Development of a GIS System Prototype for Spatiotemporal Analysis of Seismic Events

Justyna Gorniak-Zimroz 1, Agnieszka Wylomanska 2, Pawel Stefaniak 2, Anna Michalak 2

1 Wroclaw University of Science and Technology, Faculty of Geoengineering, Mining and Geology, 27 Wybrzeze Wyspianskiego St., 50-370 Wroclaw, Poland
2 KGHM Cuprum Research & Development Centre, System Analytics Department, 2-8 Gen. Wl. Sikorskiego St, 53-659 Wroclaw, Poland

justyna.gorniak-zimroz@pwr.edu.pl

Abstract. The paper presents the selected functionality of the GIS system prototype developed to support the tracking of seismic activity related to the mining activity of the copper ore underground mine. The system allows to perform spatiotemporal analyzes related to seismic activity, spatial mapping of seismic shocks as well as reporting of a set of statistics describing the nature of these events and their dependence on mining activity. The database design takes into account a number of factors affecting the appearance, course and intensity of seismic shocks. For this purpose, the identification of information needs in the field of monitoring of seismic events was carried out and the data sources were determined, the structure of the data stream was proposed, and methods for their integration and auto-validation were developed. The analytical module of the system focuses a number of analytical methods in the field of statistics and spatial analysis. The article presents an example of the functionality of the analytical module on the example of a selected mining plant.

1. Introduction

Underground exploitation of minerals has many consequences, and one of them is a violation of the natural balance in a rock mass. What follows are mining tremors, accompanied by the triggering of seismic vibrations that spread from the source of the tremor in all directions in the rock mass. These phenomena, called seismic phenomena, are not only of significance in underground mining plants, during the exploitation of the deposit, but may also have an adverse effect on the infrastructure located on the surface [1]. The problem of mining tremors impacting mining operations and surface facilities is very complex, and the factors that have the greatest impact on the scale of losses caused by seismic phenomena are the geological structure of the rock mass, technical condition of the facilities in the danger zone and the nature of the tremor [2]. The seismic phenomena in underground mining plants are generated by various natural and anthropological factors. This diversity causes a high complexity of the structure of mining tremors, which are usually not uniform, and they consist of various effects of the fracturing process. It is assumed that tremors in mines involve two main types of events, i.e. those directly related to mining activities and those originating from the redistribution of stresses, which are usually found within a larger area [3]. During the mining plant operation extracting the mineral using the underground method, as a rule, there are no problems to predict the distribution of elastic stresses and the associated deformations on the basis of the given geological and mining conditions. The
difficulty arises when one evaluates the risk of seismic phenomena. It should be remembered that these are inelastic phenomena, which are additionally accompanied by changes in physical properties, conditioning the development of dynamic processes leading to tremors. The development of inelastic deformation that precedes the tremor consists of several stages. However, there is no rule that says the existence of subsequent stages must lead to the occurrence of a seismic phenomenon, as it only may increase the probability of its occurrence. The next stages are identified on the basis of changes in the physical properties of the rock mass, mainly using geophysical methods. Mining tremors are an inseparable part of the underground mining of deposits. One of the most important elements when selecting minerals is registering the mechanisms of the seismic phenomena occurrence and assessing its risk in a mining plant in the future. The highest importance is attached to the possibility of predicting strong seismic events under specific geological and mining conditions [4].

In order to increase preventive actions within mining areas affected by seismic activity, it is necessary to extend the functionality of current seismic monitoring systems enabling the identification of risk factors before the occurrence of undesirable phenomena from the point of view of occupational safety by class modules of decision support systems (DSS). This article presents a spatiotemporal information system prototype that provides integration of data from various sources, as well as analysis and reference in geographical space. The structure and functionality of the system were described. Particular attention was paid to the analytical module and reporting module. Instances of analysis results were presented on the example of a selected part of an underground mine.

2. System design

Figure 1 shows a functional diagram of the proposed system prototype for spatiotemporal analysis of seismic events. This system is characterized by a multi-module structure ensuring automation and integrity of procedures based on the following modules:

- **module 1** - automatic downloading of data on seismic events, which includes procedures and tools for data integration, cleaning, format conversion, spatial reference and further archiving,
- **module 2** - a set of tools for validation and processing of seismic signals in the GIS environment,
- **module 3** - construction of spatial analysis models, calculation procedures on attribute and spatial data in the GIS system,
- **module 4** - visualization procedures, i.e. thematic mapping,
- **module 5** - reporting, sharing information.

![Figure 1. Functional diagram of the GIS system prototype.](image)

The main idea of system functionality is to obtain data, synthesize information and knowledge about seismic activity of the rock mass, and make them available to employees of the mining geophysics station. The data fusion process includes:

- **spatial data on mining galleries and technical infrastructure**, i.e. the boundaries of excavations, exploitation sections, areas and fields, seismic sensors, access roads, belt conveyors, mining, transport and ventilation shafts,
• **descriptive data for seismic events**, i.e. signal characteristics: time and date with time reference to blasting, energy, shock class, secondary features of the signal (e.g. principal components of PCA analysis) and event location,

• **historical data on seismic events**, i.e. the causes, effects and time of the event and mining works, i.e. the direction of exploitation, time and scope of the blasting works, location of the cavities caused by extraction and the plan of mining works,

• **geological conditions**, i.e. lithostratigraphy at the deposit level, changes in discontinuity, gradient and others.

The database assumes an open architecture based on a file data management scheme, thanks to which it is possible to further develop the system according the end-user needs. The database model takes into account the hierarchization of spatial objects, their relations and factors affecting the presence and intensity of seismic phenomena. A vector data model was proposed related to the concept of discrete fields. The choice is justified by uncomplicated and convenient data collection, high precision of environmental representation, availability of topological principles and wide functionality that gives the possibility of spatiotemporal analysis. Initially, the appropriate objects of the described reality were classified as the appropriate geometrical type: **point** (seismic sensors), **polylines** (access roads, belt conveyors and others), **tested area** (excavations, exploitation sections, areas and fields and others). In the logical representation, the mentioned simple elements with the same geometry are stored in the geodatabase in the form of object classes. In turn, the physical model consists of attribute tables of all objects, in which individual elements are arranged in rows and their features in columns.

Figure 2 shows the sources of input data for building the database in the proposed system. Also proposed were examples of seismic signals (time-frequency maps based on group-delay [5, 6]; the representation of the instantaneous frequency IF [7], ARMA model), which can be included in the system, based on which it will be possible to analyze seismic signals, e.g. in the context of their segmentation or classification.

![Figure 2](image)

**Figure 2.** Input data for the construction of a GIS database to support monitoring of seismic events in mining areas.

3. **Description of the analytical module**

3.1. **Basic analytical functions**

The analytical module proposes procedures for data processing for observed seismic events. For their construction, selected geoprocessing tools available in GIS systems were used. Their functionality was tested using the data obtained from a selected part of the underground mine. As a result of the analysis, thematic maps and reports were provided to assess the state of seismic activity in the considered area of the deposit mining. Figure 3 shows a flow chart when selecting the data processing and analysis procedure in the GIS system prototype for spatiotemporal analysis of seismic events [8-10].
The information used in the GIS system is closely related to their reference in the geographical space. Therefore, it is possible to perform spatio-temporal analyzes and visualize their results using thematic maps. Some statistical analyzes are spatial and raster analyzes that are inseparable from maps and visual analysis, and some of them can be considered separately as basic statistics as the mean or standard deviation. Statistical analyzes can be divided into:

- **descriptive statistics** (statistics menu options), i.e. those that can be considered in isolation from the maps and graphical approach, and those are: maximum, minimum, sum, mean, standard deviation, frequency histogram and number of elements; depending on the needs, the results are presented in a table or graphic form.
- **spatial statistics** (measuring geographic distributions) is a set of tools that are designed to answer the questions: where is the data center, what shape they take, in which direction they are oriented and whether they are scattered.
- **central feature** is responsible for finding a central point, line or tested area for a given set of data,
- **directional distribution** creates an ellipse that aims to determine the dispersion and directional trend; depending on the needs, calculations can be based on Euclidean or taxicab distance measures.
- **linear directional mean** determines the mean direction, length and geographical center for given lines.
- **centroid** identifies the geographical center (or center of concentration) for the given data.
- **mapping cluster** makes data clustering possible to identify the location of statistically significant hotspots, cold spots, outliers, etc.
- **grouping analysis** is a function that is to enable one to better understand the data; it performs the classification procedure, trying to divide the data into clusters; in addition, one can define spatial constraints; it is possible to specify the number of groups to which the observations are to be divided.

Geostatistical tools are an important functional feature of the system (geostatistical analyst) making analysis of distributions, generation of interpolation models or assessment of data quality possible before performing further analyzes. The most popular and simplest tools are going to be presented. These include: histogram, QQplot, Voronoi map.

### 3.2. Proposed data processing models

In the analytical module of the GIS system prototype for spatial analysis of seismic events, procedures of spatiotemporal analysis using the above-described data analysis methods were proposed. These are, among others:

- **identification of seismic events** consisting in obtaining spatial and descriptive information on the selected seismic event from the database. As a result, one gets a cartographic document with the location of the event together with a report describing it,
- **analysis of seismic events in a selected exploitation field** based on searching for seismic phenomena from the database in the exploitation field selected by the user. As a result, we obtain a cartographic document with the location of the selected exploitation field along with the location of seismic events and a report with data characterizing these events. One can also perform a statistical analysis of the number of events taking place in the analyzed field in a given calendar year,
• **analysis of seismic events in the selected time interval** (years, year, month, day, hour) consisting in selecting the time interval interesting to the user from the database and analyzing events occurring at that time. As a result, one gets a map with the location of events, a report and statistics on the occurrence of events along with the distribution of released energy (e.g. analysis of the number of events on a given day, broken down by hour),

• **analysis of seismic events with the greatest energy released** consisting in searching for events with a given energy in the database and obtaining a result map with their location, report and statistics with the distribution of released energy;

• **analysis of seismic events according to the location of** their occurrence consisting in the identification of the number of seismic events in the undisturbed soil, gallery, cavities caused by extraction, and face or in the exploitation fields. As a result, one gets a report on the classification of locations of event occurrence together with statistics,

• **analysis of the place of occurrence of the event with the largest number of registered seismic events.** As a result of this analysis, one gets a result map and a report with statistics,

• **analysis of seismic events relative to the duration of exploitation in exploitation fields.** As a result, one obtains a classification of fields of exploitation with respect to its duration together with event statistics.

4. **Testing of analytical models - description of the selected data analysis**

One exploitation field was selected for the analysis, in which seismic events of various energies took place in 2002-2016. Data from this field were subjected to statistical analysis. First, two classes of objects were selected from the database: *Seismic Events* and *Exploitation Fields*. Then, using the tool **Selection by Attributes**, an exploitation field was selected in which one wants to identify all seismic events. For this purpose, one builds an SQL query in the geodatabase containing data from the field class and select it. One gets the result of the query (Figure 4a). In the next step, using the tool **Selection by Location** one searches for seismic events inside the exploitation field selected for analysis (Figure 4b). As a result, one obtains a cartographic document (Figure 4c). In the next step, statistical analyzes were performed for seismic events occurring in the selected exploitation field. First, the so-called "scatterplot" graph was made in the GIS environment, where the event number was marked on the X axis, and the event energy on the Y axis (Figure 5a) and using the histogram, the energy generated by events (Figure 5b) and the number of events (Figure 5c) in the years 2002-2016 were shown. Subsequently, data on the amount of the energy of events was grouped using the **k-nearest neighbours algorithm** applied in statistics for predicting the value of a random variable and for classification. The *k*-nn algorithm consists in: comparing the values of explanatory variables for *C* observation with the values of these variables for each observation in the learning dataset, *k* selection (a predetermined number) of the nearest to *C* observations from the learning dataset and averaging the value of the explanatory variable for selected observations, resulting in a forecast. Figure 6 shows the result of energy amount classification for events in 2002-2016.
In the next step, it was verified whether the amounts of energy for seismic events can be described by the classical Gaussian distribution. In order to verify the Gaussian distribution, a visual test based on QQplot graph was used here. Figure 7a shows the QQplot graph for the energy of events in the years 2002-2016 for the selected exploitation field. As one can easily notice, quantile graphs for data and Gaussian distribution do not coincide. Therefore, energies cannot be modelled by this distribution. This conclusion is obvious. The function of Geostatistical wizard Kriging was selected for further analysis making the testing of several other distributions possible. This testing consists in comparing the basic characteristics obtained from the empirical data with the characteristics for the tested distributions. Probability densities (theoretical and empirical), cumulative distribution functions and QQplot graphs
were taken into account here. As an example for modelling the amount of energy for events in 2002-2016, the t-student distribution was chosen here. One should bear in mind that the t-student distribution has one parameter, n, specifying the number of degrees of freedom. The random variable X has a t-student distribution if its density is given by the following formula:

\[
f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad t > 0
\]

(1)

In the above formula, \(\Gamma\) is a Gamma function. With the increase of parameter n, the t-student distribution takes on the Gaussian distribution features. Figure 7b-d shows the comparison of density, cumulative distribution functions and quantile plot determined on the basis of data describing energy quantities with appropriate characteristics for the t-student distribution. Also in this case, the distribution is inadequate for the analyzed data.

Figure 7. QQplot graph for event energy amount (a), comparison of histogram (b) and empirical cumulative distribution function (c) for energy amount with t-student distribution density as well as QQplot graph comparing the following quantiles: empirical determined on the basis of energy amount after transformation using t-student distribution, and theoretical from normal distribution (d) for seismic events from 2002-2016 for a selected exploitation field.

It is worth noting that one should not model seismic data over a long period of time with one model. It was noticed that the nature of the data changes over time and it is necessary to propose models whose coefficients change over time. Therefore, in further analyzes, data will be modelled describing the waiting times for seismic events and the amount of event energy for the selected period (calendar year). The energy amounts for 2015 were the first to be analyzed. The function of Statistics menu options makes it possible to determine simple descriptive statistics and to illustrate a histogram of the analyzed data (Figure 8a). Like for the data from the period 2002-2016, it was also verified here whether the Gaussian distribution can be used for data modeling in 2015. Using the function of Geostatistical Analyst, a QQplot graph was made comparing empirical quantiles of data for 2015 and quantiles of the Gaussian distribution (Figure 8b). It is easy to notice that also here Gaussian distribution is inadequate for the analyzed quantities. In the next step, it was verified whether other available distributions could describe energy related data for 2015. The t-student distribution was taken into account here again.
together with the lognormal distribution. Figure 8c-d shows a comparison of QQplot characteristics for the two mentioned distributions. The obtained results unambiguously indicate that the data describing the amount of event energy for 2015 do not show the properties of either t-student distribution or lognormal distribution. Therefore, to model this type of data it is necessary to use a different distribution - Pareto distribution.

![QQplot comparison](image)

**Figure 8.** Descriptive statistics determined for the energy amount from events in 2015 for the selected exploitation field

The MATLAB environment was used to make the Pareto distribution. Figure 9a shows the QQplot graph comparing empirical quantiles for data describing event energies in 2015 with the theoretical quantifiers of the Pareto distribution, and Figure 9b shows comparison of tails (1- cumulative distribution function) of distributions for event energy in 2015, Pareto distribution and Gaussian distribution (log-log scale). This graph clearly indicates that the Pareto distribution reflects the nature of the data to a much greater extent. The analysis in the GIS environment also included data on waiting times for the next event. Statistical analysis, as in the case of event energy, consists in defining basic data descriptive statistics and an attempt to select the appropriate distribution for the data. Because the data describing the waiting times also exhibit volatility over time, the data for 2015 were analyzed as an example. The waiting times divided into 5 groups are marked on the map in Figure 8a. Figure 9 shows simple descriptive statistics for waiting times together with a histogram. As in the case of the energy of
events, it was first verified whether the Gaussian distribution can be used to model the waiting times. Like before, QQplot visual test was used here (Fig. 9b). It is easy to see that Gaussian distribution is not adequate for the analyzed data. In the next step, the t-student distribution was also tested and densities, cumulative distribution functions and empirical quantiles were compared for waiting times and for t-student distribution.

**Figure 9.** QQplot graph comparing empirical and theoretical quantiles from Pareto distribution (a) and comparison of tails of empirical distribution, Pareto distribution and Gaussian distribution (b) for data describing event energies in 2015.

In this case, the t-student distribution cannot be used to model data describing waiting times for seismic events in 2015. Therefore, the Gamma distribution available in the MATLAB environment was used in

**Figure 10.** Simple descriptive statistics determined for the waiting times for events in 2015 in the GIS environment: histogram (a), QQplot (b), comparison of the histogram with the t-student distribution density (c), comparison of the empirical cumulative distribution function with the cumulative distribution function of t-student distribution (d) and QQplot comparing empirical quantiles after transformation using t-student distribution and the theoretical ones from normal distribution (e)

In this case, the t-student distribution cannot be used to model data describing waiting times for seismic events in 2015. Therefore, the Gamma distribution available in the MATLAB environment was used in
the next step. The results of the analyzes are shown in Figure 10. Gamma decomposition can definitely be used to describe the waiting times for seismic events in the analyzed period. The Kolmogorov-Smirnov test for the Gauss distribution returned the p-value of 0, whereas for the Gamma distribution of 0.65.

5. Conclusion
The article presented a prototype of a system based on GIS technology making a spatiotemporal analysis of seismic phenomena occurring in a selected area of mining activity possible. This activity concerns the underground mining for copper ore by the room-and-pillar system. A modular structure was proposed for its construction ensuring automation and integrity of procedures. The system structure consists of five modules. The article described selected data processing procedures for seismic phenomena taking into account spatial and descriptive statistics available in the proposed analytical module. Sources of obtaining input data for the construction of a database reflecting the features of the analyzed research facility in the form of an underground copper ore mine were proposed. The functionality of the proposed GIS system prototype within the selected area of the mining plant was also shown. For this purpose, data analysis was performed in a selected exploitation field for a given period of time (years, year) as a result of which cartographic documents were obtained with the location of the selected exploitation field along with the location of seismic events occurring in this field.

Acknowledgment
This article was financed with the research projects: no. 0401/0174/16 Data mining in modelling environmental changes and no. 0401/0123/2017 Research on natural and anthropogenic environmental changes using geographic information technology.

References
[1] J. Blachowski, W. Milczarek, and P. Stefaniak, “Deformation information system for facilitating studies of mining-ground deformations, development, and applications,” Natural Hazards and Earth System Sciences vol.14(7), pp.1677-1689, 2014.
[2] L. Stolecki, “Studies on the distribution of vibration parameters generated by mining shocks in LGOM mines,” CUPRUM Scientific and Technical Mining Journal of Copper Ore, vol. 60/2011, 2011.
[3] S. Lasocki, B. Orlecka-Sikora, “Seismic hazard assessment under complex source size distribution of mining-induced seismicity,” Tectonophysics vol. 456, pp. 28-37, doi: 10.1016/j.tecto.2006.08.013, 2008.
[4] H. Marcak, “Reverse problems in modeling mining shocks,” Conference materials "Natural hazards in mining", 2009.
[5] M. Polak, J. Obuchowski, M. Madziarz, A. Wyłomanska, R. Zimroz, “Differences In time-varying group delay of seismic signals,” Vibroengineering PROCEEDIA, Vol. 6., pp. 206-211, 2015.
[6] M. Polak, J. Obuchowski, M. Madziarz, A. Wyłomanska, R. Zimroz, “Time-varying group delay as a basis for clustering and segmentation of seismic signals,” Journal of Vibroengineering, Vol. 18, Issue 1, pp. 267-275, 2016.
[7] J. Sokolowski, J. Obuchowski, M. Madziarz, A. Wyłomanska, and R. Zimroz, “Features based on instantaneous frequency for seismic signals clustering,” Journal of Vibroengineering, Vol. 18, Issue 3, pp. 1654 1667, 2016.
[8] M. N. Demers, “Fundamentals of Geographic Information Systems,” 3rd Edition, John Wiley and Sons, 2005.
[9] I. Heywood, S. Cornelius, S. Carver, “An Introduction to Geographical Information Systems”, 3rd Edition, Pearson Prentice Hall, 2006.
[10] D.J. Maguire, J. Dangermond, "The functionality of GIS, in: Maguire D.J., Goodchild M.F., Rhind D.W. (editors) Geographical Information Systems: principles and applications,” Longman, Londyn, 1991.