RESEARCH ARTICLE

Spatial Database Designing for Environmental Monitoring and Decision Making in Mitrovica Region, The Republic of Kosovo

Bashkim Idrizi1,2, Edon Maliqi3,*, Lyubka Pashova4
1Department of Geodesy, University of Prishtina “Hasan Prishtina”, Prishtina, Kosovo
2Geo-SEE Institute, Skopje, North Macedonia
3Faculty of Architecture and Spatial Planning, UBT – Higher Education Institution, Pristina, Kosovo
4Department of Geodesy, National Institute of Geophysics, Geodesy and Geography - Bulgarian Academy of Sciences, Sofia, Bulgaria

Received 18 April 2021/Revised 21 July 2021/Accepted 2 August 2021/Published 17 August 2021

Abstract
The integration of spatial data analysis methods and thematic map models is an approach to reduce the negative impact of anthropogenic pressure on the environment due to mining and waste generation. The large amounts of industrial waste from mining in the Mitrovica region in northern Kosovo lead to serious environmental problems with organic and inorganic water and soil pollution. This study aims to design and establish a geospatial database for long-term environmental monitoring, provide analytical tools, and support appropriate management decisions by local authorities and agencies. The database contains topographical elements and ecological parameters collected from different national and open access international sources. All collected data have been analyzed, standardized and harmonized within the open-source QGIS ver.3 software. The results showed that in developed datasets were organized in different GIS layers and compiled several thematic maps. The designed database is unique by its architecture, providing an opportunity for periodical monitoring of the environment near the mining areas.

Keywords: Environmental monitoring; Spatial database; Open source software; QGIS; Kosovo

1. Introduction
Geospatial information has become an indispensable element in several studies related to monitoring, analysis, modelling, visualization, and mapping environmental parameters. Thirty years ago, terms such as ‘environmental monitoring’, ‘risk assessment’, ‘contamination management’ and ‘risk management’ were not commonly used (Maliqi, 2020). In recent decades, significant progress has been made in management approaches and tools used to promote sustainable environmental management and protection in the context of global change, natural and human-made hazards.

*Corresponding author.
Email address: edon.maliqi@ubt-uni.net (Edon Maliqi)
Most environmental data would be almost valueless without their spatial reference component (ESRI, 2010; Sanka, 2015). According to Ilyushchenko & Kamberov (2004), public officials and environmental managers need suitable tools for the visualization, manipulation, and analysis of ecological data. Effective monitoring and improved understanding of the state of the environment require the collection of essential information (ESRI, 2010; Shrestha et al., 2016). Geographic Information System (GIS) is a technology that provides an opportunity for ecological scanning and analysis of the environment and creating models for assessment and forecasting for its future state. According to several studies, the GIS database has many advantages for environmental management. Modern GIS supports automated processing, increases productivity with streamlined workflows in compliance with regulatory requirements, and improves management decision-making. It deals with collecting, storage, retrieval, manipulation, updating, analysis, display, and disseminating spatially referenced data and information. GIS systems are vital tools for managing an ecological database at all scales, ranging from local to global and mapping the environment (Lein, 2006; ESRI, 2010; Thakur et al., 2011; Hendricks, 2014; Maliqi, 2020). The list of needed source data depends on the environmental characteristics, planned calculations, and existing data. Collecting spatial data, preparing the data for GIS use, and documenting those are usually the most expensive and time-consuming aspects of any GIS project (Lo et al., 2001). More specifically, many modern satellite missions are solely focused on collecting datasets using instrumentation on board. In contrast, the relevant satellite platforms can provide accurate remotely sensed data and products covering more expansive areas that it is impossible to continuously monitor using other instrumentation or technologies (Kolios et al., 2017).

The growing concern about sustainable mining practices, environmental monitoring in and around industrial areas is a critical process that can be successfully supported through advanced geospatial monitoring practices. Nowadays, geospatial technologies have been applied for the preparation of environmental impact assessments for different mining areas. Roumenina et al. (2007) represent a geodatabase organization model for large-scale mapping of geocological aspects of land use conflicts at the level of settlement land mapping. The developed geodatabase eliminates sources of conflicts in land use or mitigates their impact in areas where metallurgical, mining and chemical production are concentrated. Similarly, in other countries with developed mining and subsequent environmental problems, such as Nigeria (Ochelebe et al., 2020), Morocco (Khalil et al., 2014), China (Qi et al., 2017), Indonesia (Maryati et al., 2012), South Africa (Orimoloye & Ololade, 2020) Sierra Leone
(Kamara, 2020), the use of GIS with combination with modern remote sensing technologies and tools is becoming routine.

Mining companies are mandatory to perform ecological management efforts to minimize the negative impact on the environment. There is no previous study for the Republic of Kosovo, which attempts to analyze some of the leading environmental concerns using geospatial technologies. Until now, independent researchers did not answer the spatial and temporal environmental questions. As a result, there is a significant gap in the spatial and temporal distributions of the ecological phenomena. Furthermore, the environmental data for a specific purpose were not found for the Mitrovica region, located in northern Kosovo. There are many problems in finding and obtaining environmental data to solve any environmental issue.

This study focuses on establishing a geospatial database for monitoring the mining environment, including groundwater, surface water, vegetation and erosion in the area of Mitrovica, the Republic of Kosovo. The region has a long history of lead and zinc metallurgy productions, where, unfortunately, these mining operations have resulted in adverse environmental impacts (Dekonta, 2009). The natural environment in and around the industrial area is affected widely by heavy metals such as Arsenic (As), Chromium (Cr), Copper (Cu), Nickel (Ni), Lead (Pb), Zinc (Zn) and Cadmium (Cd). After a brief description of the research area, the rest of the article is organized as follows. In Section 2, the primary sources of geographic and specialized data were collected from governmental authorities and departments, private sector and companies, technical reports, literature review and conducted campaign measurements in three months continuously. Three field works were carried out during three monitoring months May, June and July 2018. These input data are used for establishing a database QGIS environment. The conceptual, logical and physical design of the spatial database for environmental analyses in the study area is also outlined. Section 3 presents the main results and their discussion. In the last Section 4, some general conclusions for maintaining, updating and improving the created information database are outlined.

2. Methods
2.1 Study Area

The study area covers part of the Mitrovica region in the northern part of the Republic of Kosovo (Maliqi, 2020), with about 41 km² (Maliqi & Singh, 2019). It is defined based on the environmental influence of two tailings nearby the town of Mitrovica and Kelmend village population centers, the Sitnicariver and the Trepca mine complex (Figure 1). The study region is located within 42°57'N and 42°52'N and 20°55'E and s20°54'E and has an
altitude between 500 and 950 m. The area is situated in the Vardar zone of the dinaride alpine belt (Maliqi, 2020), consisting of palaeozoic basement rocks, jurassic-cretaceous sediments and rocks of ophiolitic affinities foliated during the early tertiary (Heinrich & Neubauer, 2002). The weather in this region is continental, with warm summers and cold, snowy winters.

Figure 1. Study area in Mitrovica region, North Kosovo

The town of Mitrovica has a long history of lead and zinc metallurgy productions (Dekonta, 2009). Cities like Mitrovica with an industrial pollution legacy face many complex challenges. This region has provided economic opportunities and many jobs, but this is at the expense of high costs for environmental protection and human health. The primary sources for soil degradation and environmental pollution are the Trepça Mines, which is a large industrial complex in Kosovo and the largest lead, zinc and silver mining mine in Europe. Pollution in this region comes from various sources but is mainly due to the industrial legacies, the influence of winds and water movement.
2.2 Source Data and Methodological Approach

Having the source data's technological and institutional challenges can affect serious environmental monitoring commitment in the studied area. Therefore, for this study, three sources of environmental data are used to perform environmental analyses: (1) Authoritative information developed by experts/governmental institutions without accessibility possibilities by the public; (2) Information developed by specialists and scientists with options for open use by the public; (3) Information is developed by involving the public as a participatory contribution, e.g. Volunteered geographic information (Sadeghi-Niaraki et al., 2020).

A systematic approach was followed to obtain the source data, including (1) identifying the data needed in terms of environmental monitoring, (2) identifying the open-access geoportal and organizations of data collection, (3) searching the open-access geospatial databases and visiting the local organizations/institutions, (4) downloading and gathering available geospatial data (Maliqi, 2020). Priority was given to environmental data available in the Republic of Kosovo. Geosensors are defined as devices for receiving and measuring environmental parameters that can be georeferenced (Retscher & Kabashi, 2014).

The lack of open geospatial data in the ecological context, the difficulties to have ecological data in digital formats, lack of systematized data in local/central institutions or agencies, lack of vision for such data from environmental monitoring institutions made to have difficulties in ensuring a large amount of needed ecological datasets. Therefore, it was impossible to collect datasets required to be migrated in the spatial database. Primary data for the performed research have been obtained by direct field measurements of some ecological parameters. In contrast, additional data are extracted from different sources such as professional literature, open databases, and responsible institutions. As a set of primary data, surface and groundwater data were collected in three monitoring months in six monitoring stations by using GPS equipment for defining geolocation and bottles for water sample collection. A list of spatial coordinates for six locations and eighteen water samples for surface and groundwater data has been composed. Water samples have been used for analyzing the heavy metal pollution and metal index. Surface water is one of the most vital Earth resources changing in time and space due to land use/land cover changes, climate change, and other environmental factors (Kaplan, 2020).

For monitoring of vegetation and erosion, the Landsat satellite images provided by the U.S. Geological Survey (USGS, 2020), the Digital Elevation Model with 20 m spatial resolution provided by ALOS (Japan Aerospace Exploration Agency, 2020), Land Cover map provided by Copernicus (2020), the study area Geological map, the Corine Land Cover database, temperature data from World Clime web database (World Climate, 2021), soil
texture (Food and Agriculture Organization, 2021), the Mitrovica municipality boundary in shapefile from the Kosovo Cadastral Agency (2020), the mining tailings boundaries from the Ministry of Environmental and Spatial Planning (2011) in Kosovo, and the transportation and hydrology data (main roads, railway and rivers) from Open Street Map (2021).

2.3 Spatial Data Mathematical Elements

Mathematical rules of the spatial database in QGIS are defined according to the Republic of Kosovo's national legislation related to geospatial data (Kosovo Cadastral Agency, 2020), technical requirements for environmental monitoring of area with about 41km$^2$ and diversity of social and natural phenomena in the study area. They are as follows: (1) Coordinate Reference System: Kosovoref01 with EPSG code 9141 (Idrizi, 2020); (2) The scale of vector data is 1:10,000; (3) The spatial resolution of raster data is 10m.

2.4 Spatial Data Structure

Spatial data structures manipulate spatial data, that is, data with geometric coordinates (BOWDOIN, 2008). GIS-based environmental applications became prominent with shifting from static paper maps to dynamic digital maps and geographic databases. GIS became one of the most applied computer-based tools for storing and manipulating environmental data (Sadeghi-Niaraki et al., 2020). In the present study, the spatial database was designed and developed using QGIS software ver 3.0 and its tools and visual paradigm such as Unified Modeling Language (QGIS, 2021).

The datasets were generated to organize the monitored environmental data datasets, the type of representation and their relationship. The basic schema of spatial data content architecture for the Mitrovica region is given in Figure 2. This scheme includes Basemap, Hydrology, Transportation, Meaning Features, Groundwater Quality, Surface Water Quality, Vegetation Monitoring and Erosion Monitoring layers. Basemap and Erosion monitoring layers are stand-alone layers directly linked to a database, while the other six layers are grouped in two groups per three layers. The first group contains hydrology, transportation and mining features (left side of Figure 2). In contrast, the second group has Groundwater Quality, Surface Water Quality, and Vegetation Monitoring layers (right side of Figure 2). Based on requirements for performing environmental spatial analyses and available source spatial datasets, an appropriate data architecture was developed.
Figure 2. Scheme of the environmental database content

The spatial database has been designed based on the Arctur & Zeiler (2004) approach, with three designing phases (I - conceptual, II - logical, and III - physical), which contains ten steps (I - 1. Identify the information products produced with the GIS system; 2. Identify the key thematic layers; 3. Specify the scale range and spatial representation for each thematic layer; 4. Group representations into datasets; II - 1. Define the tabular structure and behaviour of descriptive attributes, 2. Define the spatial properties of the datasets, and 3. Propose a spatial database design; III - 1. Implement the design; 2. Design workflow for building and maintaining each layer, and 3. Design documentation).

Layers are followed by metadata based on ISO 19115:2003 standard (ISO, 2021), containing information on identification, categories, keywords, access, extent, contact, link, history and validation of the spatial data layers.

2.5 Data Harmonization

Input data obtained from several existing databases and field measurements have different structures and formats (Idrizi et al., 2018). To have a homogenous spatial database for performing research, the input data from other datasets should fit geometry and semantics. The entire source data process was organized according to the five steps of the data harmonization framework, defined by Janecka et al. (2013), as follow: (1) Theory of spatial
data harmonization; (2) Source data understanding; (3) Target data understanding; (4) Definition of necessary harmonization steps; (5) Practical realization.

All collected source data have been converted and transformed based on the predefined spatial data structure and spatial data mathematical elements (Lubishtani & Idrizi 2016) described in the text’s upper points. Harmonized data cannot integrate spatial data in one place, with the same data structure, in the same format, in the same coordinate system, in the same scale and with similar accuracy. The data harmonization process between 27 feature classes given in figure 3 ensured full interoperability between source data from different resources. It enabled performing complex spatial environmental analyses by common usage and data combination between 27 feature classes of 8 datasets.

Figure 3. Structure of spatial datasets
3. Results and Discussion

3.1 Environmental Database for Mitrovica Region

To address the challenges of managing multi-parameter and multi-time spatial data and information for environmental monitoring in the studied region, open-source QGIS is an effective tool. Creating an integrated unified database allows avoiding redundant data, reducing data volume, homogeneity data, easy finding and tracking of data, integrated spatial and attribute data, their use for GIS spatial analysis for decision making and development strategies. A spatial database has been developed focusing on the environmental geospatial data in full accordance with standard procedures, including conceptual, logical and physical design phases (Arctur & Zeiler, 2004; Sepehr, 2013) described above in the text. Developing a spatial database involves defining the vital thematic layers regarding the usage, content, and representation of each thematic layer or defining how geographic features are represented (as points, lines, polygons or attribute and tabular data) (Sepehr, 2013). Arctur & Zeiler (2004) pointed out that spatial database design specifies how data is organized into feature classes and attributes and specifies the GIS behaviours by establishing the spatial relationships in datasets using rules. Developed environmental spatial database for the Mitrovica region of the Republic of Kosovo contains both the vector and raster layers, according to the spatial data structure. The extract of feature classes are given in Table 1 and Table 2. Data feature classes based on the geometry of vector data classified in point, line and polygon categories are given in Table 1, while raster feature classes grouped in five datasets are shown in Table 2.

Table 1. Vector data feature classes in ‘Environmental Spatial Database’

<table>
<thead>
<tr>
<th>Point data</th>
<th>Line data</th>
<th>Polygon data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trepca Mine</td>
<td>Sitnica River</td>
<td>Republic of Kosovo</td>
</tr>
<tr>
<td>Flotation</td>
<td>Roads</td>
<td>South Mitrovica</td>
</tr>
<tr>
<td>Groundwater Monitoring Stations</td>
<td>Railway</td>
<td>The study area</td>
</tr>
<tr>
<td>Surface Water Monitoring Stations</td>
<td>-</td>
<td>Tailings</td>
</tr>
<tr>
<td>Trepca Mine</td>
<td>-</td>
<td>Sitnica river (50 m)</td>
</tr>
</tbody>
</table>

Table 2. Raster data feature classes in ‘Environmental Spatial Database’

<table>
<thead>
<tr>
<th>Base map</th>
<th>Groundwater Quality</th>
<th>Surface water Quality</th>
<th>Vegetation Monitoring</th>
<th>Erosion Monitoring</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Heavy Metal</td>
<td>Heavy Metal</td>
<td></td>
<td>Soil Loss (2018)</td>
</tr>
<tr>
<td></td>
<td>Pollution Index**</td>
<td>Pollution Index***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Metal Index**</td>
<td>Metal Index***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * Spline, ** Inverse distance weighting (IDW), *** Kriging
The ‘Environmental Spatial Database’ developed in this research consists of eight feature datasets, structured according to the conceptual and logical design phases, as an extendable database that can be upgraded with new thematic layers and data processing algorithms. Using QGIS as free and open-source software, the developing procedure has been performed, part of the OSGeo family that contain all OGC standards and possibilities for integrating spatial data within the INSPIRE directive specifications for establishing the spatial data infrastructure. Vector and raster datasets are organized in one project (Maliqi, 2020), containing project metadata information for performed research and specific settings for supporting spatial data. The datasets are grouped in separate thematic groups and subgroups in a hierarchical model. Four theme models for automatic visualization change from one to other thematic parts (groundwater, surface water, vegetation and erosion). For each vector and raster layer, besides metadata information, QGIS symbol styles in *.qml and *.sld were created, later exported as cartographic keys in *.xml format. Vector layers are fully interoperable between them, as well strict rules for geometry, topology, joins and relationships were used for creating preconditions for performing spatial analyses. Raster data have been aligned and then analyzed to avoid the no data and sieve raster cells. The final *.qgz project has all technical opportunities for lunching as a web service for public use by stakeholders. Such solid rules for establishing the environmental spatial database were able to define and perform using QGIS tools. The data (thematic layers) were loaded in terms of the entity types and their completeness in the loading data procedure, shown in Figure 4. The hierarchical organization in the established environmental spatial database is given as an extract from QGIS software, giving an overview of spatial data hierarchy and links between datasets and feature classes.

This spatial database offers a wide range of environmental monitoring and decision-making applications and interdisciplinary analysis in the Mitrovica region. Available GIS database allows to:

- Show the spatial-temporal distribution of environmental contamination and degradation; the numerous serial thematic maps were compiled through this spatial database.

- Having such spatial data and attribute data as one system helps show the spatial distribution and environmental pollution values and identify the most degraded areas. At the same time to identify areas potentially endangered by the natural and anthropogenic factors.
- Easily to be compared the spatial and non-spatial data with national and international standards and directives.
- Use the current spatial database as base series for further environmental analysis and subsequent scientific studies to perform a trends analysis for this region.

Figure 4. Feature datasets and classes in 'Environmental Spatial Database'

The developed spatial database has importance because all concerned parties can be involved; it becomes a new system for authorities in the Mitrovica region and the Republic of
Kosovo. An important issue is knowing why the spatial database is being designed and how it can be used. If not, there is a risk that this environmental spatial database will not be utilized and developing in the future.

There are a few attempts to develop spatial databases or geo databases by different researchers and authors in the past. Hence, a comparison approach (previous studies) speaks that in the past, various researchers (Kranjcic et al., 2020; Roumenina et al., 2007) have done their analysis focusing on the environment but separately for each parameter such as waters, noise, air, erosion, vegetation. While, the present study encompasses a wide range of parameters, including groundwater, surface water, soil erosion and vegetation, even on different time series. As a common thing compared to previous studies, the source data were gathered mainly from the same open-access sources and used the same mathematical models to ensure the valid input data (water monitoring indices). The comparative analysis has shown that the developed spatial database is in full accordance with the theory of geoinformation science and systems with full support by QGIS software. It is confirmed by other authors as well, for example, a previous study by Plassin et al. (2019). Therefore, the given results at the end of the present study remain within the theory limits and are comparable to other authors. However, they reflect the environmental specifics of our study area.

On the other hand, the newly geospatial database in this study remains unique because we have a wide range of environmental aspects/parameters united into one database. The difference from previous studies is the number of environmental elements and parameters included in this spatial database. In this way, the present study reflects the current ecological conditions of the study area based on the last two decades monitoring. However, at the same time, it can be used for forecasting analysis for the future. Environmental impact assessment and measures to decrease negative impacts on the environment (Krasniqi et al., 2018), by using the established environmental spatial database for the Mitrovica region in this research, will be a bridge between the current and future investigations (Idrizi & Zhaku, 2013) as well as a measure for environmental protection of studied area in next period. The developed database enables defining algorithms for performing spatial analysis for the study area and provided opportunities for periodic monitoring of groundwater quality, surface water quality, vegetation and erosion, with integrated multi-stage iterative processing of spatial and non-spatial data.

For the Mitrovitsa mining region, there was no available environmental spatial database. Therefore, this is the first attempt to have a database in the ecological context, including groundwater, surface water, vegetation and erosion in different time series for the
last two decades. In addition, the spatial database included other environmental aspects to support the ecological monitoring in the studied region. The spatial database output can be visualized as time series thematic maps, graphs, reports, or three-dimensional views.

The spatial database elaborated in this study has many advantages for environmental monitoring. It includes the integration of different essential thematic layers into a single database. The redundancy data is avoided, the volume of information is reduced, and the data are uniform. It can be served for environmental planning because all the environmental data can be overlaid with other data to view spatial relationships in the ecological context. To undertake “what if” scenarios, analyzing different environmental conditions can be used in this spatial database. It is not only for present environmental procedures but also for future ecological analysis and studies. Furthermore, the ‘Environmental Spatial Database’ will answer the many questions for environmental monitoring and assessment in the Mitrovica region and its change ratio in the future. However, several conditions must be fulfilled to implement this spatial database by local or central institutions. First, the spatial database implementation as assistance for environmental monitoring should be articulated by the local or governmental institutions. Second, it is essential to consider the datasets' potential with key thematic layers into the spatial database for decision and policy-making for this region. An important issue is administrative officers’ ability to use and manipulate geospatial data to generate intelligent solutions.

4. Conclusion

In the present study, an overview, theoretical, and implemented spatial database design approach was given for particular ecological issues in the Mitrovica region, Kosovo. Based on ISO and OGC standards, the created ‘Environmental Spatial Database’ integrates the different spatial and attribute/tabular environmental data types allowing their easy manipulation and performing complex geospatial analyses. During the conceptual, logical and physical design phases, the geodatabase structure was significantly improved. The environmental entities can be visualized as thematic tabular and cartographic results in QGIS. Its availability will benefit stakeholders involved in environmental monitoring and management. In addition, after updating in the future, the multi-functional database allows managing other aspects of the environment such as air, soil and noise with further time-series data. The compiled thematic maps can serve as a means of communication and comments on environmental issues to local and central institutions in Kosovo, support management decisions, and implement environmentally responsible policies. The established spatial database is open and extendable, enabling improvement by involving new mathematical
models and spatial data to analyze the environmental and other related spatial phenomena by all interested experts and institutions/organizations without any limitation.

**Conflict of interest**

The authors declare no conflict of interest.

**References**


