Tua Nylén and Harri Tolvanen

Deliverable D.T3.2.1 Guidelines for harmonising, producing and managing cross-border spatial data

Cross-border case

28.7.2017

PLAN4BLUE

SYKE.FI/PROJECTS/PLAN4BLUE

CONTENTS

BACKGROUND

Spatial information forms the basis of spatial planning. It is present in the pre-planning stage, in the mapping of the current state of the planning area, analysing the interactions between interests and evaluating and communicating planning options. Compared to land-based spatial planning, Marine Spatial Planning (MSP) faces many additional challenges, including cross-border and data availability issues and the addition of the vertical dimension.

This guide aims at presenting a step-by-step approach to utilising spatial data in cross-border MSP. It seeks to complement existing MSP guides in a concise and practice-oriented way, instead of forming a comprehensive manual. Particularly, the guide refers to UNESCO IOC's guide "Marine Spatial Planning – A step-by-step approach to ecosystem-based management" (Ehler & Douvere 2009). This guide introduces each step of the MSP spatial data process in a general way and includes practical tips, arising from experiences of an MSP pilot project (Plan4Blue in the Gulf of Finland). It is prepared in four steps, conforming to the workflow of an ordinary MSP process:

- Guidelines for identifying and sharing relevant spatial data in MSP (ready in April 2017)

- Guidelines for harmonising, producing and managing cross-border spatial data (July 2017)

- Guidelines for spatial modelling and analysis to produce data and identify interactions of interests (January 2019)

- Guidelines for visualising spatial MSP information effectively on maps (June 2019)

The guide is prepared as part of the MSP development project Plan4Blue - Maritime Spatial Planning for Sustainable Blue Economies (2016-2019; http://www.syke.fi/projects/plan4blue) funded by the EU European Regional Development Fund (Interreg).

FREQUENTLY USED TERMS

CRS – Coordinate reference system

Data collection plan – A preliminary plan of which phenomena should be presented on maps, aimed at guiding data inventory and collection

Map plan – A preliminary plan of the map outputs and other products based on spatial data, aimed at guiding the data collection plan and creating a time table

MSP – Maritime/Marine spatial planning

PGIS – Participatory GIS, collecting spatial information in digital format using participatory methods

GUIDELINES FOR HARMONISING, PRODUCING AND MANAGING CROSS-BORDER SPATIAL DATA

Step 4. Collect data

(Related to "Step 5" and "Step 6" in Ehler & Douvere 2009)

4.1 Obtain existing data

Practical tips

- Utilise the data collection plan developed in Step 2.1 to guide the acquisition of existing data. Utilise the inventory of data providers (Step 3.1) to access data.

- Consider the use of OGC web interface services (e.g. WMS, WFS) which enable a direct access to updated data, provided by many national and international authorities (e.g. interfaces of HELCOM: http://www.helcom.fi/baltic-sea-trends/data-maps). When using web interface services, the data layers do not need to be archived as files. On the other hand, with using web interfaces, the version of the data layers cannot be controlled in a similar way as with working with files.

- In case interface surfaces are impractical, download and store spatial data files on hard drives and/or cloud services.

4.2 Derive new data

Practical tips

- Utilise the data collection plan developed in Step 2.1 to guide the work.

- In many cases, useful data can be derived from existing datasets.

- For example, business registers often include addresses (full street address, postal code or municipality) that can be used to geocode the postal addresses of firms (and create point data). GIS software typically have plugins for geocoding addresses using Open Street Map or Google Maps data as reference. The success of geocoding is strongly dependent on the quality of the address data.

- Point based population data etc. can be used to calculate population densities for administrative units (joining) or transformed to density surfaces of different types (point density, kernel density) for visualisation purposes. Land areas can be derived from sea area by using geoprocessing tools.

- Spatial modelling can be utilised to cost-efficiently construct new datasets based on limited observations. For example, limited observations of the species richness of seabed flora can be related to physical and chemical characteristics of the seabed. The more detailed data of these seabed characteristics can then be utilised to calculate to probable distribution of species richness.

- See a more detailed guide for spatial modelling and geoprocessing methods in the next part of the guidelines: "Guidelines for spatial modelling and analysis to produce data and identify interactions of interests".

Practical tips

- Utilise the data collection plan developed in Step 2.1 to guide the work.

- Resources for the collection of new data are typically minimal. Efforts should be targeted to largest gaps in the knowledge, which are often the silent knowledge of private actors and true location of economic and social activities on the sea (registered addresses of firms are often insufficient for example for analysis of human impact on the sea). Moreover, data production should aim at balancing the differences in the amount of information between administrative units and sectors.

- Utilise participatory GIS (**PGIS**) e.g. for locating activities that do not leave a mark in official datasets (such as fishing areas, tourist destinations), interests and future scenarios of experts. Many applications exist for collecting digital spatial data from participants either in workshops, in the field or online. PGIS questions can also be built in in a comprehensive questionnaire.

- Participatory planning on poster maps ("MSP board game") can be used in events (workshops, stakeholder meetings; Figure 1). Current activities or future scenarios can be marked on the map by drawing or settings pins or board game pieces on the fabric (see Figure 1 for some ideas). The resulting plan can be photographed and digitised into digital spatial data. In the Plan4Blue case, one spatial layer was created for point, line and polygon features, separately for the four main Blue economy sectors dealt with in the project (maritime cluster, blue bioeconomy and subsea resources, energy, tourism). Conflicts and "matches" between three groups were analysed.

Figure 1. Participatory planning ("MSP board game") in Plan4Blue cross-border workshop in June 2017. A basemap (restricted military areas, nature conservation areas, fairways, ports, pipelines, dumping sites, main roads and railroads, urban areas and place names) of the plan area was printed on a 151 cm x 91 cm canvas. The map was covered with plastic and sketched on. The drawings were later digitised to create digital spatial data, including point, line and polygon layers for each four blue economy sector (photograph: Anne Erkkilä-Välimäki).

Step 5. Harmonise data

(Related to "Step 5" and "Step 6" in Ehler & Douvere 2009)

5.1 Decide principles for data harmonisation

Practical tips

- Note that the following section mainly deals with technical harmonisation of spatial data. However, as discussed in section 3.3, discrepancies in the content of data products should be carefully evaluated and not hidden with harmonisation tricks. As an example, when collecting participatory GIS data of expert future visions, there may be large differences in the number of participants from different parts of the area. If there are two expert participants from Finland and 32 participants from Estonia, it will be inappropriate to summarise the results in a way that masks the spatial imbalance. If for example data from two sides of an administrative border originate from different decades, treating the datasets as equal would be misleading.

- Firstly, the main question for deciding the principles is if different parties accept each other's data. If they do, there may be no need to harmonise them. If they don't, the data could be harmonised in different ways to reduce the discrepancies. More importantly, this is the point were inadequate data should be identified and left out from analyses.

- Secondly, the aimed use of the data should be considered: if the data are only used internally for the analyses and map production of the MSP process, there may be no need to harmonise e.g. coordinate reference systems (CRS) and data formats. However, if the datasets are publicly redistributed, e.g. published in a geoportal for further use, even the more technical characteristics of the data should perhaps be uniform.

- Thirdly, analyses needs dictate the need for data harmonisation further. For example, some statistical analyses, geoprocessing methods or raster calculation require a uniform resolution of the data, a uniform cell size of raster data or a uniform CRS. These harmonisation needs can be fixed along the way through the process.

- Classify identified discrepancies to three groups: 1) those that do not require harmonisation 2) those that may require harmonisation for specific uses and 3) those that need harmonisation for all uses.

- Typical discrepancies that may not (depending on the specific conditions of the MSP process) require harmonisation include: CRS, data formats, naming the spatial data layers.

- Typical discrepancies that may need harmonisation: resolution (observation density and cell size) of the data, discontinuities in the data from different areas, data type (feature or raster) of the data and language and terminology of the attributes.

5.2 Resize data products

Practical tips

- Usually, the best available data originates from different sources (e.g. national or municipal) and is therefore spatially discontinuous. This may not be a problem when visualising the data on maps, but usually prevents all analyses. Geoprocessing methods can be used to "merge" feature layers of the same data type (point, line or polygon) or to "mosaic" raster layers.

- Often data supplied by officials of a larger administrative area contain large amounts of data outside the plan area. There may be a need to exclude those data either for visualisation purposes or for speeding up data analysis (larger data are more time-consuming to process). Feature or raster layers can be "clipped" to a specific extent for example by using a polygon layer of the plan area as a "clip feature".

- Moreover, spatial data layers may contain large amounts of features that are not useful for the MSP process or for a specific analyses. For example, only the sea areas included in a polygon layer of multiple land use classes, or only largest lakes may be needed for an analysis. These features can either be temporarily "filtered" from the layer or queried and permanently saved as a new layer.

5.3 Harmonise attribute tables

Practical tips

- Utilise the inventory of data discrepancies (Step 3.3).

- Attribute tables may contain discrepancies and errors and different classifications across national borders (even international datasets).

- When data from different sources are combined to create merged data layers, the attribute tables are not automatically usable. Unless all columns in all initial layers are identical in terms of column types and names, the new attribute table will include all columns from all initial layers and a large amount of empty cells. If the resulting layer is to be publicly redistributed, the entire attribute table should the harmonised and appropriate metadata prepared. If the resulting layer is only used for analysis or visualisation on a map, it may be more practical to calculate one or two new columns that combine information from several initial columns (e.g. a new column "Size" that contains information either from an initial column "Koko" (in Finnish) or "Suurus" (in Estonian), depending on the origin of a feature).

- Note that often local expertise is needed to interpret and translate the attributes. In addition, sometimes there may be notable differences in local classifications (of e.g. municipalities, degree of conservation status or types of fairway areas). Local expertise should be used to harmonise the attribute table in these cases, which will influence the selection of comparable classes across borders for visualisations and analyses.

5.4 Harmonise resolution

Practical tips

- Utilise the inventory of data discrepancies (Step 3.3).

- Utilise international (often larger spatial scale) data to smooth national boundaries and fill gaps in the EEZ. If this is done for visualisation purposes only, the individual data layers may not need to be physically merged into one. However, if a single layer is needed for analyses, geoprocessing tools are used to merge (feature layers) or mosaic (raster layers) the data. In the latter case the overlap of e.g. international and national datasets should first be removed. Moreover, the resulting layer should perhaps be resampled to a uniform spatial resolution, usually the one used in the coarsest dataset or at least the coarsest resolution should be documented. This is vital particularly if statistical analysis or modelling will be applied on the merged layer.

- Utilise spatial data analysis methods to convert existing data to different resolutions.

5.5 Harmonise data type

Practical tips

- Utilise the inventory of data discrepancies (Step 3.3).

- Utilise spatial data analysis methods to convert existing data to different data types and formats; i.e. between point, line and polygon feature types and between feature and raster data types.

5.6 Harmonise coordinate reference system and file format

Practical tips

- When necessary (e.g. the data will be publicly redistributed), the data can be reprojected to a uniform CRS. Usually it is practical to choose a CRS that is designed for a slightly larger area around the project area. For example, Finnish and Estonian national datasets are mostly prepared to a national CRS. When data are stored for use in the broader (in this case international) context, a practical choice is the same CRS that is used by HELCOM, an international organisation operating in the Baltic Sea region. However, for map visualisation and often also for analyses, the data do not have to been reprojected, since GIS software are usually able to handle data with different CRS's.

- When necessary, the data can be transformed from one file format to another. For example, file formats of less often used GIS software need to be converted to more universal formats before they can be viewed and analysed in another software. Tools for the transformations can often be found online or as plugins. When e.g. the data will be publicly redistributed, the file formats can be harmonised for the convenience of the users. Uniform file format for can be chosen from a variety of choices for raster (e.g. GeoTIFF, ESRI grid, binary file, JPEG2000) and feature data (e.g. xyz, MapInfo TAB, ESRI Shapefile, geodatabase).

Step 6. Manage data

 (Related to "Step 5" and "Step 6" in Ehler & Douvere 2009)

6.1 Manage file names

Practical tips

- Create a clear naming system for spatial data files. The file name should specify the origin, resolution and version of the data.

6.2 Manage coordinate reference system

Practical tips

- Use one CRS for all original datasets created in the MSP process. See section 5.6 for the choice of the most appropriate CRS.

6.3 Create metadata

Practical tips

- Create detailed metadata for original datasets created in the MSP process, as well as for heavily edited datasets that will be publicly redistributed. When possible, follow INSPIRE standard (INSPIRE Maintenance and Implementation Group 2017; for an INSPIRE metadata editor see page http://inspire-geoportal.ec.europa.eu/editor/).

6.4 Store data

Practical tips

- Store spatial data files both on hard drives and cloud services.

- Use database software to organise the data or work out a systematic structure for organising them on hard drives and cloud services.

- Keep copies of existing spatial layers, derived layers and original layers (created in the MSP process) separate.

- Keep your data archive up to date frequently, preferably using an automatic back up system.

REFERENCES

Ehler, C, Douvere, F (2009). Marine Spatial Planning: a step-by-step approach toward ecosystem-based management. Intergovernmental Oceanographic Commission and Man and the Biosphere Programme. IOC Manual and Guides No. 53, ICAM Dossier No. 6. UNESCO, Paris.

INSPIRE Maintenance and Implementation Group (2017). Technical Guidelines for implementing dataset and service metadata based on ISO/TS 19139:2007. 157 p. Inspire, Infrastructure for Spatial Information in Europe.