Cartography in Central and Eastern Europe

Selected Papers of the 1st ICA Symposium on Cartography for Central and Eastern Europe

Bearbeitet von
Georg Gartner, Felix Ortag

ISBN 978 3 642 26158 9
Format (B x L): 15,5 x 23,5 cm
Gewicht: 896 g

Zu Inhaltsverzeichnis
schnell und portofrei erhältlich bei

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.
Chapter 2

EuroRegionalMap –
How to Succeed in Overcoming National Borders

Andreas Pammer\(^1\), Anja Hopfstock\(^2\), Anita Ipša\(^3\), Jarmila Váňová\(^4\), Igor Vilus\(^5\) and Nathalie Delattre\(^6\)

\(^1\) Bundesamt für Eich- und Vermessungswesen, Vienna, Austria
  andreas.pammer@bev.gv.at
\(^2\) Bundesamt für Kartographie und Geodäsie, Frankfurt am Main, Germany
  anja.hopfstock@bkg.bund.de
\(^3\) Geodetski inštitut Slovenije, Ljubljana, Slovenia
  anita.ipsa@gis.si
\(^4\) Český úřad zeměměřický a katastrální, Sedlčany, Czech Republic
  jarmila.vanova@cuzk.cz
\(^5\) Državna Geodetska Uprava, Zagreb, Croatia
  igor.vilus@dgu.hr
\(^6\) Institut Géographique National, Brussels, Belgium
  e-mail: nde@ngi.be

Abstract

EuroRegionalMap (ERM) is a pan-European multifunctional topographic reference dataset at scale 1:250 000 based on national contributions from EuroGeographics member organisations. In one of its regional production groups data providers of eight Central European countries faced the challenges of harmonizing their national medium-scale databases in matters of content, geometry and quality standards.

This paper presents a progress report about the chosen approach and the realization of this multinational co-production. In this context the used project organisation and the given ERM specifications and guidelines are introduced. Thereafter the different national ways and means of deriving an ERM dataset by converting national data to ERM are exemplarily shown. Furthermore, to demonstrate a best practice approach the edge-matching of data at the cross-border area of countries to reach a seamless dataset is presented in more detail. The last focus is set on how
quality assurance is achieved. In a validation phase all data is checked by several validation procedures to identify errors in the data’s logical structure as well as to insure feature/attribute compliance with the current ERM specifications. On-going changes and future developments complete the paper.

**Keywords:** EuroRegionalMap, topographic dataset, harmonisation, edge-matching, pan-European

### 2.1 Introduction

EuroGeographics with its 52 members from 43 countries represents nearly all European National Mapping and Cadastral Agencies (NMCAs). Currently, four pan-European geographic datasets are available as products that were produced in project cooperation by the NMCAs (EuroGeographics 2009). Among them EuroRegionalMap (ERM) is the multi-functional topographic reference dataset at scale 1:250 000. ERM data is intended to meet two objectives. At the one hand the data should be fit for geospatial analysis in GIS applications, on the other hand their usage as topographic map backdrop for thematic mapping is envisioned. In Figures 2.1 and 2.2 ERM data is visualized in different scales to give an impression of the topographic content.

Looking back in the history of ERM the first product distributed on the market as release 1.0 in June 2004 was only covering seven countries (France, Germany, Belgium, Luxemburg, Denmark, Ireland and Northern-Ireland), but demonstrated the feasibility of harmonising the medium scale topographic databases owned by the NMCAs at a European level. In a next step a new project phase called the “extension phase” of ERM was established aiming to extend the coverage to at least the European Union and the EFTA countries. During that phase the organisational bases as well as the management and production workflow for an enlarged production of EuroRegionalMap over Europe were set up. The harmonisation approach, the data model and specifications were adapted and enlarged to include the contributions of the new project members. Furthermore a new management and production organisation was established setting up regional groups of production and quality assessment. As a result of these efforts, release 2.1 was published in March 2007.

Currently, the project is facing the “consolidation phase” that will end in December 2010. There are two main objectives: to provide regular updates and to enhance data quality and improve the level of harmonisation (Hopfstock 2008). The detailed update plan as well as further changes in the specifications are agreed and prioritised with the main customers, the European Commission with its sub-organisation Eurostat and the European Environment Agency.
Fig. 2.1. ERM data visualised in the cross-border region between Austria, Czech Republic and Germany at scale 1:250 000

Fig. 2.2. ERM data visualised in the cross-border region between Austria and Slovenia at scale 1:100 000
In the following the organisational and technical framework is introduced in more detail and a progress report about the chosen approach and the realization of this multinational co-production is presented from the perspective of producers in the Central-European regional group.

### 2.2 Organisational Framework

In the first place the project management structure which is working at four different levels with different organisations involved has to be mentioned.

- **The EuroGeographics Head Office (EUROHO):** This central permanent unit of the association of the European NMCAs acts as the contracting party and contract manager for the EC (Eurostat).

- **The Project Manager** at the National Geographic Institute of Belgium (IGNB): IGNB works as the project coordinator and is responsible for the overall project management and tasks like project awareness, product exploitation and the maintenance of EuroRegionalMap data. IGNB is also responsible for the final data collection of the ERM national data components and for the integration of the data into a seamless dataset, which is delivered to customers and data providers. IGNB is supported by the Project Management Team.

- **The Project Management Team (PMT):** The PMT is subdivided into
  - **The Technical Team:** It is composed of staff members of the National Geographic Institute of France (IGNF), the National Geographic Institute of Belgium (IGNB), the Bundesamt für Kartographie und Geodäsie of Germany (BKG) and the Bundesamt für Landestopografie of Switzerland (swisstopo) acting as technical experts. They are responsible for the technical development of the ERM production workflow and for the technical support to the producers by developing production tools, providing technical guidance for the producers, updating specifications and the data model and addressing any issues in production and data maintenance reported by the regional co-ordinators.
  - **The Regional Coordinators:** Owing to the amount of producing countries (32), the supervision of the production has been shared between four regional coordinators, which are the National Geographic Institute of Belgium (IGNB), the National Geographic Institute of France (IGNF), the National Centre of Geoinformatics and Remote Sensing of Lithuania (GIS-Centras) and the Federal Office of Metrology and Surveying of Austria (BEV). The Regional Co-ordinators are responsible for the regional co-ordination including the supervision of the production done by a group of partners and the final data
validation and data assembly at regional level. The area of responsibility for each regional co-ordinator is presented in Figure 2.3.

– The Business Manager: The Business Manager from the EuroGeographics Head Office is responsible for all issues related to communication and marketing on behalf of the project partners.

• The ERM data Producers (NMCAs): Producing NMCAs are responsible for providing the EuroRegionalMap data of the national territory and of other territories that might be under their production responsibility.

The members of the different management levels have to play their roles according to the general production workflow described below (The different steps mentioned in this overview are described in more detail in the following chapters):

NMCAs harmonise their data by re-engineering their existing national data collections according to the ERM specifications and data model and update the ERM data model according to the agreed update plan. Each national producer is responsible for the conversion and upgrade of the national database of its national territory into ERM and for the update and maintenance of this national part of the ERM dataset. The production limits between produced areas are represented by international boundaries, which have been commonly agreed and fixed between both

---

**Fig. 2.3.** Production area and the division in regional groups
neighbouring producers and approved by the project co-ordinator. Project partners of neighbouring countries should collaborate to get cross-border consistency of the data at both sides of the international boundaries. National producers have to check the quality of the dataset according to the Validation Specifications before delivery to regional co-ordinators. After the validation and correction phase the national datasets including delivery report and metadata are sent to regional co-ordinators who provide the final validation and data assembly at regional level. Finally the project manager integrates all regional results into a seamless pan-European dataset.

The used communication tools complete the organisational framework. As the project is still developing, the producers need to be informed about changes as soon as possible. ERM producers get information about news, events and updates in documentation via e-mails and they are available at the EuroGeographics website, too.

An important communication tool is the ERM Discussion Forum. Its main function is to notify any technical issue or question raised for open discussion or comments. This system allows getting a quick answer from the technical team in a way transparent for every member. All information related to the project, i.e. contacts, news, events, production plans, agreements, templates, data, documentation, tools, metadata, are stored in ERM Repository on IGNB web server.

In the Central-European regional group (group D) data exchange during the edge-matching and validation phase is realized by means of Projectplace, a web based project management tool for downloading and uploading data and documents (Projectplace, 2009).

Meetings and workshops represent another way to support co-ordination between data producers or inform about substantial changes in specifications. They are organised either by the project manager (for all producers) or by the regional co-ordinators (for producers within one production group).

2.3 Technical Framework

The ERM specifications and data model, technical guidelines, and database templates build the technical framework of the multinational ERM co-production to ensure the harmonisation of the existing national data collections (Hopfstock et al. 2007a, Hopfstock et al. 2007b).

The ERM specifications define the data model, the spatial reference, the positional accuracy, and the data content. The feature and attribute catalogue includes portrayal and quality criteria. Table 2.1 states the basic parameters of the specifications.
Table 2.1. Basic Parameters of the ERM Specifications

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>1:250 000</td>
</tr>
<tr>
<td>Data Model</td>
<td>Georelational based on DIGEST</td>
</tr>
<tr>
<td>Geodetic Datum</td>
<td>ETRS89 (~WGS84)</td>
</tr>
<tr>
<td>Coordinate System</td>
<td>Geographic Coordinates in Decimal Degree</td>
</tr>
<tr>
<td>Positional Accuracy</td>
<td>125m</td>
</tr>
<tr>
<td>Themes</td>
<td>Administrative Boundaries, Hydrography, Transportation, Settlements, Vegetation, Miscellaneous, Named Locations</td>
</tr>
<tr>
<td>Delivery formats</td>
<td>ArcGIS Geodatabase, Shapefiles</td>
</tr>
</tbody>
</table>

The ERM conceptual model is a geo-relational data model based on the DIGEST vector data model. The basic topological relationships are set up at the level of geometric primitives (node, edge, face), and can be described as edge-to-node, face-to-edge, and node-to-face.

The ERM data is made available in geographic coordinates (latitude and longitude) with a decimal fraction based on the European ETRS89 spatial reference system. At scale 1:250 000 this corresponds to the world reference of WGS84. The spatial reference is chosen to support both transformations from national coordinate systems to ERM as well as from ERM into European coordinate systems for cross-border applications.

In general, the quality of a dataset is defined by the positional accuracy and the completeness of the data. According to the ERM specifications, national contributions shall be consistent with topographic maps at scale 1:250 000 in terms of positional accuracy and completeness. Features shall be provided with a positional accuracy at least of 125m and a horizontal geometric resolution of 5m. The minimum length of a line feature shall be more than 50m. Area features less than 0.06km² shall be merged or omitted. These definitions are guidelines for data collection and generalisation of ERM data, but do not overwrite national production rules. Thus, differences in positional accuracy and completeness of ERM occur.

The ERM data is organised into seven data layers (themes, see Table 2.1) containing several feature classes. A feature class shares the same geometrical type and a set of attributes. A feature class may hold one or more feature codes representing different geographic entities. Besides, the ERM database contains a number of related tables:

- EBM_NAM is related to the administrative areas holding the names of the administrative units;

\(^\text{1}\) DIGEST: Digital Geographical Information Exchange Standard is a NATO standard developed by the Digital Geographic Information Working Group (DGWIG) to support efficient exchange of digital geographic information among nations, data producers, and data users.
EBM_ISN is related with EBM_NAM string the national hierarchical level of the administrative units;

- SYMBOL_RAT is related to the Named Locations providing information for a classified symbolization of cartographic text for representation on hardcopy output at scale 1:250 000;

- FERRY_LINK is a cross-link table connecting the ferry stations with the ferry lines;

- ERM_CHR is a general table holding the information which ISO character sets were used to encode names attributes. This is important to display the national characters of the names attributes in a system that does not support Unicode encoding.

Figure 2.4 shows an extract of the ERM data schema.

The ERM feature catalogue contains mandatory and optional features and attributes. This concept on the one hand leverages the provision of a consistent core dataset by the producers, and on the other hand takes into account the geographical diversity and the particularities of the national data production within Europe (Hopfstock et al. 2007a).

The demand for a multilingual international dataset is met by providing two attributes for geographical names per feature. Geographical names of a feature can be populated in the first and – if applicable – second national/regional official languages. Additionally, the language code is stored according to ISO639-2/B for each name, and the names are transliterated into the ASCII-7bit format. This ensures displaying the names in a non-Unicode environment for all countries.

The ERM technical guide contains guidelines to support the re-engineering of the national databases according to facilitate geometric and semantic harmonisation. It provides guidelines and examples from best-practice for data capture, generalisation, and attribution.

![ERMDiagram.png](https://via.placeholder.com/150)

**Fig. 2.4.** EuroRegionalMap Data Schema (Hopfstock et al. 2007b)
The ERM edge-matching rules provide guidelines for handling features crossing or consistent with international boundaries to ensure the final pan-European dataset is seamless across borders. Meaning, there are no gaps or overlaps between country datasets, and the hydrographic and transportation network are continuous. For more information refer to Section 2.5.

The ERM specifications and data model describe the conceptual data model of the pan-European dataset. Into which physical data model the producers would like to implement the conceptual model depends on their national production system. However, the implementation may lead to differences in data structure and topology. Therefore The ERM PMT provides a physical implementation of the conceptual data model for the preferred data production and exchange formats (see Delivery formats of Table 2.1). Pre-defined empty datasets (i.e. ArcGIS Geodatabase Templates) are maintained to accommodate the ERM production and to support data quality, a hassle-free internal and external data exchange, and the final data assembly of the pan-European dataset from national contributions.

2.4 National Data Production

On the basis of the technical framework requirements different ways to produce the national ERM data can be located at the contributing NMCAs. The production workflows are determined by

- Resources and organisational constraints: Can production be performed in house or must it be outsourced?
- Data sources: Are there proper national databases in scale 1:250 000 or must data be generalized based on data sources in bigger scales?
- Data maintenance method: Is data maintained at the national database level or is the ERM data directly updated?

Below the national data production of five countries are described in more detail in order to give quite good samples for used strategies that represent the entire European approach.

2.4.1 National Data Production of Austria

2.4.1.1 Data Source

The main data source of the Austrian ERM contribution is the cartographic model 1:250 000 of Austria (KM250) that has been established and is maintained by the federal office of metrology and surveying (Bundesamt für Eich- und Vermessungswesen, BEV). This multi-purpose database is focused mainly on
the production of the Austrian Military Map in scale 1:250 000 (ÖMK250). For providing certain attributes accessory data sources are needed:

- Digital Landscape Model of Austria (BEV),
- Population census (Statistics Austria),
- Information about motorways and expressways (ASFINAG),
- Information about hydrography (Hydrographisches Zentralbüro).

2.4.1.2 Hardware, Software and Human Resources

ERM production is performed on standard PCs (Intel Pentium 4 CPU 3GHz, RAM 1GB, Windows XP) using ArcGIS 9.2 and the extension PLTS 9.2. For conversion tasks FME 2008 is also in use. Two people, one technical coordinator and one cartographer are actually producing the Austrian contribution to ERM.

2.4.1.3 Production Workflow

In principle the ERM dataset is automatically derived from KM250, which is regularly updated. For this purpose the structure of KM250 was adapted to the ERM specifications. If there are any changes in the ERM specifications corresponding adjustments in KM250 are always implemented. As a result KM250 contains some features and even feature classes that are only used for the ERM delivery.

Generating the ERM dataset is actually a transformation from the edited national database KM250 done with FME. Only edge-matching needs to be performed in an interactive way afterwards. Of course, quality checks attend all procedures.

2.4.2 National Data Production of Croatia

2.4.2.1 Data Source

Due to the lack of recent data in smaller scale, the main source for production of the ERM dataset is the large scale (1:10 000) topographic database (TTB) as the only up to date source of topologically correct and attributed vector data.

Data in TTB covers approx. 60% of the territory of Republic of Croatia, so the ERM dataset covers only that area where the data is currently available.

Another relevant data source used for ERM production is the Registry of spatial units which contains inner and outer administrative boundaries of Republic of Croatia.
2.4.2.2 Hardware, Software and Human Resources

For production of the ERM dataset a standard PC (Intel Core 2 CPU 2.13 GHz, RAM 2 GB, Windows XP) is used. Software used is ArcGIS 9.2, FME, AutoCad Map and GeoMedia (due to specific input data formats). The work on ERM production was shared in a three men team (1 technician, 2 experts).

2.4.2.3 Production Workflow

As described above, the main source of ERM data is the large scale topographic database. So the main work that has to be done is the generalisation of this primary dataset. After transforming from national coordinate system to reference coordinate system of ERM using FME software by applying a 7 parameters transformation, the generalisation is adopted to get in conformance with the ERM specification.

Generalisation is conducted in a step by step fashion keeping in mind the 1:5 rule, meaning that from the original data set in 1:10 000 scale, we have to generalise to 1:50 000 scale and then once more to 1:250 000 scale.

The resulting data set does not include cartographic generalisation (repositioning of objects) but only the generalisation based on the simplifying of the data model and of the individual objects. Other tasks related to ERM dataset production consist of remapping of the original dataset and edge matching – where required.

2.4.3 National Data Production of Czech Republic

2.4.3.1 Data Source

The main data sources are:

- Digital Landscape Model 1:250 000 (DLM250) – produced by the Military Mapping Service
- Fundamental Base of Geographic Data 1:10 000 – produced by the Land Survey Office

Ancillary data sources are:

- Register of Administrative Units (The Czech Statistical Office)
- Base Dataset of Administrative Boundaries (Land Survey Office)
- Database of Roads (The Headquarters of Roads and Motorways of the Czech Republic)
- Data of Agency for Nature Conservation and Landscape Protection of the Czech Republic
- Aeronautical Information Publication (Air Navigation Services of the Czech Republic)
2.4.3.2 Hardware, Software and Human Resources

The ERM project is running on standard PCs (Intel Core 2 Duo E8400 3GHz, RAM 2 GB, GPU NVIDIA Quadro FX 370 256 MB, Windows XP). ArcGIS 9.2 and PLTS 9.2 are available since 2008. There are 2 persons (1 expert, 1 technician) engaged in ERM project in the Land Survey Office of the Czech Republic.

2.4.3.3 Production Workflow

The DLM250 was transformed from WGS84 to the national coordinate system using a 7-parameter transformation. The data was modified with regard to the ERM Specifications and updated according to the national database at 1:10 000 scale and other ancillary sources. A positional accuracy was improved to satisfy 1:200 000 scale. After an editing and validation process the database was transformed to WGS84 and submitted to an edge-matching process. However the edge-matching had to be performed both in WGS84 (ERM) and in national coordinate system (national database).

The first final database became source for a new official national database at 1:200 000 scale. In future all topographic databases at a corresponding scale (including ERM updated versions) will be derived from this national database based on ERM data model and specifications.

2.4.4 National Data Production of Germany

2.4.4.1 Data Source

The German contribution to the ERM database is mainly derived from the most up-to-date version of the Production Database 250 (PD250), a topologically structured vector dataset at resolution 1:250 000. Additionally, the following ancillary data sources are used to enhance the geometry and attribution:

- VG250, a vector dataset with related tables containing all administrative units of all hierarchical administrative levels starting from state to the municipalities. The geometry of the borders follows the scale 1:200 000 concerning accuracy and resolution.

- GN250, the geographical names book.

2.4.4.2 Hardware, Software and Human Resources

The ERM production is performed with standard PCs (Intel Core 2 CPU 2.66 GHz, RAM 3.6 GB, Windows XP) equipped with ArcGIS 9.2, PLTS 9.2 and FME 2008. There are two persons (1 expert, 1 technician) involved in the production
of the German ERM contribution. The maintenance of the national topographic data is continuously executed by the production group of the department of geoinformation.

2.4.4.3 Production Workflow

The German ERM dataset is derived by converting the national data to ERM using extraction and conversion rules which are determined by the ERM specifications. A cross-reference documentation between PD250 and ERM is maintained to record the production, facilitate the maintenance and future update process of ERM contribution, and as a source for compilation of the ERM metadata. The national data is derived and transformed using FME. The resulting ERM database is validated and if necessary corrected using ArcGIS and the PLTS extension.

2.4.5 National Data Production of Slovenia

2.4.5.1 Data Source

The main data source is the National General Map of Republic of Slovenia at scale 1:250 000 (last updated in August 2008).

Additional ancillary data sources:

- Register of spatial units (The Surveying and Mapping Authority of the Republic of Slovenia),

![Fig. 2.5.](image)

Fig. 2.5. Part of the National General Map of Republic of Slovenia at scale 1:250 000 (© Geodetska uprava Republike Slovenije)
Register of public roads (Slovenian Roads Agency, Ministry of Transport),
Statistical databank (Statistical Office of the Republic of Slovenia).

2.4.5.2 Hardware, Software and Human Resources

For the ERM production standard PCs (Intel Core 2 CPU 2.13 GHz, RAM 2 GB, Windows XP) are in use. ArcGIS is the main program for loading the data into the ERM template geodatabase. Additionally other software for preparing the data was used.

Mainly three persons are involved in ERM production. One is the national coordinator from The Surveying and Mapping Authority of the Republic of Slovenia, which is the ERM contractor. ERM production was outsourced to the Geodetic Institute of Slovenia where two persons (1 technician, 1 expert) share the work.

2.4.5.3 Production Workflow

After collecting all the source data, it is generalized according to the ERM specifications and some basic topology editing is done. In the next step the data is transformed from national coordinate system to reference coordinate system of ERM. Then additional attributes are filled in and geometry is edited regarding the topology constraints. Finally data quality control and edge-matching with all the neighbouring countries is performed as described in Sections 2.5 and 2.6.

Two country specific advantages of the ERM production are worth mentioning. First, the main source, the National General Map of the Republic of Slovenia at scale 1:250 000, was generated in vector format completely new in 2005 (updated in 2008) and was at the same time harmonized with ERM specifications. Thereby almost all layers can be derived directly from this main data source; additional sources are needed mostly for some attribution and for the administrative units. Second, because Slovenia is a small country it is easier to handle the data volume and less automation is needed (some editing and corrections can be made manually).

Production deficiencies are mainly related to software restrictions. For instance in ArcGIS it is hard to define the 7 parametric transformation, that would be necessary to perform correctly the transformation between the national coordinate system and the ERM reference coordinate system. In the end this problem is solved by performing this transformation with other software. The second deficiency is that some validation checks cannot be executed, because some checking tools would only be available in PLTS. Therefore these checks can only be carried out by the regional coordinator.
2.5 The Edge-Matching Procedure

The objective of the edge-matching procedure is to get continuous and seamless data at cross-border zone implying a consistent selection of features at both sides of the international boundary.

2.5.1 Edge-Matching Rules

The edge-matching process intends to ensure a coherent spatial and attribute continuity between features at both sides of the international boundaries. It is necessary to fit together both the geometrical representation of the features and their attribute values or just significant attributes. The producers have to follow the “Edge Matching Rules” and respect the topological relationships between features. Obviously the new matched features must be in conformance with ERM Specifications.

The matching process is performed on the line features and on the area features crossing or overlapping the international boundary and on the point features located on the boundary. All producers follow the directives containing a set of rules for edge-matching the geometry and a list of attributes that need to be consistent at both sides of the boundary. The guidelines document gives specific examples how to match the geometry and attributes.

Exemplary rules for matching the geometry and attributes are:

• When a feature is stopped at the international boundary with no counterpart at the other side, the decision whether to erase it or to extend it should be decided in common agreement between neighbour countries (see Figure 2.6a).
• Two line features are automatically (without discussion) moved to match each other if they are of the same feature type and if the gap is no greater than approximately 125 meters.
• When a line segment (usually a watercourse) is consistent with the international boundary, the segment should have exactly the same geometry in both national components.
• The matching distance between borderlines of area feature is a maximum of 125m. If bigger distance is noticed, it is not needed to do the edge-matching except for the water bodies and islands.
• When the extended part to be added to the feature is less than 20ha (for area feature), this can be ignored and not added (see Figure 2.6b).
• When country A portrays a watercourse located on the international boundary as a line, while country B portrays the same watercourse segment as an area with a fictitious axis, then the watercourse segment should be commonly portrayed as a line feature (see Figure 2.6c).
• Features overlapping the international boundary must have the same attributes in both country datasets (see Figure 2.6d).

Fig. 2.6. Edge-matching rules


### 2.5.2 Edge-Matching Workflow

The edge-matching phase is performed at the end of the production workflow. The preferred approach is to perform an accurate edge-matching when both neighbouring datasets have already been submitted to a first checking phase by the regional co-ordinator. When the edge-matching is done, the data can be delivered again to the regional co-ordinator for a quality control of the edge-matching.

Preliminary the responsible persons of neighbouring countries contact each other and decide how to share the work to avoid duplicate efforts (see examples in Table 2.2). Most edge-matching work is in themes “Hydro” and “Transportation”, less in other themes. The sharing is done so that the amount of work is shared approximately equally between two neighbouring countries.

**Table 2.2.** Sharing the edge-matching work along international borders

<table>
<thead>
<tr>
<th>Border between Country A</th>
<th>Who is going to perform edge matching along border for specified theme</th>
<th>Country B</th>
<th>Hydro</th>
<th>Transportation</th>
<th>Vegetation</th>
<th>Settlement</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Austria, Slovenia</td>
<td>Germany</td>
<td>Austria</td>
<td>Slovenia</td>
<td>Germany</td>
<td>Germany</td>
<td>Czech R.</td>
</tr>
<tr>
<td>Czech R.</td>
<td>Germany, Czech R.</td>
<td>Austria</td>
<td>Germany</td>
<td>Czech R.</td>
<td>Germany</td>
<td>Germany</td>
<td>Czech R.</td>
</tr>
</tbody>
</table>

The actual edge-matching workflow is presented in Figure 2.7. The country responsible for edge-matching gets the latest data from the neighbour to perform edge-matching according to the directives. The country performing the edge-matching may only displace the features within its national territory. This can be systematically done when the matching distance is stated within the limits of the indicated parameters. If the edge-matching between features has to be done within a bigger distance, which requires the displacement of a feature belonging to the neighbouring country, the issue has to be notified in a report to the neighbouring country in order to discuss and negotiate the displacement of features.

After edge-matching by editing only national data, the responsible country provides a report with all remaining discrepancies to the neighbour. The discrepancies are divided in two types, changes and issues. Changes refer to necessary modifications of the neighbour dataset; issues represent unsolved edge-matching problems. The report should contain suggestions for a solution and a picture identifying the problem. Based on this report, discussion and negotiation can start between countries. The neighbour country sends back the report indicating the discrepancies that have been approved and/or corrected with a new version of corrected data or the discrepancies not approved and commented (see Figure 2.8). The responsible country will check again the data and restarts the procedure with reporting until all discrepancies will be solved and all changes approved.
Fig. 2.7. Workflow of edge-matching procedure

Fig. 2.8. Example of reported discrepancies at international boundaries
ERM producers can use automatic validation tools for checking points and lines. The ERM technical team has developed two Python scripts for checking line edge-matching and point edge-matching. Areas have to be checked visually. The principle of the automatic tool is to locate features within a very small distance from the international boundary and look if they have a companion feature in the neighbour country dataset. The scripts generate a list of suspect features, which have to be reviewed visually because not all findings are errors. The list contains the attribute values of matching features, too. When the edge-matching is done, the data can be delivered again to the regional coordinator for a quality control of the edge-matching.

### 2.6 Quality Assurance

For the purpose to ensure good data quality the ERM production process includes two independent validation controls. On one hand producers have to perform a comprehensive check after data production themselves in order to deliver clean data to the regional coordinator. On the other hand the regional coordinator validates the received data and reports about the results. If any severe errors have been found data has to be corrected by the producer and delivered to the regional coordinator again. Data may pass this cycle presented in *Figure 2.9* several times till an acceptable result is reached.

![Fig. 2.9. Validation process at ERM production](image)

The validation procedures (Hopfstock et al 2007a, Hopfstock et al. 2007b) consist of a series of checks against the following criteria:

- Compliance with the ERM Specifications regarding topology, data model, allowed attribute values, selection criteria and the geometrical resolution
- Coherence and consistency of feature and attributes
- Homogeneity of attribute values in a feature network
- Consistency between themes
- Cross-border continuity between neighbouring countries
The completeness of data collection is in the only responsibility of the producers, because the review by the regional coordinator is done without maps and ancillary sources.

Because producers work with different GIS and mapping software and additionally with different versions there is no directive for using certain validation tools to perform the checks. In contrast, the regional coordinators use the same validation tools for the specified checks to achieve a comparable quality standard for the whole ERM dataset. These validation tools can be classified in four groups:

- **Standard tools of ArcGIS:** checking i.e. topological rules or connectivities
- **Automated checks with PLTS GIS Data ReViewer:** checking i.e. allowed attribute values, attribute combinations, geometrical resolution or spatial relationships
- **Non standard tools of ArcGIS:** additional extensions of ArcGIS that must be installed separately and self-developed programs, i.e. Python Scripts for checking edge-matching
- **Visual quality checks:** for all those criteria which can’t be checked automatically i.e. generalization and selection rates

The regional coordinator reports the validation results both by means of an Excel-sheet stating the number of errors relating to the validation criteria (see Table 2.3) and additionally sending a ReViewer table.

### Table 2.3. Example of validation report (Excel sheet)

<table>
<thead>
<tr>
<th>Qualita criteria</th>
<th>Validation results of RC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Attributes structure</strong></td>
<td></td>
</tr>
<tr>
<td>allowed attribute values for FERRYL (DETN, DETA, DNLN, USE, RSU, FerryID)</td>
<td>ok</td>
</tr>
<tr>
<td>allowed attribute values for RAILRDL (EXS, FCO, GAW, LLE, NAMxx, RCO, RGC, RRA, RRC, RSU)</td>
<td>62 errors (LLE)</td>
</tr>
<tr>
<td>allowed attribute values for ROADL (EXS, LLE, LTN, MED, NAMxx, RST, RSU, RTE, RTN, RTT, TOL, TUC)</td>
<td>3 errors (TUC)</td>
</tr>
<tr>
<td><strong>Topology</strong></td>
<td></td>
</tr>
<tr>
<td>AIRFLDA must not overlap with HARBORA</td>
<td>ok</td>
</tr>
<tr>
<td>FERRYL, HARBORL, RAILRDL, ROADL, RUNWAYL: Must not intersect or touch interior</td>
<td>RoadL: 14 errors</td>
</tr>
</tbody>
</table>

The ReViewer table is a database table written by PLTS that contains the features that have been found using automated checks and visual review. This table contains a link (ObjectID) to the found feature and therefore, when using ArcGIS, errors can easily be corrected and evaluated as needed. In addition the ReViewer table contains several columns that provide information on the records such as the feature class, the used check and an error description (see Figure 2.10).
2.7 Conclusion and Outlook

This paper presented a progress report about the best-practice approach of the multinational ERM co-production for the central-European region.

A new release ERM v3.0 providing a regular update will be published in the first quarter of 2009. The ongoing consolidation phase aims at improving the data quality and, therefore, increasing the level of geometric and semantic harmonisation to meet the user requirements in spatial analysis and visualisation. Further on, the ERM specifications are reviewed in the context of the ESDIN project.

**ESDIN – European Spatial Data Infrastructure with a Best Practice Network** aims to develop, test, and implement services providing transformations to European data specifications, support for multi-lingual aspects, transformation of coordinates to common spatial reference systems, generalisation from ‘large’ to ‘small’ scales, and edge-matching to address inconsistencies at national borders (ESDIN 2009). Besides achieving technical interoperability, the project also addresses business interoperability seeking for simple data licensing, digital geo-Rights Management providing fast and easy user access to the data. Finally, data quality is a prime concern of users and ESDIN is looking into establishing a standard approach to reporting data quality.

In a long-term vision the ESDIN developments could lead to a distributed European Network running services that automatically join proper national databases in scale 1:250 000 at a European level and generate the EuroGeographics’ products, amongst others ERM, meeting the requirements of INSPIRE.
References

Cartography in Central and Eastern Europe
Selected Papers of the 1st ICA Symposium on Cartography
for Central and Eastern Europe
(Eds.) G. Gartner; F. Ortag
2009, XIX, 570 p. 236 illus., Hardcover
ISBN: 978-3-642-03293-6