Polish Geological Institute Special Papers, 15 (2004): 13–18
Proceedings of the Conference "Risks caused by the geodynamic phenomena in Europe"

LANDSLIDE SUSCEPTIBILITY ASSESSMENT: A CASE STUDY FROM BESKID NISKI MTS., CARPATHIANS, POLAND

Teresa MROZEK[1], Simone POLI[2], Simone STERLACCHINI[3], Lesław ZABUSKI[4]

Abstract. This study involves the integration of multiple thematic datasets for landslide susceptibility assessment through spatial prediction models. The proposed methodology has been applied in the Bystrzanka-Biczyska area (Beskid Niski Mts., Carpathians Mts., Poland), characterised by a very high density of landslides. The susceptibility assessment has been based on an indirect bivariate statistical analysis ("Weights of Evidence" modelling technique, Bonham-Carter *et al*., 1989) performed in order to predict the occurrence of an event (landslide) where well-known evidences (predictor variables) are available. According to the relative importance of each evidence, a landslide susceptibility map has been produced. Observing final prediction results, it is concluded that the susceptibility map gives useful information both on present instability of the area and its possible future evolution in agreement with the morphological evolution of the area.

Key words: landslide susceptibility map, spatial analysis and prediction, Beskid Niski Mts.

Abstrakt. W artykule omówiono analizę danych z kilku warstw tematycznych do oceny zagrożenia osuwiskowego przy wykorzystaniu modelu predykcji przestrzennej. Metoda została zastosowana do danych zebranych w obszarze testowym Bystrzanka-Biczyska, w Beskidzie Niskim w Karpatach Polskich, charakteryzującym się dużą liczbą osuwisk (29% badanego obszaru). Podatność osuwiskową oceniono używając metody *weights of evidence* (Bonham-Carter i in., 1989), w której predykcję wystąpienia nowego wydarzenia (osuwiska) wyprowadza się na podstawie znanych czynników pasywnych, kontrolujących ruchy masowe (dostępne predyktory). Opracowana mapa podatności osuwiskowej jest wypadkową oddziaływania przeanalizowanych czynników.

Słowa kluczowe: mapa podatności osuwiskowej, analiza i predykcja przestrzenna, Beskid Niski.

INTRODUCTION

Landslides belong to natural hazards which bring about significant environmental and socio-economical losses in many regions of the world. This is also the case of the Polish Flysch Carpathians which are prone to mass movements. To make mitigation and prevention measures effective and to allow for sustainable development of such regions, it is substantial to delimit landslide susceptible areas. This paper discusses an application of weights of evidence model as an effective data integration method in a GIS environment in the prediction of landslide occurrences in the Bystrzanka-Biczyska area, Beskid Niski Mts., Poland.

[1] Polish Geological Institute, Carpathian Branch, Skrzatów 1, 31-560 Kraków, Poland; e-mail: teresa.mrozek@pgi.gov.pl
[2] University of Milan-Bicocca, Department of Environmental Science (DISAT), Piazza della Scienza 1, 20126 Milan, Italy; e-mail:simone.poli@uni.mib.it
[3] Italian National Research Council, Institute for the Dynamic of Environmental Processes, Piazza della Scienza 1, 20126 Milan, Italy; e-mail: simone.sterlacchini@uni.mib.it
[4] Polish Academy of Sciences, Institute of Hydro-Engineering, Kościerska 7, 80-953 Gdańsk, Poland; e-mail: lechu@ibwpan.gda.pl

PRINCIPLES OF "WEIGHTS OF EVIDENCE" MODELLING TECHNIQUE

The "Weights of Evidence" (WoE) modelling technique (Bonham-Carter *et al*., 1989), the log-linear version of the general bayesian theorem, is based on the idea of prior and posterior probabilities to solve the problem of combining multiple datasets. The prior probability is that a terrain unit contains the response variable (the landslides in this study) before considering any existing predictor variables (conditioning and triggering factors). This model is fundamentally based on the calculation of positive and negative weights (W^+ and W^-), the magnitudes of which depend on the measured association between the response variable and each class of each predictor variable. The difference between W^+ and W^- defines the contrast (C), one of the parameters for accepting or rejecting a class of a predictor variable. Given that the model is in a log-linear form, the weights can be added. In fact, after calculating the W^+ and W^- values for all the selected classes of each predictor variable, it is possible to define the posterior probability, which updates (increases or decreases) the prior probability. When several classes are combined, the areas with the greatest coincidence of low/high weights produce the lowest/greatest probability of occurrence of the response variable.

PREDICTION OF LANDSLIDE OCCURRENCES

GEOLOGICAL AND GEOMORPHOLOGICAL SETTING OF THE STUDIED AREA

The case study area (17 km^2) is located in the Polish Flysch Carpathians, close to the border between the Beskid Niski Mts. and the Carpathian Foothills (Fig. 1), drained by Bystrzanka and Biczyska streams which join the Ropa River, close to Gorlice.

The studied area is located in the marginal part of the Magura Nappe which thrusts here over the Silesian Nappe. The flysch folds are dissected by numerous faults, and then secondarily folded (Świdziński, 1973). In the marginal zone the thickness of the Magura series is 50–250 m. In the tectonic

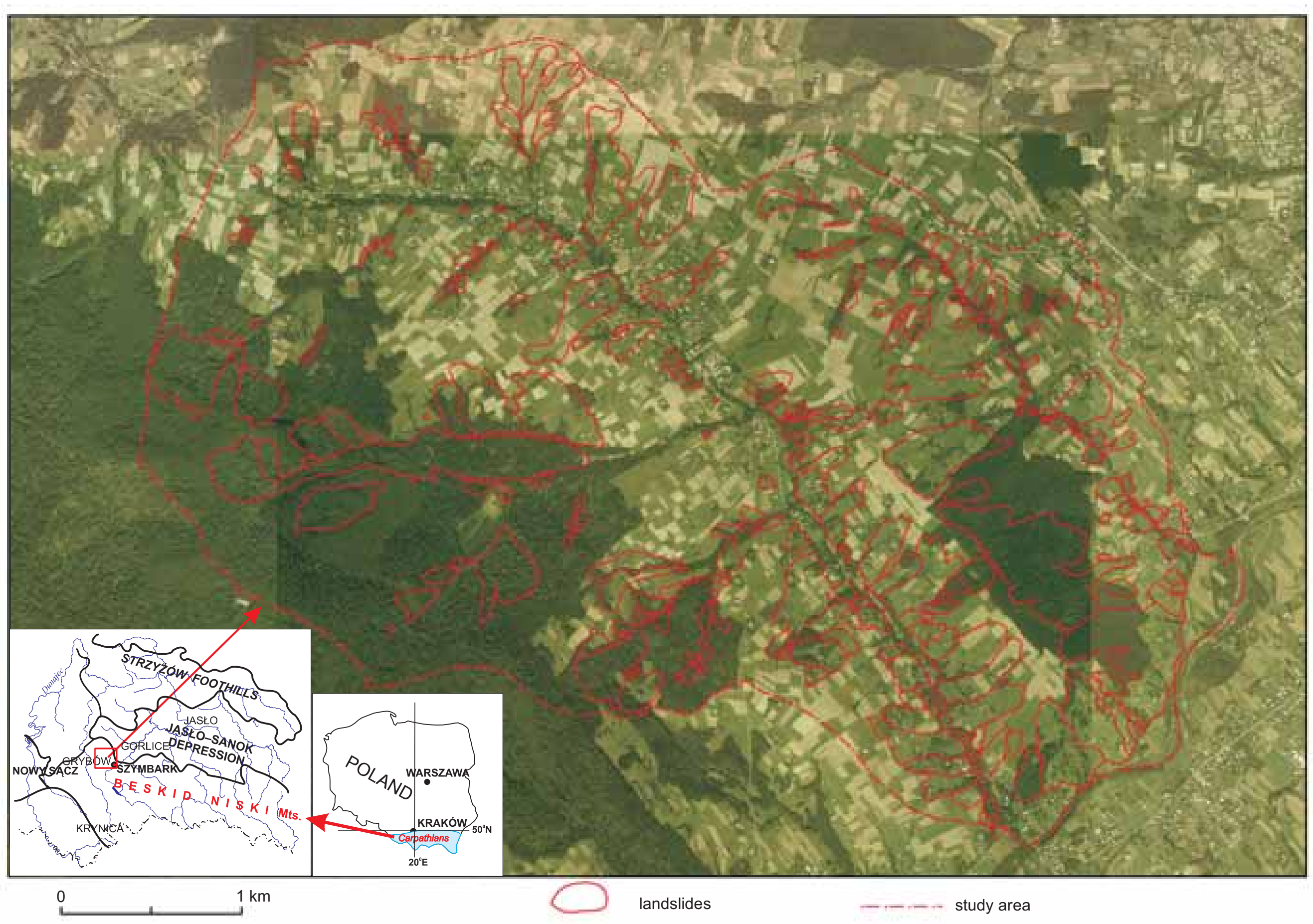


Fig. 1. Study area and landslide occurrences

window between the Magura and Silesian nappes, middle Miocene deposits are present. A syncline of Maślana Góra and an asymmetric fold of Heddy, occurring in the western part of the studied area, are the main tectonic elements. The above-mentioned Magura Unit is a major lithostratigraphic unit which comprises rock series from Upper Cretaceous to Eocene–Oligocene. Quaternary deposits are represented by loams and/or loams with debris or rubbles, colluvium and alluvia. The lithological characteristics of these units are strictly related to the relief pattern. The relief features of the studied area comprise rounded ridges (some with flattening in summit parts), flat bottom valleys of major streams and V-shaped valleys cut in bedrock, alluvial fans, landslides, terrace plains and a flood plain. Whereas hill ranges and fragments of summit flattening represent the Tertiary relief modelling, the side valleys dissecting the lower slopes are attributed to the Pleistocene development (Kotarba, 1970). Further dissection of slopes and vivid landsliding, especially during wet periods, are outcomes of the Holocene morphogenesis. The present-day relief changes result from erosion, denudation and repetitive mass movements. In fact, the studied area is characterised by a very high density of landslides (~ 29% of the studied area), mainly complex slumps and translational slips, which are rejuvenated or triggered by rainfall events.

CONCEPTUAL MODEL

The analysis is performed to evaluate the relative weight of each factor contributing to the instability considering only shallow complex landslides: this choice is strictly related to the geological-geomorphological knowledge of their conditioning factors, derived from many years of studies. The overall analysis is based on the hypothesis that failures have occurred due to the increase of pore water pressure in the soil that reduce the soil strength, which can be eventually overcome by slope driving forces. So, rainfall is the main triggering factor for the initiation of shallow landslides.

DATA COLLECTION AND PROJECT DATABASE

The dataset used to perform the susceptibility analysis consists of landslide inventory maps (training data) together with a number of thematic or evidential (predictor) maps, carrying information on landslide conditioning factors, as follow:

— landslide maps, considering a time period from 1969 to 2000, where depletion (scarp/rupture) and accumulation zones were distinguished for each landslide in the database;

— geological map, obtained by field mapping at 1:5,000 scale, where a total of 7 lithostratigraphic units was identified;

— tectonic map, depicted as faults, overthrusts and dips of the strata as derived from field measurements;

— distance from tectonic lines, buffering the main tectonic discontinuities (overthrusts and faults);

— landuse, resulted in a subdivision of the studied area into 7 classes, indicating the main landuse units that might affect the hydrological condition and soil strength;

— slope geometry, such as altitude, slope angle and aspect automatically extracted from a high precision Digital Elevation Model, originated from the State Geodetic Service, Małopolska Province.

DATA PROCESSING

Data input, storage and management for both spatial and attribute data have been carried out using ILWIS (Integrated Land and Water Information System), a GIS software developed by the International Institute for Aerospace Survey and Earth Sciences in the Netherlands. In this step, a unit cell of 10 m by 10 m seem to be a good choice in the analysis, representing a realistic size to ensure that only one landslide deposit can be present in any given pixel.

The "core" of the work, the map analysis (data integration), has been carried out using Arc-SDM (Spatial Data Modeller), an ArcView extension developed at the Geological Survey of Canada (Kemp *et al.*, 2001) for combining evidential themes to generate a response theme. This extension automatically calculates all the parameters input to the model: prior probability, positive and negative weights, variances, contrast values, and posterior probability. In addition, the main advantage of this extension is the automatic calculation of a statistical parameter (chi-squared, χ^2), that defines whether the assumption of conditional independence, which is assumed to exist among the evidential themes input to the Bayesian model, is satisfied or not. The null hypothesis of conditional independence is tested by determining if the measured χ^2 value exceeds a theoretical χ^2 value, given the number of degrees of freedom and the level of significance.

CONSTRUCTION OF THE PREDICTION MODEL AND PRELIMINARY SUSCEPTIBILITY MAP

First of all, the landslide dataset has been randomly divided in two spatial sub-datasets: the first one to build the model, the second one to validate it. Then, the "best" number of points, useful to accurately model each landslide deposit, has been pointed out.

Once tables of weights have been computed within each statistical simulation, a critical evaluation of the weights has been performed to determine the degree of association between landslide deposits and each class of each evidential map. As indicated in documentation, weights between 0.1 and 0.5 are mildly predictive, 0.5 and 1 are moderately predictive, 1 and 2 are strongly predictive, and greater than 2 are extremely predictive for the analysis. After obtaining the weights, the predictor factors are combined to generate a posterior probability map.

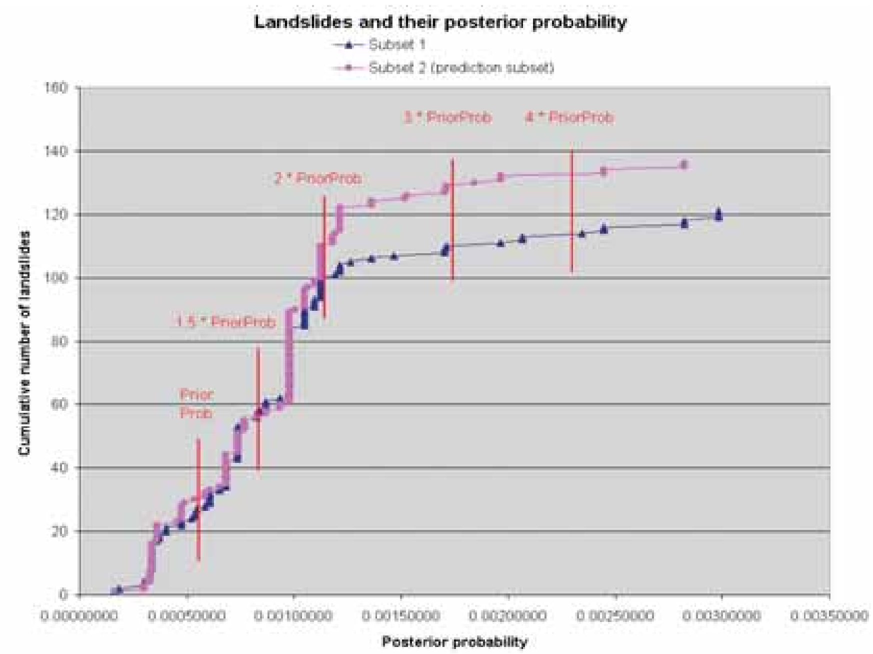
Table 1

Positive and negative weights, contrasts and their standard deviations for each class used for building the model

Evidential theme	Classes	W^+	$s(W^+)$	W^-	$s(W^-)$	C	s(C)
Slope	5–7°	–0.7187	0.3536	0.9187	0.2675	1.6373	0.4434
	18–20°	0.9187	0.2675	–0.7187	0.3536	–1.6373	0.4434
Aspect	north	–0.731	0.4083	0.1715	0.1563	–0.9025	0.4372
	east	0.6556	0.3165	–0.109	0.1645	0.7646	0.3567
Landuse	arable lands	–0.7091	0.2295	0.3958	0.1292	–1.1048	0.2633
	permanent grasslands	0.3573	0.1526	–0.3066	0.1667	0.6639	0.226
	tree clumps	0.5003	0.2427	–0.1025	0.127	0.6028	0.2739
Fault	buffer 25 m	0.7448	0.2675	–0.082	0.1079	0.8268	0.2884
	buffer 50 m	0.9749	0.2502	–0.1121	0.1091	1.087	0.273

Other useful parameters have been calculated for each simulation: standard deviations of W^+, W^-, C, and the Studentised values of contrasts — s(C). These data are the basis for the calculation of an overall test of conditional independence, which is assumed to exist, between the predictor variables with respect to the landslide occurrence points and a chi-squared (χ^2) statistical test for each combination of these themes. In this study, the level of significance of the test is taken at 95%.

Even if further studies will be necessary to focus on the definition of the "best" number of points, the choice to represent each entire landslide deposit considering only one point (its centroid) seems to give the best results and provide a more representative map of the real slope conditions, as witnessed by the



Distribution of training points (landslides) into classes of susceptibility

Subset of landslides	Total number of points	< PriorProb	< 1.5 * PriorProb	< 2 * PriorProb	< 3 * PriorProb	< 4 * PriorProb	> 4 * PriorProb
Subset 1	121	27	31	42	10	3	8
Subset 2	136	30	26	54	19	3	4

Fig. 2. Distribution of points, representing landslides, in each class of susceptibility

The computation has been made for both the subsets

χ^2 test. All the predictor factors used for constructing the susceptibility map and their values of weights and contrast are shown in Table 1. The performed pair-wise χ^2 test did not suggest that the null hypothesis of conditional independence should be rejected for any pair of themes involved in the analysis. Finally, the percentage of pixels with or without landslides (considering the first landslide dataset) distributed in each susceptibility class has been calculated, obtaining the correctly classified pixels (CCP).

At the end, 78% of landslides introduced in the model have been recognized in susceptible areas, getting a good fitting between the first landslide dataset, used to construct the model, and the calculated susceptibility map classified considering a ratio between posterior probability and prior probability (Fig. 2).

Therefore, the final landslide susceptibility map can be crossed with the second landslide dataset in order to verify the prediction power of the model. The results have been quite satisfactory, given that 79% of the pixels observed with landslides are situated in areas with a posterior probability superior than the prior probability (Fig. 3).

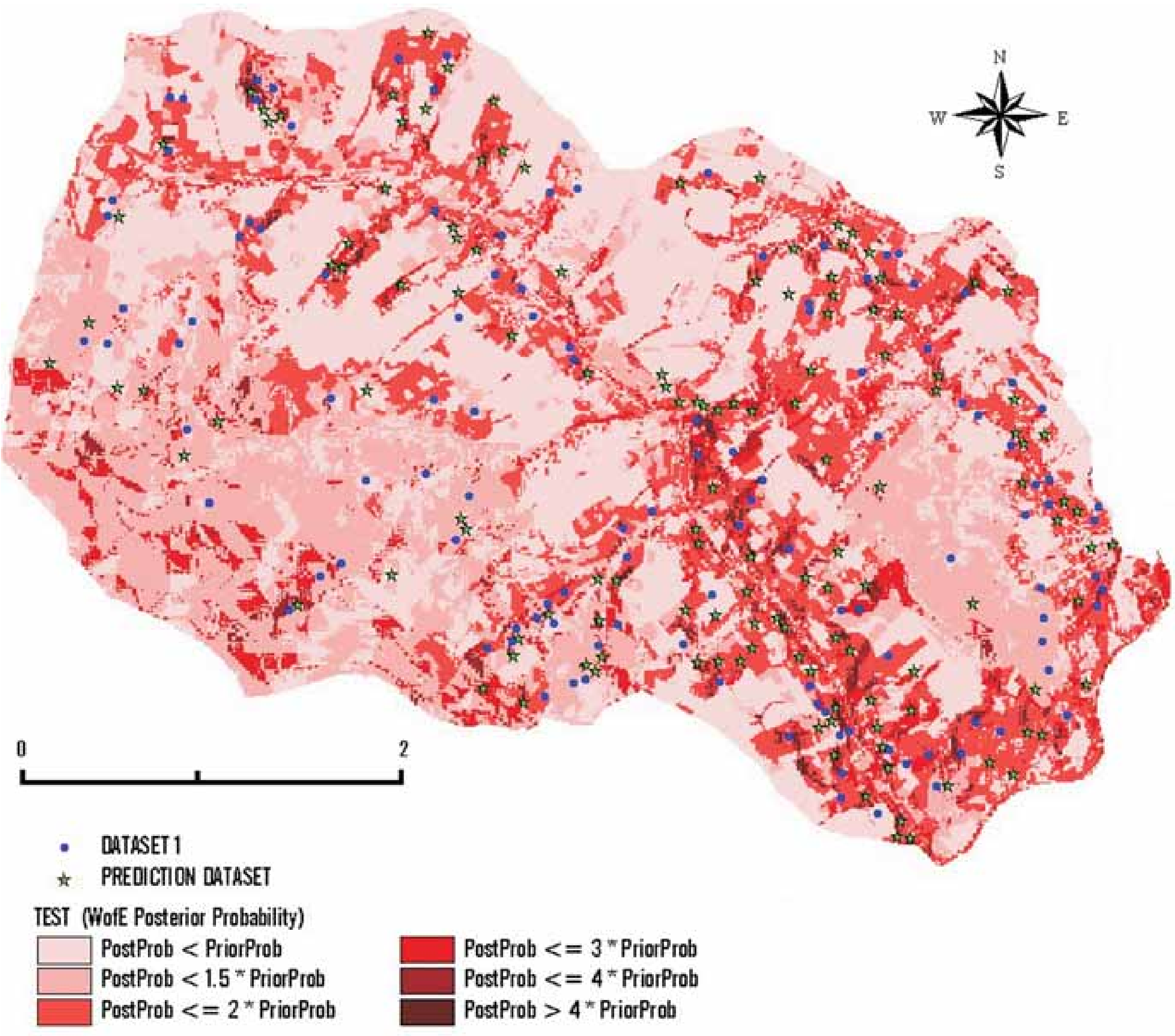


Fig. 3. Final susceptibility map constructed using 9 classes belonging to 4 different variables

DISCUSSIONS AND CONCLUSIONS

In this paper, an application of the Bayesian approach of combining predictor patterns for landslide susceptibility mapping in the Polish Flysch Carpathians is presented. After classifying the posterior probability map in terms of multiple values of prior probability, the percentage of correctly classified pixels has been calculated, resulting equal to ≈78%. Moreover, the so-obtained final susceptibility map has been furthermore validated by a cross-comparison with the second landslide dataset, evaluating the prediction capabilities of the model: ≈ 79% of the pixels are correctly classified. In this work, we have considered as correctly classified pixels the pixels falling within classes characterised by posterior probability values higher than 1.0 prior probability.

At the end, it is possible to observe how much more detailed studies and considerations have to be conducted in those areas where, within high or very high susceptibility classes, no landslides have been mapped, up to the time of the last field survey (2002). These specific situations, interesting targets for future studies, could be related to conditional independence problems (probably not completely solved) or to the absence of clear morphological evidences in the case of ancient landslides, or to "latent" instability conditions which will evolve in future occurrences.

The final susceptibility map could give useful and effective information both on present instability and its possible future evolution, in agreement with the morphological evolution of the area. This final document could be used for urban and regional territorial management and planning according to the new European policy based on the concepts of landslide prediction and prevention.

An approach used in this study is believed to stimulate devising of similar landslides susceptibility maps in a regional scale, especially in other Carpathian landslide threatened municipalities.

Acknowledgments. This contribution is completely framed within the ALARM Project "Assessment of Landslide Risk and Mitigation in Mountain Areas" EC Contract No. EVG1-CT-2001-00038, a European Community funded project that aims to have progress in prevention and mitigation of hydrogeological risks affecting inhabited areas.

REFERENCES

BONHAM-CARTER G.F., AGTERBERG F.P., WRIGHT D.F., 1989 — Weight of evidence modeling: a new approach to mapping mineral potential. Statistical Applications in the Earth Sciences (eds. F.P. Agterberg, G.F. Bonham-Carter). *Geological Survey of Canada Paper*, **89**, 9: 171–183.

KEMP L.D., BONHAM-CARTER G.F., RAINES G.L., LOONEY C.G., 2001 — Arc-SDM: Arcview extension for spatial data modelling using weights of evidence, logistic regression, fuzzy logic and neural network analysis. http://ntserv.gis.nrcan.gc.ca/sdm.

KOTARBA A., 1970 — Charakterystyka rzeźby okolic Szymbarku. *Dokumentacja Geograficzna*, **3**: 7–24.

ŚWIDZIŃSKI H., 1973 — Z badań geologicznych w Karpatach. *Pr. Geol*., **80**.