Analysis of the national ecosystem database of Bulgaria: (Mis)matches with the MAES framework

Gergana Petkova, Hristina Prodanova *, Vanya Stoycheva
National Institute of Geophysics, Geodesy and Geography – Bulgarian Academy of Sciences, Sofia, Bulgaria
Corresponding author: hristina.zh.prodanova@gmail.com

ABSTRACT

The mapping of ecosystems is a significant element in the European Biodiversity Strategy and the results of its implementation should support the maintenance and restoration of ecosystems and their services. The quality of the spatial data is of crucial importance for the achievement of these goals. A methodological framework for Bulgaria in the form of nine separate methodologies has been developed in recent years. In this paper, we analyze the ecosystem typology for Bulgaria and the GIS database to assess the possibilities to develop a common database for the needs of integrated water management. The data analyses were carried out in two dimensions: 1) the typology and attributive data were analyzed by cross-tabular approach; and 2) the spatial data were analyzed by topology rules. The results of the study reveal three main problems of the typology: 1) for some types it is developed to the fourth level while for others it is to the third level; 2) in some of the ecosystem types, especially in the freshwater ecosystems, different categories are mixed within a single hierarchical level; 3: there are duplicated numerical designations between grassland and forest ecosystems. This necessitates a revision of the typology and the development of a correct uniform classification to be used for the needs of integrated assessment. The topology analyses of the merged data from the eight ecosystem GIS layers show extremely large numbers of gaps and overlaps. The main reason is the use of different sources for the mapping of different ecosystem types. The main conclusion is that it is practically impossible to generate topologically correct integrated GIS layers from the eight ecosystem type layers. Therefore, it is necessary to develop a new approach for mapping all ecosystem types into a uniform database.

Key words: ecosystem services, INES, integrated water management, MAES, topology

1. Introduction

The mapping of ecosystems is a significant element in the European Biodiversity Strategy which sets a target to EU member states to map and assess the state of ecosystems and their services in their national territory. The results of this mapping and assessment should support the maintenance and restoration of ecosystems and their services. The quality of the spatial data is of crucial importance for the achievement of these goals. A methodological framework for Bulgaria in the form of nine separate methodologies has been developed in recent years. In this paper, we analyze the ecosystem typology for Bulgaria and the GIS database to assess the possibilities to develop a common database for the needs of integrated water management. The data analyses were carried out in two dimensions: 1) the typology and attributive data were analyzed by cross-tabular approach; and 2) the spatial data were analyzed by topology rules. The results of the study reveal three main problems of the typology: 1) for some types it is developed to the fourth level while for others it is to the third level; 2) in some of the ecosystem types, especially in the freshwater ecosystems, different categories are mixed within a single hierarchical level; 3: there are duplicated numerical designations between grassland and forest ecosystems. This necessitates a revision of the typology and the development of a correct uniform classification to be used for the needs of integrated assessment. The topology analyses of the merged data from the eight ecosystem GIS layers show extremely large numbers of gaps and overlaps. The main reason is the use of different sources for the mapping of different ecosystem types. The main conclusion is that it is practically impossible to generate topologically correct integrated GIS layers from the eight ecosystem type layers. Therefore, it is necessary to develop a new approach for mapping all ecosystem types into a uniform database.
The quality of the spatial data is of crucial importance for the achievement of the above-mentioned goals. A methodological framework for Bulgaria has been developed under the MetEcosMap project (Methodological assistance for ecosystem assessment and biophysical valuation). It includes nine separate methodologies (Apostolova et al., 2017a, 2017b; Kostov et al., 2017; Karamfilov et al., 2017; Sopotlieva et al., 2017; Uzunov et al., 2017; Velev et al., 2017; Yordanov et al., 2017; Zhiyanski et al., 2017) each of which covers a specific ecosystem type according to the MAES typology. Based on these methodologies, an assessment and mapping by ecosystem types was performed for the territories of the country falling outside the NATURA 2000 network. The INES (Integrated assessment and mapping of water-related ecosystem services for nature-based solutions in river basin management) project is set up to develop a methodological framework for mapping, modeling and evaluation of water-related ecosystem services in order to implement nature-based solutions (NBS) in water management activities. One of the main objectives of the project is to create a common spatial database for the main ecosystem types, linked to the global classification of ecosystems and the ecosystem accounts scope (extend account). In order to achieve this objective it is necessary to analyze national typology and the database of the ecosystems in Bulgaria. The analyses are performed in two case studies representative for the terrestrial ecosystems in the country: the Ogosta river basin and the city of Sofia. The database contains data for nine ecosystem types found in the country. The main goal of this paper is to present the first results of the ecosystems database analyzing the possibilities to develop a common database for the needs of integrated water management.

2. Database analyses

The data analyses were carried out in two dimensions: 1) the attributive data and the classification itself, published in the methodologies, were checked including its hierarchical levels, their names, and the separation criteria; and 2) the spatial data was reviewed - geometry, the presence of gaps and overlaps between the layers and between the graphical elements within them. The results of the validation are reported and commented in the following subsections of this short communication. The results are part of a bigger work as presented on Fig. 1. Here we focus on the analysis of the ecosystem data – both the attribute data, and the spatial one.

The attribute data analyses were performed simultaneously in two items: 1) in Excel sheet (Fig. 2), where on the one hand all available data for the database, for the methodologies classification, and for the EUNIS names were entered; and 2) in a Geographic Information System (GIS) - in ArcMap and ArcCatalog, where inconsistencies were checked and resolved, as well as missing records were filled in according to the tables described in the methodologies.

The ecosystem type comparison Excel sheet (Fig. 2) contains horizontally all levels of the national level classification (columns A to C) and column D, which contains the full names from the methodologies. In order to achieve this objective it is necessary to analyze national typology and the database of the ecosystems in Bulgaria. The analyses are performed in two case studies representative for the terrestrial ecosystems in the country: the Ogosta river basin and the city of Sofia. The database contains data for nine ecosystem types found in the country. The case studies cover eight of them excluding the marine ecosystems. The main goal of this paper is to present the first results of the ecosystems database analyzing the possibilities to develop a common database for the needs of integrated water management.
Ecosystem and their Services in Bulgaria of all 8 ecosystem types present in the Ogosta river basin (Apostolova et al., 2017a, 2017b; Kostov et al., 2017; Sopotlieva et al., 2017; Uzunov et al., 2017; Velev et al., 2017; Yordanov et al., 2017; Zhiyanski et al., 2017).

Completion of column E includes the ecosystem types and subtypes names from the official database for the nine mapping projects under the MAES methodological framework. The database was provided for the purpose of the INES project by the Ministry of Environment and Water through the Executive Environment Agency - ExEA (requested by letter No. 33-00-150/20.06.2022 of the Ministry of Environment and Water). In order to perform the verification, the data were loaded into ArcCatalog, version 10.6.1. One by one, the layers were compared and the names of the types and subtypes from the database were transcribed into the Excel spreadsheet with the guiding criterion being their meaningful match. In the last column F, the EUNIS ecosystem type names were added. The official website (https://eunis.eea.europa.eu/habitats-code-browser.jsp?expand=182#level_182) of the European Environment Agency was used for this purpose.

A primary check in Excel and GIS revealed several mismatches in general related to certain ecosystem types, the most significant of which are related to woodland and forest, grassland, and freshwater ecosystems. Problematic areas requiring correction were removed and missing information was completed.

The detection and removal of the errors and inconsistency of spatial data in vectors are the main concerns of the Geographic Information Systems (GIS) since the GIS templates in vector form are obtained from the raster maps and plans (Maras et al., 2010). The topology describes the spatial relationships between objects using sets of rules for how vector entities (points, lines, polygons) share geometry and space and its importance derives from the objectives pursued by defining and enforcing conditions on vector geometric entities (Herbert et al., 2015). In the current work we used two main topology rules to analyze the quality of the ecosystem database: must not overlap and must not have gaps.

3. (Mis)matches in the typology and attributes’ data

3.1. Duplicate numerical designations between grassland and woodland and forest ecosystems

Duplicate numerical designations error between grassland and woodland and forest ecosystems occurs when proceeding with merging all available ecosystem layers from the eight methodologies into one composite layer. When classifying them according to the digital records from the attribute field “EcosystemF”, it appears that in the area of the Ogosta river basin, the woodland and forest ecosystems are missing, which is incorrect. The following detailed data examination revealed that the numeric designations in both layers begin with code 301+. This code, according to MAES, belongs to grassland ecosystems, while woodland and forest ecosystems are defined to start with 401+. We, therefore, undertook a renaming of all woodland and forest ecosystem codes beginning with "3" to "4". In doing so we have ensured:

1. Consistency with Level 2 classification numbering;
2. Ability to subsequently merge the two layers without the risk of losing records and being unable to classify (distinguish) them from each other.

The changes made to the attribute tables of the GIS layers were also reflected in column “C” of the table CLC_MAES_class_work.xlsx, containing the codes of the individual ecosystem subtypes.

3.2 Mismatches between the forest ecosystems classified in the Bulgarian methodology and those in the MAES framework

When the woodland and forest ecosystems were checked, it was found that the alphanumeric indices "G1" (Broadleaved deciduous woodland) and "G2" (Broadleaved deciduous woodland - coppice) in the methodology are at the same common level, but in the Bulgarian version, they are distinguished. In order to be able to distinguish the GIS data for the later classification, Roman numerals were added after the existing records in both indices, so that Broadleaved deciduous woodland - coppice becomes "G1 -G2".
deciduous woodland – coppice became index “G1.I” and Broadleaved deciduous woodland became index “G1.II”.

Additional rows were added to the Excel spreadsheet to write the descriptions for each code. It was found that codes 401-407 corresponded to deciduous broadleaved coppice woodlands, 408 to 414 were deciduous broadleaved woodlands, and 415 to 417 were coniferous.

In the official database for the nine ecosystem mapping projects, opened in ArcCatalog 10.6.1, the code names are as follows: 1) Code 423 - 'Highly artificial coniferous plantations'; 2) Code 424 - 'Mixed deciduous and coniferous woodland'; 3) Code 427 - 'Mixed forestry plantations'. After comparison in the table, code names 423 and 427 match the description of the methodology and EUNIS classification. The code name 424 only matches the methodology description, whereas there is no match in the EUNIS classification.

3.3. Mismatches between codes and names of freshwater ecosystems classified in the Bulgarian methodology and those in the MAES framework

The following problems were identified in the classification of freshwater ecosystems: 1) incorrect classification - the "Level 1" column of the Bulgarian methodology contains two categories that do not exist in the original MAES methodological framework; 2) at "Level 2" in MAES there is only one category "Rivers and lakes", which in the Bulgarian methodology is given as a total of four categories (C1, C2, X01, X03). These four categories should be assigned to "Level 3" and the current "Level 3" entries should become "Level 4". Moreover, it was found that the authors had used the definitions of the Water Framework Directive 2000/60/EEC (WFD), which describes four categories of surface water bodies: rivers, lakes, transitional (brackish) and coastal marine waters, in addition to artificial and heavily modified (modified/man-made) water bodies.

The current codes in the "Level 2" and "Level 3" columns have been checked for consistency with the EUNIS codes. It was found that there are no entries for codes 811 and 812 in the worktable in column "C" of the main sheet called "ec-bg" and "freshwater", whereas there are entries for these codes in the spatial data. Code 855 is also missing. Following this, a thorough check of all codes was made. Leading the check was the separate ecosystem-type methodology.

3.4. Mismatches between Bulgarian MAES and EUNIS

The results obtained from the comparison between the MAES methodological framework and the EUNIS classification are reproduced in a new column (M) in the table (Fig. 2). The comparison is represented by a 5-point scale showing the level of correlation between both data. When filling-in column M, several omissions and/or discrepancies were found, among data: 1) within agricultural ecosystems, in the sub-type "Livestock farms for large agricultural ecosystems, in the sub-type "Livestock farms for large

degree of the scale. Third degree express the correlation at the second EUNIS level, but not at the third level. It includes three of the subtypes considered. The fourth degree of the scale also features a slightly larger number. Here, subtypes are distinguished that have no correlation between the compared levels of the MAES and EUNIS mapping projects, but there is correlation at other levels. The fifth level assesses subtypes where there is no correlation, with only one ecosystem in this category. As shown on the figure 3, about 80% of the entries have a complete correlation to EUNIS, and the remaining 20% have partial correlation. The most ecosystems cover grades 1 and 4, and the lowest grade 5.

4. (Mis)matches in spatial data

The topology analyses of the spatial data are performed using ArcGIS software. The analyzes are made for both case studies, the Ogosta river basin and the urban part of the city of Sofia. The results are generated in two formats: 1) statistics in tabular format, and 2) spatial vector data (polygon and linear objects representing the gaps and overlaps of the data).

4.1. Gaps and overlaps in the spatial data for Ogosta river basin

There is about 80% data coverage of the basin. The rest of the area falls within the NATURA 2000 network which was outside of the national ecosystem mapping. The topology data analysis carried out for the Ogosta river basin area identified more than 57 000 gaps and overlaps (Fig. 4). The cropland, grassland, heathland and shrub ecosystem types have the most frequent errors in the data (Fig. 5).

More detailed topology analyses are carried out for the upper part of the Ogosta river basin, which is one of the case studies in the INES project. In contrast to the whole basin, the upper part of the basin has only about 10% data coverage. The reason is that there are larger NATURA 2000 sites in this part. The results of the analysis show a total of 1034 gaps and overlaps (Fig. 6).

Most of the available ecosystem mapping data in this area have some kind of errors. The highest number of errors are observed in the grassland, woodland and forest, and cropland ecosystems, which have 105 up to 130 gaps. The highest number of overlaps is observed in grassland, heathland and shrub, and woodland and forest ecosystems (Fig. 7). Although there are fewer errors all other ecosystem types show a particular number of gaps and overlaps. One of the reasons for the overlaps is due to a misinterpretation of code 501 from heathland and shrub, which has been assigned to both woodland and forest and heathland and shrub ecosystems. The predominant number of overlaps are between cropland ecosystems (codes 201 or 202) and the deciduous coppice forests (code 301).
Analysis of the national ecosystem database of Bulgaria: (Mis)matches with the MAES framework

Figure 4. Overall view and statistics of identified gaps and overlaps in the Ogosta river basin data.

croplands_ogosta_basin,Must Not Have Gaps,4004.0
wetland_ogosta_basin,Must Not Have Gaps,7.0
urban_ogosta_basin,Must Not Have Gaps,636.0
sparcely_vegetated_ogosta_basin,Must Not Have Gaps,27.0
rivers_lakes_ogosta_basin,Must Not Have Gaps,214.0
heathland_ogosta_basin,Must Not Have Gaps,3318.0
grassland_ogosta_basin,Must Not Have Gaps,3761.0
woods_ogosta_basin,Must Not Have Gaps,3391.0
croplands_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,6.0
croplands_ogosta_basin,Must Not Overlap With,urban_ogosta_basin,486.0
croplands_ogosta_basin,Must Not Overlap With,sparcely_vegetated_ogosta_basin,0.0
croplands_ogosta_basin,Must Not Overlap With,rivers_lakes_ogosta_basin,161.0
croplands_ogosta_basin,Must Not Overlap With,heathland_ogosta_basin,2605.0
croplands_ogosta_basin,Must Not Overlap With,grassland_ogosta_basin,3335.0
croplands_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,7898.0
ggrassland_ogosta_basin,Must Not Overlap With,heathland_ogosta_basin,1786.0
ggrassland_ogosta_basin,Must Not Overlap With,rivers_lakes_ogosta_basin,8.0
ggrassland_ogosta_basin,Must Not Overlap With,sparcely_vegetated_ogosta_basin,13.0
ggrassland_ogosta_basin,Must Not Overlap With,urban_ogosta_basin,1.0
ggrassland_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,5.0
ggrassland_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,8846.0
heathland_ogosta_basin,Must Not Overlap With,rivers_lakes_ogosta_basin,190.0
heathland_ogosta_basin,Must Not Overlap With,sparcely_vegetated_ogosta_basin,2.0
heathland_ogosta_basin,Must Not Overlap With,urban_ogosta_basin,207.0
heathland_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,0.0
heathland_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,6212.0
rivers_lakes_ogosta_basin,Must Not Overlap With,sparcely_vegetated_ogosta_basin,0.0
rivers_lakes_ogosta_basin,Must Not Overlap With,urban_ogosta_basin,29.0
rivers_lakes_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,3.0
rivers_lakes_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,215.0
sparcely_vegetated_ogosta_basin,Must Not Overlap With,urban_ogosta_basin,8.0
sparcely_vegetated_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,0.0
sparcely_vegetated_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,84.0
url_ogosta_basin,Must Not Overlap With,woods_ogosta_basin,1174.0
urban_ogosta_basin,Must Not Overlap With,wetland_ogosta_basin,2.0

Figure 5. Statistics for gaps (A) and overlaps (B) between ecosystem types in the Ogosta river basin (C).
Figure 6. Overall view of the identified gaps and overlaps for the case study of the upper Ogosta river basin.

Figure 7. Statistics for gaps (A) and overlaps (B) between ecosystem types in the upper Ogosta river basin (C).
4.2. Gaps and overlaps in the spatial data for the city of Sofia

The results from the topology analyses of the city of Sofia show the predominant number of errors located in the periphery of the case study (Fig. 8). The main reason for this spatial distribution is the fact that the same areas appear as boundaries between urban ecosystems and other subtypes, e.g. grassland or cropland (Fig. 9).

The total number of overlaps is 815. The most frequent errors are between urban, cropland, or heathland and shrub overlapping any of the other types (Table 1). Urban ecosystems have the largest overall proportion of overlaps, with a total of 700 errors, of which only 563 errors are between urban and woodland and forest ecosystem types.

Figure 8. Overall view of gaps and overlaps in the data for the city of Sofia.

Figure 9. Close-up view of the gaps and overlaps between selected ecosystem types in the data for the city of Sofia.
Table 1. Ecosystem type overlaps in the data for the city of Sofia.

<table>
<thead>
<tr>
<th>Ecosystem type</th>
<th>Ecosystem type</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cropland</td>
<td>Grassland</td>
<td>24</td>
</tr>
<tr>
<td>Cropland</td>
<td>Heathland and shrub</td>
<td>1</td>
</tr>
<tr>
<td>Cropland</td>
<td>Woodland and forest</td>
<td>8</td>
</tr>
<tr>
<td>Grassland</td>
<td>Woodland and forest</td>
<td>36</td>
</tr>
<tr>
<td>Heathland and shrub</td>
<td>Grassland</td>
<td>18</td>
</tr>
<tr>
<td>Heathland and shrub</td>
<td>Freshwater</td>
<td>4</td>
</tr>
<tr>
<td>Heathland and shrub</td>
<td>Woodland and forest</td>
<td>24</td>
</tr>
<tr>
<td>Urban</td>
<td>Cropland</td>
<td>55</td>
</tr>
<tr>
<td>Urban</td>
<td>Heathland and shrub</td>
<td>56</td>
</tr>
<tr>
<td>Urban</td>
<td>Freshwater</td>
<td>26</td>
</tr>
<tr>
<td>Urban</td>
<td>Woodland and forest</td>
<td>563</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td></td>
<td>815</td>
</tr>
</tbody>
</table>

5. Conclusions and future steps

The MAES ecosystem typology is developed at third and fourth level for the territory of Bulgaria following EUNIS habitats categories and other classes specific for the country. It enables detailed large-scale mapping of all nine main ecosystem types. However, our study identified several discrepancies between which may cause confusion in cases of integrated assessment of all ecosystem types. The first problem is that for some types it is developed to fourth level while for the other it is to the third level. The second problem is related to the hierarchical levels in some of the ecosystem types especially in the freshwater ecosystems, where different categories are mixed within a single hierarchical level. The third problem comes from the duplicated numerical designations between grassland and forest ecosystems. The reasons for such problems can be both objective and subjective. Objective reasons are related to the different nature of the ecosystem types which determines differences in the ecosystem categories at different levels. The objective reasons are caused by the different interpretations of the teams which prepared the typology of the different ecosystem types. Therefore, a revision of the typology is necessary in order to solve the above-mentioned problems and to prepare a correct uniform typology to be used for the needs of integrated assessment. This will be in line with the recommendations towards better consistency of the mapping efforts (Maes et al., 2020).

The database of the ecosystems is designed for ESRI Geodatabase format and includes a scheme consisting of seven interrelated tables. This ensures a uniform structure of the nine ecosystem databases and easy relation between the different ecosystem types. However, our study again identified several problems which impede the use of the database for integrated assessment. Firstly, the coding system does not fully correspond to the hierarchical structure of the typology. It is well developed for three level hierarchy, but could hardly incorporate the fourth level of the typology. Secondly, there are some discrepancies between the numerical codes from the database and the indexes used in the methodology. Thirdly, there are some discrepancies between codes used in the database and the EUNIS codes which may cause confusion in the interpretation of the results from ecosystem mapping. The conclusion again is towards a revision of the database coding system preferably by developing a new design of the codes in correspondence with the revised typology. Such an effort will ensure better interrelation with the data about ecosystem condition which is recommended in order to fill the gap about spatial mapping of condition (Erkhard et al., 2016; Maes et al., 2018).

The topology analyses of the merged data from the eight ecosystem GIS layers show extremely large numbers of gaps and overlaps. The main reason is in the use of different sources for the mapping of different ecosystem types. In most cases the data sources are in the form of GIS vector datasets developed from other initial sources such as satellite images, orthophoto images, cadastral data etc. The manipulation for these procedures causes additional spatial errors which predefines most of the identified gaps and overlaps. The main conclusion is that it is practically impossible to generate topologically correct integrated GIS layers from the eight ecosystem type layers. Therefore, it is necessary to develop a new approach for mapping of all ecosystem types into a uniform database. The first task for such an approach is to define initial spatial data-source with full coverage of the country. This source can be used to develop an initial spatial dataset of the ecosystem types at level 2. The incorporation of the other levels of the typology should be arranged as further steps based on the available data sources for the different ecosystem types.

Funding program

The study was carried out within task 4.1 of the WP4 of the INES project (INtegrated assessment and mapping of water-related Ecosystem Services supporting nature-based decisions in river basin management), funded by the National Science Fund of the Bulgarian Ministry of Education and Science, under contract No. KP-06-N-54/4.

Acknowledgements

We would like to express our gratitude to the Ministry of Environment and Water (MOEW) of the Republic of Bulgaria and the Executive Environment Agency (ExEA) for providing us with the official GIS database for the nine mapping projects under the MAES methodological framework.
References


Author contributions (CRediT roles)
GP - Data curation, Formal analysis, Investigation, Validation, Visualization, Writing - original draft; HP - Conceptualization, Methodology, Resources, Supervision, Validation, Writing - original draft, Writing - review and editing; VS - Data curation, Formal analysis, Investigation, Validation, Visualization, Writing - original draft, Writing - review and editing.

Conflict of interest
The authors have declared that no competing interests exist.

ORCID
https://orcid.org/0000-0001-7412-0097 - G. Petkova
https://orcid.org/0000-0003-2453-8975 - H. Prodanova
https://orcid.org/0000-0001-7354-1711 - V. Stoycheva