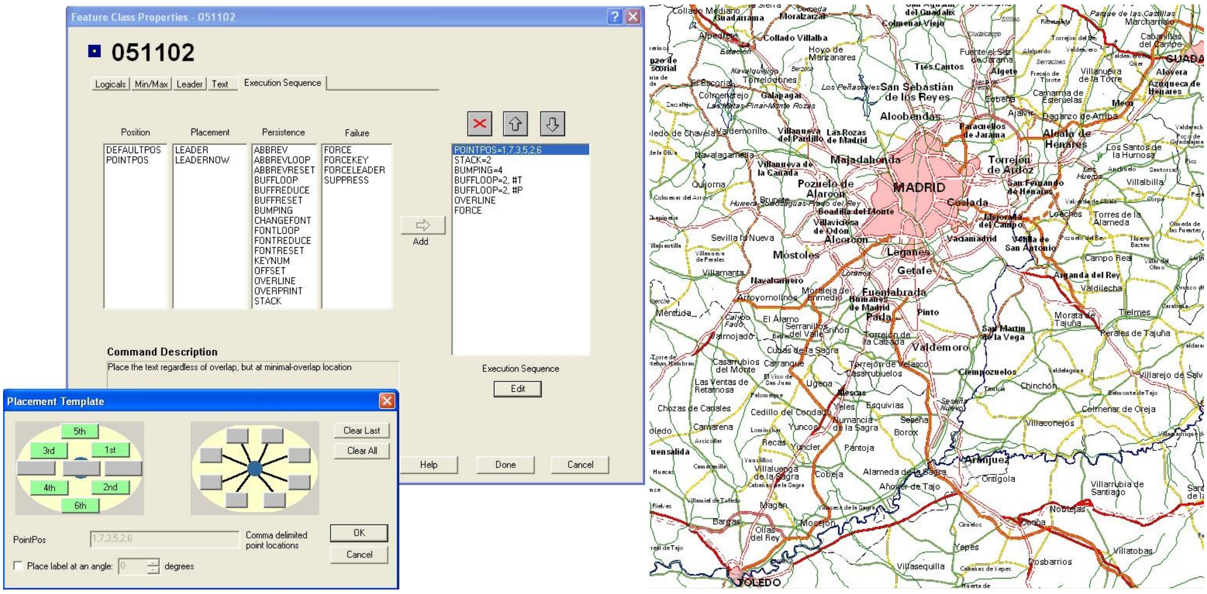
The text style management is remarkable. We have created a table that store the necessary style definition for each class of text and we have to setup an advance style management based on attributes. The text style on the map is obtained joining the feature class to be labeled with the “table of styles” and applying a “base style” which uses the attributes from “table of styles” (font, size, black…). This configures an easy way to apply styles stored in the same place that geometries.

In relation with hypsometry generation it is defined the pixel size of the raster layer because should be greater than 3/4 of scale resolution and final raster resolution should not be lower than 150 dpi to avoid perceive pixel size when printing. The selected pixel size was 75 m.

We have applied two different processes for label placement: one for the automatic labeling on features that has geometric and text representation on map, and another for virtual features where it is not necessary automatic label placement because the geometry is the guideline to place the text along this geometry. Automatic placement rules are defined for each feature class considered and some iterative processes were made to solve the best place, detecting overlaps and replacement texts.



Picture 4. Cartographic edition: static label placement and results



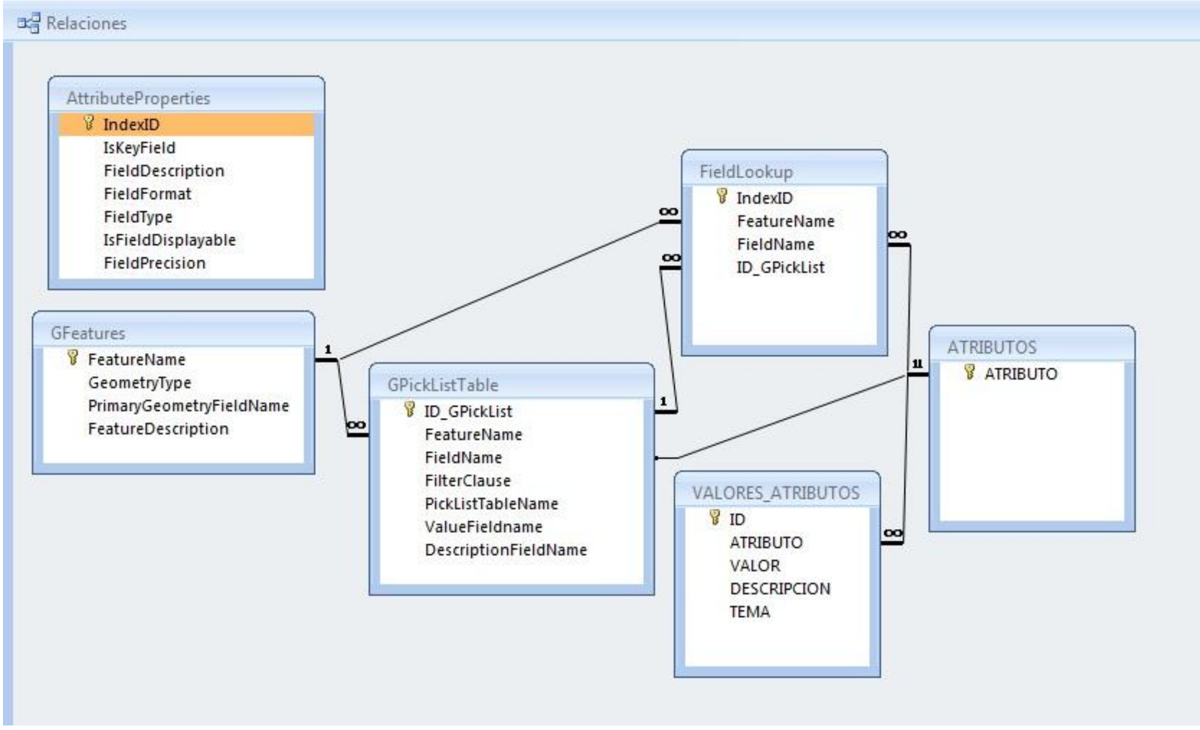
Picture 5. Cartographic edition: automatic label placement rules and results

7. - QUALITY CONTROL AND REVISION

To evaluate the quality in the generated geographic information we have considered checking and updating tasks to ensure logic consistency and completeness within compilation procedures. Positional and thematic accuracy has been considered to make a quantitative evaluation of quality in geometric and semantic characteristics. We will use different representative features class generating datasets from randomly sample of at least 20% of the total elements.

Within each feature compilation procedure it is included the necessary controls to debug and update each feature class iteratively, so we check geometries, attributes, topology, completeness and proceed with edition and updating, finishing with data validation. Geometric and semantic quality controls has been defined and applied with validation tools to check for example, empty geometries, zero-length lines, area loops, loops in line, dead ends, overshoots, undershoots, geometries under resolution, null attributes assignation. We have structured and validated connectivity and sharing geometries relations and checking completeness based on rules defined for this scale taking advance of query and analysis tools provided by GIS software.

Once compilation processes are completed we validate the logic consistency of the database, checking and updating the model definition, feature names, attribute names, attributes definition and it is built and integrity relation model providing coherency to the implemented data model.



Picture 6. Logic consistency: data model integrity

Within edition processes a manual revision is carried out from printings. This manual revision is based on check each element represented on the map, interactions between elements, name errors, etc. This revision is the final control and updating made to this cartography.

8. - METADATA

To facilitate the maintenance and updating of the information contained in the Map of Spain 1:500,000 scale, we generated the metadata to describe: the characteristics of the product, the data sources, processes applied to the sources, entities responsible for their preparation, etc.

Metadata must be consistent with Metadata International rules ISO 19115:2003, ISO/FDIS19115-2 e ISO/TS 19139:2007 especially with Metadata Spanish Core (NEM) and INSPIRE Implementation rules.

This document is still not ready because metadata should be generated once the project has completely finished representing final product.

RESULTS

We considered: specifications, production, delivery, use and update as the life cycle of the project. Now we are finishing the production phase just ready to deliver and use the new ME500 GIS based cartography. The final product will be a digital georeference image to represent on a map all of Spain at scale 1:500.000 with a symbology applied to represent two-dimensional space and a raster hipsometry with shaded relief to represent the elevation. Next picture will give us an idea of the final result.

Picture 7. Screenshoots from ME500 workspace.

Once we have realized the quality evaluation of the considered quantitative variables, we can show some of the results to show an idea about ME500 positional and thematic quality. Results from the analyzed random datasets are shown in the following tables. Table 4 shows the description about quality evaluation and Table 5 shows the calculated values.

Element	Positional Accuracy. RMSE	Positional Accuracy. Percentage	Thematic Accuracy ATTRIBUTE TYPE	Thematic Accuracy ATTRIBUTE LABEL
Subelement	Absolute	Absolute	Classification correctness	Non-quantitative attribute correctness
Scope	Random dataset	Random dataset	Random dataset	Random dataset
Measure Desc.	RMSE	Percentage of items with coordinate error greater than specification limit.	Percent correctly classified	Percentage of items with incorrect geographic names
Eval Method Desc.	For each node, measure the error distance between absolute coordinate values of the node in the dataset and those in BCN200. Compute RMSE from the error distances.	For each node, measure the error distance between absolute coordinate values of the node in the dataset and those in BCN200. Count the number of the nodes whose error distance exceeds the specification limit (100 m). Divide the number of the nonconforming nodes by the number of the nodes in the data quality scope. Multiply the result by 100.	For each item in the dataset, compare the assigned class against true class BCN200. Count items which are correctly classified. Divide the results by the total number of the items in data quality scope and multiply it by 100.	Compare the geographic names in the data quality scope against those in BCN200. Count items with incorrect geographic names. Divide the result by the total number of items in the data quality scope and multiply it by 100.
Value Type	Number	Percentage	Percentage	Percentage
Value	(Table 5)	(Table 5)	(Table 5)	(Table 5)
Unit	Metre	Percentage	Percentage	Percentage
Date	March 2011	March 2011	March 2011	March 2011
Conformance level	<= 100 m	<=5%	Less than 5% items may have incorrect class type	Less than 5% items may have incorrect names
Quality result meaning	RMSE of distance of the Nodes.	% of the nodes within the data quality scope have error distance more than 100 m.	% of Items misclassified	% of Items with errors in names.

Table 4. Quality Evaluation: positional and thematic accuracy legend.

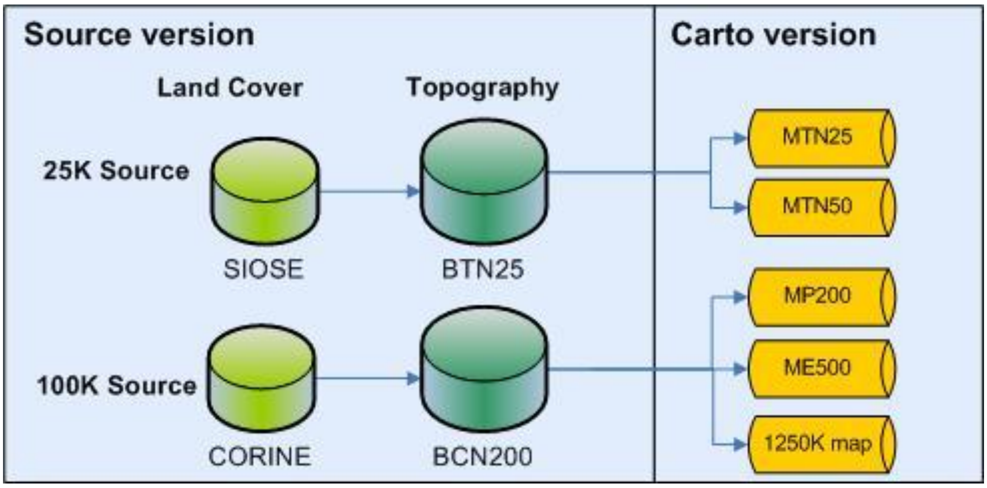
FEATURE	Positional Accuracy. RMSE	Positional Accuracy. Percentage	Thematic Accuracy ATT. TYPE	Thematic Accuracy ATT. LABEL	Eval method description of Positional accuracy	Quality result meaning
S_POPULATION	47	0%	2%	4,8%	Distance between boundary point well indentified with source.	Dataset pass test.
P_POPULATION	151	53%	1%	0,5%	Distance between centroids	Dataset fails positional test but pass thematic
L_ROAD	162	51%	1%	3,5%	Average distance between phenomena.	Dataset fails positional test but pass thematic

Table 5. Quality Evaluation: positional and thematic accuracy results (Value).

To evaluate positional accuracy we have considered the cartographic features selected from ME500 with the applied symbology. Each feature can be displaced from each other the necessary distance in order to distinguish different elements on map in relation with the resolution parameter. Population areas were not displaced but points elements of this feature are displaced in relation to roads and roads are displaced in relation to others: roads, hydrography and rail roads. BCN200 is the data source compiled from an orthophoto image with a positional accuracy of 40 m but ME500 is a cartographic product and features can be displaced from real positions. We should have selected for example, hydrography (elements from this feature cannot be displaced) to evaluate positional accuracy but this feature is not still ready updated on BCN200.

MULTIREPRESENTATION AT DIFFERENT SCALES

One of the main tasks in ME500 workflow is to have completely tested a multiscale production and updating environment. The different scale ranges of the cartography produced at IGN-Spain varies from 1:25.000 to 1:1.250.000 and smaller, such variety of scales and databases associated needs a solid link between them, if possible, to keep the coherence among the geographic objects contained in it. The goal is to have at least two main geographic databases as reference, mainly BTN25 and BCN200, so called "source version", and derivate several mapping products from them. Such derived products ("carto version") are stored as edited geometry but the original geometries are kept untouched but the link between them is always alive and guaranteed. Such derived products can be not only cartographic versions (with cartographic conflicts solved) but even generalized if needed, or combination of them. This process has a high cost at first time, but any update of the "source version" is clearly marked in the "carto" ones derived from them.



Picture 8. Multirepresentation at different scales environment in IGN Spain.

CONCLUSIONS

The ME500 project started with two purposes: to update this cartography and to test the new GIS based cartographic technologies. This cartography will be completely updated and finished in a few months but the goal of this project has been the implementation of a "test cartography workbench" on a GIS environment and the experience obtained. This project represents a change in the digital cartographic production for the IGN of Spain, going from discrete CAD files to a centralized and seamless geographic information management taking advantage of the new tools adopted.

GIS software has provided the tools to easily integrate the different data sources and a complete query and analysis environment to edit, update and use the information. It also makes the composition tasks easier through a designed layout environment, allowing "maps on demand" from a seamless geographic space.

We are working on 25K and 200K source databases compilation, developing at this moment procedures and specific tools to customize corporative software using the experience gained in the ME500 project. It is too early to think in carto products from 25K but they have been already tested on 200K for the Spanish Map at scale 1:200.000. ME500 is based on its own database as it is shown in this paper but the next update will be a "carto version" from BCN200.

REFERENCES

[1] ISO 19113:2002. Geographic information - Quality principles

[2] ISO 19114:2006. Geographic information - Quality evaluation procedures

[3] National Topographic Database 1:25.000 (BTN25) specifications and feature catalog. National Geographic Institute, 2010.

[4] Numerical Cartographic Database 1:200.000 (BCN200) specifications and feature catalog. National Geographic Institute, 2010.

[5] Spanish Spatial Data Infrastructure (SDI). http://www.idee.es

[6] Spanish viewer IBERPIX: http://www.ign.es/iberpix/visoriberpix/visorign.html

[7] Spanish Core Metadata: http://www.idee.es/resources/recomendacionesCSG/NEM.pdf

[8] Spanish Map 1:500.000, Cartographic rules. National Geographic Institute, 2006.
[9] SIOSE land cover: http://www.siose.es/siose/documentacion.jsp
[10] CORINE land cover: http://www.eea.europa.eu/publications/COR0-landcover