The value of geological data, information and knowledge in producing landslide susceptibility maps.

Tsangaratos P.  National Technical University of Athens, School of Mining and Metallurgical Engineering, Department of Geological Studies

Koumantakis I.  National Technical University of Athens, School of Mining and Metallurgical Engineering, Department of Geological Studies

http://dx.doi.org/10.12681/bgsg.10992

Copyright © 2016 P. Tsangaratos, I. Koumantakis

To cite this article:

THE VALUE OF GEOLOGICAL DATA, INFORMATION AND KNOWLEDGE IN PRODUCING LANDSLIDE SUSCEPTIBILITY MAPS

Tsangaratos P.1 and Koumantakis I.1

1National Technical University of Athens, School of Mining and Metallurgical Engineering, Department of Geological Studies, ptsag@metal.ntua.gr, koumantakisoannis@gmail.com

Abstract

The paper attempts to illustrate the value of Geological data, information and knowledge in studies that investigate landslide phenomena and how this is reflected in the production of landslide susceptibility maps. In particular, the concepts of data, information, knowledge and wisdom are defined and also the process of converting data into information and information into knowledge. The presented study continuous with a summary that introduces landslide phenomena and the available techniques and methods for identifying and classifying these natural phenomena. It describes the significance of landslide analysis and how this process is essential in order to orient intelligence decisions regarding strategies for regional and urban development. The susceptibility maps that are considered as the outcomes of the over mentioned process, reflects the Geological Knowledge that has been produced by involving expert knowledge and processing data.

Key words: Landslide assessments, data processing, expert knowledge.

1. The Concept of Data, Information, Knowledge and Wisdom

The concepts of data, information, knowledge and wisdom, which can be found in the process of Knowledge Discovery, are concepts mentioned in the research field of Information Science. Many definitions have been presented for these four concepts. The existence of different definitions is due to the different background of the researcher and their specific objectives that they analyze (Zins, 2007). One of the first reports on the hierarchical structure that these concepts follow and the potential link between them was made by Zeleny (1987) and Ackoff (1989) (Figure 1). According to Zeleny (1987) and Ackoff (1989), the data are non - organized or not - processed...
events, associated with objects identified in the real world. They are considered as measurable values of variables and parameters that describe objects or events, characterized as "static" since, after being recorded, they do not change. To convert Data into Information requires a frame of reference and a conceptual model that allows their interpretation. Information is a set of Data that has undergone some form of processing and formatting. Unlike "static" data, information has meaning, purpose and relevance. Data are converted into Information mainly by following five (5) basic procedures (Davenport & Prusak, 1998):

- By placing them within a conceptual frame of reference, determined by the reason and the purpose for which the collection of data is done.
- By managing data in a more compact format, eliminating thus unnecessary details.
- By analyzing data, probabilistic or statistical data analysis.
- By categorizing or classifying data in specific classes or categories.
- By processing data.

Knowledge is referred as the Information that has undergone specific validation process. Knowledge is actually based on Information that has been organized, synthesized or summarized in order to increase understanding and awareness. Knowledge highlights the importance of Information and links Information with actionable conclusions (Benet & Benet, 2004). The usual questions about Data and Information are related to who, what, where and when, while questions referring to Knowledge are related mainly to the how and why.

Figure 1 - Hierarchical structure of data, information, knowledge and wisdom.

To reach the required wisdom reasoning is needed. Reasoning is the process by which we use existing knowledge in order to make inferences an integral component of the process that we call intelligence. It is the attempt to identify the consistency between two or more judgments, which are appreciated as the logical continuation of these judgments. It is the ability to recognize the truth and make sound judgments based on prior knowledge, experience and insight (Leibowitz, 2003). In particular, Knowledge is in a sense the familiarity which one acquires from an object or a phenomenon, the theoretical and practical understanding.

2. The Concept of Geological Data, Information and Knowledge

In complete correspondence with the foregoing, the Geological Information is the result of processing the primary geomorphological and geotechnical data. Raw data, non-organized associated with the geological environment, geological phenomena, geological and physical processes. Specifically, they could express values of measurable physical and mechanical parameters corresponding to certain geological formations. The collection of primary raw geological data and geological information is undertaken through:
Reviewing the literature, studying geological maps, reports and geotechnical studies, environmental studies, examination of aerial photographs and satellite images, etc.

Mapping of geological structures, lithological formations, tectonic features and areas of risk.

Explore through sampling techniques, and performing appropriate field and laboratory tests (to assess the physical and mechanical properties of rocks and soils).

Geophysical exploration through which one can achieve rapid preliminary identification of subsurface conditions.

The Geological Information is reflected primarily on maps, analogue or digital, adopting characteristics and technical symbols in diagrams and tables. Geological and Geotechnical particular maps are the main means of communication and flow of geological information between different disciplines (van Westen, 2004). The use of Geological Information can solve practical and theoretical issues in geosciences. Issues related to slope stability studies, forecasting geological hazards (Rosenbaum & Culshaw, 2003, Koumantakis et al., 2005), estimation of hydrocarbon reserves or aquifers (Mukerji et al., 2001, Schon, 2004), assessment of the geotechnical behavior the rock mass and soil formations (Tzamos & Sofianos, 2006, 2007, Ili et al. 2009), assessment of suitability for the construction of underground space for various uses (Benardos & Kaliampakos, 2006, Koumantakis et al., 2007), decision making related to the design of underground infrastructure (Panou & Sofianos, 2002a, 2002b) etc. According to Wood & Curtis (2004), in most cases, the Geological Information is provided as an advance element, a priori component, a key component of the "function" that solves any geo-environmental problems. The use of reliable Geological Information is the key to successful implementation of the solution that is proposed for a given problem. As the authors note, where Geological Information is detailed and accurate, then and only then you can provide rational decisions with respect to critical environmental issues. The importance of the Geological Information and Knowledge in the methods and practice of Engineering Geology can be summarized into three (3) "equations" that are considered the central point of the philosophy of Engineering Geology, as currently configured towards safety and economics in any geotechnical project (Rozos, 1989, Rozos & Tsiambaos, 1991):

- Geological and geotechnical properties of the geological materials + 'mass characteristics' = geotechnical properties of geological 'mass'.
- Geotechnical properties of geological 'mass' + Environmental (climate, etc.) conditions = determining the precise location of the structure.
- Determine the exact location of construction + induced changes in the geological environment of the construction work = final mechanical response.

2.1. Transforming Geological Data into Geological Information and Knowledge

According to Wood & Curtis (2004), the transition from Geological Data to Geological Information requires three (3) basic elements:

- Input data managed by experts and specialists.
- The incorporation of uncertainty in the generated Geological Information.
- The Geological Information to be expressed in quantitative terms, whereas when plotted in a qualitative manner it must be fully defined so that the associated uncertainty has been assessed in the final output.

When these three (3) elements are present the quality of information generated increases and the configured information becomes more relevant and meaningful. For the transition of Geological Information in Geological Knowledge it is required to manage the Geological Information within a specific theoretical and mathematical background, whose most important result is the correlation
of information to actionable conclusions. The quality of knowledge generated reflects the understanding of the mechanism that drives the behavior of the environment. Increased quality implies more confident estimation of the response of the environment. The predictive modelling and simulation is designed to capture the knowledge of quantitative and qualitative characteristics, the application of which involves the necessary special checks for verification and certification of Information. Such transformation is shown in figure 2, where the 14 lithological units identified in the field surveys have been classified into 5 geotechnical units with the help of expert knowledge and analysis of laboratory and field tests. This process provides valuable knowledge.

Figure 2 – Converting Geological Information to Geological Knowledge.

3. Managing Landslide Related Data

3.1. Landslide Phenomena

Landslides are one of the most complex and chaotic natural phenomena. Landslides are considered as unexpected and in most cases unpredictable movement that usually occur on unstable surface layers, making them one of the most frequent natural hazards with significant consequences to human life and incalculable social-economic consequences. According to Hutchinson (1995), landslides are the result of evolution or progressive extreme events that occur on the surface of the earth and due to the action of geological, geomorphological, climatic processes and also the negative impact of human activities.

3.2. Data Related to Landslide Phenomena

The process of identifying the causative factors and the related data is the basis of many methods of instability assessment. As for the nature of the data that one must investigate, they can be distinguished as dynamic (e.g. pore-water pressure), or passive (e.g. rock structure) and may also be considered in terms of the roles they perform in destabilising a slope (Crozier, 1986). In this sense, the factors from which data and information are collected could be recognised as pre-conditioning factors (e.g. slope steepness), preparatory factors (e.g. deforestation) and triggering factors (e.g. seismic shaking). The selection of the appropriate factors depends on the objective of the study and the complexity of the research area, the type of mass movement and the availability of existing data. It also depends by the analysis and method contacted. Morphological factors such as geometry (steepness, height, length, form, angle, and aspect) and basin/sub-basin characteristics (order, magnitude, channel gradient) reflect the past and present geomorphological evolution of an area and provide fundamental data and information on landslide process. The acquisition of geological factors, such as rock composition, texture, structure, degree of weathering, fracture density/foliation attitude and stratigraphic settings is still not facilitated by the application of the GIS technologies but relay mainly on the traditional procedure of field
surveys and aerial photo – interpretation. These kinds of data are essential in any instability assessment. Table 1 gives a schematic overview of the main data layers that are required for a landslide susceptibility, hazard and risk assessment and also provides information about the importance of the data set for heuristic models, statistical models, physically-based models and probabilistic models (van Westen et al., 2008).

**Table 1 – Data layers related to landslide phenomena (van Westen et al., 2008).**

<table>
<thead>
<tr>
<th>Data</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main type</strong></td>
<td><strong>Data layer</strong></td>
</tr>
<tr>
<td>Landslide inventory</td>
<td>Landslide inventory</td>
</tr>
<tr>
<td>Landslide activity</td>
<td>High important</td>
</tr>
<tr>
<td>Landslide monitoring</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Environmental factors</td>
<td>DEM</td>
</tr>
<tr>
<td>Slope geometry</td>
<td>High important</td>
</tr>
<tr>
<td>Slope hydrology</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Lithology</td>
<td>High important</td>
</tr>
<tr>
<td>Geological structure</td>
<td>High important</td>
</tr>
<tr>
<td>Faults</td>
<td>High important</td>
</tr>
<tr>
<td>Soil type</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Soil depth</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Land use</td>
<td>High important</td>
</tr>
<tr>
<td>Triggering factors</td>
<td>Rainfall</td>
</tr>
<tr>
<td>Rainfall</td>
<td>High important</td>
</tr>
<tr>
<td>Temperature/ evapotranspiration</td>
<td>Not relevant</td>
</tr>
<tr>
<td>Earthquake data</td>
<td>-</td>
</tr>
<tr>
<td>Ground acceleration</td>
<td>High important</td>
</tr>
</tbody>
</table>

3.3. Transforming Data Related to Landslide Phenomena into Information


- Landslide leave distinct mark on the surface of the earth, most of which can be identified, classified and mapped by field work or the analysis of stereoscopic aerial photographs.

**XLVII, No 3 - 1533**
The morphological profile of a slope failure depends on the type and the rate of movement. In general, similar types of geological failures produce similar morphological profiles. By the presence of a geological failure, an expert can extract qualitative and quantitative information on the degree of activity, age and extent of geological failure.

Slope failure does not occur randomly. It is the result of the interaction of natural processes that is a subject to mechanical and physical laws that can be determined empirically or statistically. The knowledge generated during the examination of the slope failures can be generalized.

Slope failures of the past and present are evidence to predict future failures. Therefore, the theoretical background and the mathematical model which are involved in the interpretation of the phenomenon of landslides are based on the following three (3) critical assumptions (Varnes et al., 1984, Hutchinson, 1995, Guzzetti et al., 1999, Chanon et al., 2006):

- Areas that in the past have a history on instability problems show great potential to failure.
- These areas, share common geotechnical characteristics.
- Similar conditions produce similar patterns of failures.

Within this conceptual frame of reference, data related to landslide phenomena could be analyzed, categorized or classified and transformed into valuable information.

3.4. Transforming Information to Knowledge by Applying Methods of Analysis

Knowledge related to landslide phenomena is based on information that has been organized, synthesized or summarized by applying specific methods of analysis. These methods produce outputs of certain format that increase our understanding and awareness regarding with landslide phenomena. A variety of methods, qualitative or quantitative, direct or indirect are implemented in order to rank slope – instability factors and assign different susceptibility and hazard levels to a region that is going to be analyzed (Carrara, 1989, van Westen et al., 2006). These methods embody expert-based analyses or complex mathematical-based systems. Particularly, geomorphologic analyses and direct field mapping methods are considered qualitative methods since they don’t provide numeric output with reference to landslide assessment. On the other hand, quantitative methods such as deterministic analyses, probabilistic approaches, statistical methods and data mining techniques, rely on mathematical models and produce numeric outputs that contain high value information (Glade et al., 2005, Tsangaratos et al., 2011).

Determining the spatial and temporal extent of landslide hazard requires to identifying areas which are, or could be, affected by a landslide and to assess the probability of such landslide occurrence within a specified period of time. To specify a precise time frame for the future occurrence of a landslide can be difficult. As a result, landslide hazard has often been represented by landslide susceptibility, where only the predisposing and preparatory landslide causes are described. Susceptibility zoning could be thought as a process that provides the spatial distribution and rating of the terrain units according to their propensity to produce landslides. It refers to the process that provides highly valued knowledge that depends on the topography, geology, geotechnical properties, climate, vegetation and anthropogenic factors (Fell et al., 2008). The value of a landslide susceptibility map refers to its information and data content. The content of information depends mainly on the type of data shown, their quality and the extent to which the information is new and essential. According to Guzzetti et al. (2000), a landslide susceptibility map is valuable when the information and data shown are useful to the user and also when the map is both relevant and fully understood by the user. In more detail, reliable susceptibility maps could provide an explanation for the observed landslide spatial distribution pattern; provide in other words knowledge of the instability state of the surroundings. In addition, by incorporating information on the instability factors that are known or supposed to control landslide spatial occurrence and...
their underlying relationship, susceptibility maps are capable of predicting the location of landslides even in the areas where landslides were not recognized or mapped providing in a sense *Wisdom*.

Figure 3 – Four different methods of analyzing landslide phenomena and producing highly valued Knowledge.

The four models that are shown in Figure 3 produce different landslide susceptibility maps and provide different levels of Knowledge. The examples refer to an area that exhibits instability problems located in the prefecture of Xanthi, Greece (Tsangaratos, 2012). The level of Knowledge is clearly associated with the performance, indicated by the accuracy index. The first two models, Analytic Hierarchy Process (AHP) and Rock Engineering System (RES) are methods that embody expert-based analyses. They perform well concerning the predictive accuracy and the associated level of Knowledge they provide is of medium to high level. The level of Knowledge is determined by the fact that there is often an overestimation of the influence that certain factors have on the landslide susceptibility index and also by the fact that the mechanism of landsliding may not be fully understood. The Certainty Factor (CF) model performs better and provides additional information and knowledge about the underlying relationship that the landslide related factors have. The Artificial Neural Network model (ANN mode) has a higher performance and although it does not provide any additional information about how or why this happens it provides valuable Knowledge of where it will happen.

Table 2 shows the calculated weight of coefficient of the factors used in the four models for the same research area. The weights represent the significance each factor has during the estimation of landslide susceptibility. It is obvious that all models identify the lithological feature as the most significant despite the different approach they require during the estimation of the susceptibility.

The expert-based method, AHP seems to overestimate the influence of the lithological factor. However, the RES method gives a lower value to the lithological factor and may be closer expressing the actual significance. If one evaluates the influence of the geological boundaries the overall influence of the geological features reaches the same level of significance (0.2932) as
estimated by the AHP method. The nature of the other two models allows a more objective way to estimate the weight coefficients and the estimated level of significance may be more accurate.

<table>
<thead>
<tr>
<th>Factors</th>
<th>AHP model</th>
<th>RES model</th>
<th>CF model</th>
<th>ANN model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lithological units</td>
<td>0.3422</td>
<td>0.1466</td>
<td>0.3142</td>
<td>0.2876</td>
</tr>
<tr>
<td>Elevation</td>
<td>0.0585</td>
<td>0.0681</td>
<td>0.0865</td>
<td>0.0975</td>
</tr>
<tr>
<td>Slope inclination</td>
<td>0.1181</td>
<td>0.1361</td>
<td>0.0899</td>
<td>0.1124</td>
</tr>
<tr>
<td>Slope orientation</td>
<td>0.0252</td>
<td>0.1047</td>
<td>0.0978</td>
<td>0.1021</td>
</tr>
<tr>
<td>Fault characteristics</td>
<td>0.1700</td>
<td>0.1257</td>
<td>0.0760</td>
<td>0.0490</td>
</tr>
<tr>
<td>Hydrological characteristics</td>
<td>0.0921</td>
<td>0.1361</td>
<td>0.0888</td>
<td>0.1011</td>
</tr>
<tr>
<td>Geological Boundaries</td>
<td>0.0921</td>
<td>0.1466</td>
<td>0.0990</td>
<td>0.0980</td>
</tr>
<tr>
<td>Infrastructure network</td>
<td>0.1038</td>
<td>0.1361</td>
<td>0.1478</td>
<td>0.1523</td>
</tr>
</tbody>
</table>

4. Discussion and Conclusion
To spatially identify landslide phenomena the process involves analyzes of a large set of interrelated data layers, often ill - known, unknown or in most cases unmappable. Landslide susceptibility is a function of the degree of the inherent stability of the slope as can be indicated by the factor-of-safety along with the presence of causative factors capable of reducing the excess strength and ultimately triggering movement. As illustrated in the present study, it is essential to produce landslide susceptibility maps, since they provide an explanation for the observed landslide spatial distribution pattern and the interrelationship between the factors that control instability. To extract Knowledge from any landslide susceptibility, hazard or risk assessment, it is necessary to develop a conceptual frame of reference that will enable the transformation of data, related to landslide phenomena, into information. A conceptual frame that is based on the belief that past and present provides evidence for the future, that failures do not occur randomly, that failures share common geotechnical characteristics and that similar conditions produce similar patterns of failures. The level of Knowledge that susceptibility maps provide is analogous to their usefulness and relevance and how well the explanation that they provide concerning the mechanism of instability is understood by the user. Each factor contributes to the overall susceptibility by a different level of significance, expressed with a weighted coefficient that is estimated through specific procedures according to different models. The provided Knowledge should be also accompanied by information of when a landslide will occur, and how large or destructive the mass movement is expected to be. This information is transformed from data that concern triggering factors and when analyzed produce landslide hazard and risk maps that provide a higher level of Knowledge. When landslide susceptibility maps predict successfully the location of landslides in areas where landslides were not yet recognized or mapped they provide in a sense Wisdom. In conclusion the landslide susceptibility maps provide to engineering geology experts a cost and time effective mean for tracing the most critical slope site with inherent instability potential. They may also serve as a valuable tool in decision-making, regarding the land use and development planning processes in landslide susceptibility areas.

5. References


