Probabilistic evaluation of the extent of the unconfined aquifer

Marek Kachnic*, Andrzej Sadurski*,**

M. Kachnic

A. Sadurski

A b s t r a c t . The research mainly aimed to present a novel application of the technique based on the Dempster-Shafer theory for the determination of the aquifer extent in a nonparametric (probabilistic) scale.
Data analyses were carried out in the Geographic Information System. All the data were imported to the IDRISI 32 release 2. The Dempster-Shafer probability theory supported by the module BELIEF of IDRISI software was applied to the algebra of pixel maps.
The research area covers 1300 km^2 in the eastern part of the Pomeranian Lakeland. The geology of this area is dominated by Pleistocene postglacial sediments. The study area comprises unconfined aquifers and four main confined aquifers. Only the unconfined aquifers were taken into account in the study described in these paper. The resulting image showed a map of the aquifers' extents in a probabilistic scale, i.e., in a range between 0 (lack of the aquifer confirmed by field research) and 1 (occurrence of the aquifer, also proved by research).

Key words: *Dempster-Shafer theory, fuzzy logic, probabilistic scale, nonparametric maps, aquifer extent*

Environmental researchers often analyse "poorly-defined" (Fisher, 1999) phenomena and objects, i.e., objects which are difficult to be assigned to a specific class of objects in compliance with dichotomic rules of binary (Aristotelian) logic. The ranges of hydrogeological elements represented in the cartographic studies are based on the point sampling or local reconnaissance performed in the field. Those limits are of probable course, more or less similar to the real boundary. Error assessment of graphic presentation of the hydrogeological elements, such as the extent of an aquifer has not been expressed in values yet. Hydrogeologic cartography provides diverse information, due to the reliability of data used. It depends on the accuracy and likelihood of estimation of the extent of groundwater bodies, their amounts and quality. Information about the reliability of hydrogeologic studies is especially useful for readers from other disciplines. Maps with nonoparametric scale offers them easy readable data about quality of source information from the study area.

In the environmental studies, proper use of information (or the lack of information) requires a way to represent such data. It was argued (Leung & Leung, 1993) that the application of Boolean logic (the all-or-nothing system) in the GIS design causes the following problems: a) it imposes artificial precision on intrinsically imprecise information, graded spatial phenomena and processes, b) it fails to determine and communicate to users the extent of imprecision and error, c) it is inappropriate to human cognition, perception and thinking processes, which are generally embedded with imprecision.

An aquifer is a good example of a poorly-defined object. This is due to the lack of information on its extent (especially for the confined aquifer) and various parameters of the aquifer due to facies changes within the aquifer.

In order to correctly describe "poorly-defined" objects in modeling, proper methods should be found. They should allow intermediate values to be defined between conventional evaluations like 1 and 0, true or false. For describing "poorly-defined" objects, we can use one of the multi-valued logic systems, such as "fuzzy logic" (Zadeh, 1965), kernel-based probability density function estimation (Brundson, 1995) or other probability methods such as Bayesian theory or Dempster-Shafer theory (Shafer, 1976; Klir & Yuan, 1997; Eastman, 1999a). This paper attempts to evaluate the extent of an unconfined aquifer in a nonparametric — probabilistic scale with help of Dempster-Shafer theory.

The main objective

The main study objective was to evaluate the probability that an unconfined aquifer may be found in each pixel location within a surface represented in the studied area. Due to a large amount of data, IDRISI 32 software was used to achieve the aim.

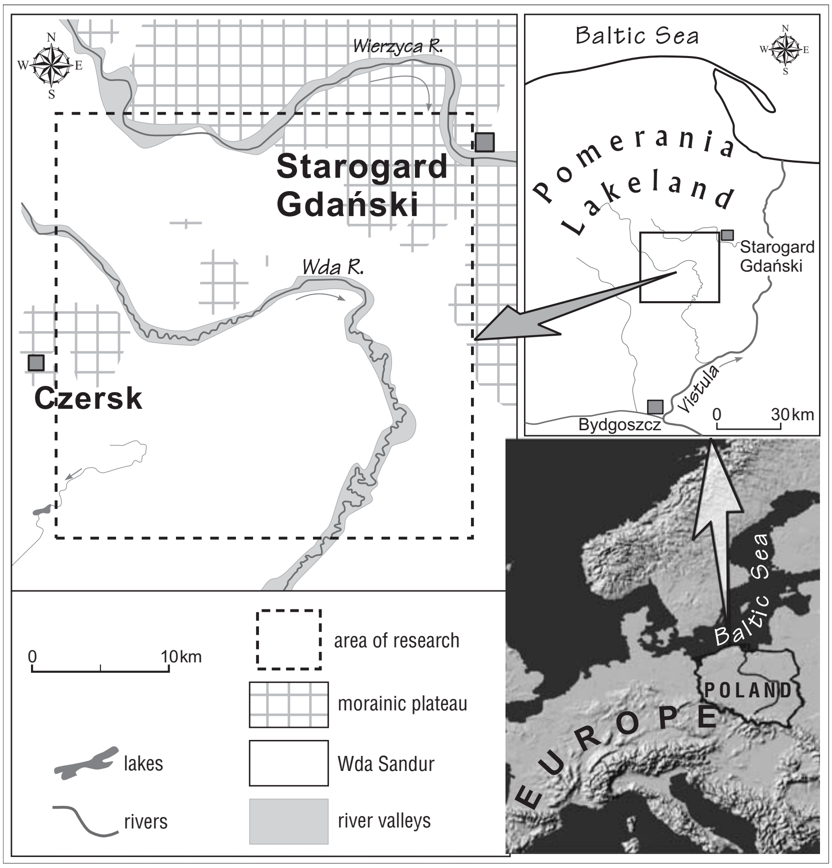
The area of research

The research area of 1300 km^2 in the eastern part of the Pomeranian Lakeland in Poland was chosen for testing this procedure. This area lies completely within the limits of the last (Weichselian) glaciation. Along with the relatively slight hypsometric differentiation, the relief of the studied area is characterized by a few forms of fluvioglacial and glacial origin. The main form is outwash sediments (the Wda sandur) and a morainic plateau (Fig. 1). Only Cenozoic water-bearing strata have been recognized within the log wells. The Pleistocene water-bearing layers form the major aquifer for the studied area. It consists of of one unconfined aquifer and a few confined units.

Methodology

The Dempster-Shafer theory is an extension of Bayesian probability theory. This theory makes a distinction between probability and ignorance and allows for the

*Faculty of Biology and Earth Science, Nicholas Copernicus University, Sienkiewicza 4, 87-100 Toruń, Poland; kach@uni.torun.pl;
**Polish Geological Institute, Rakowiecka 4, 00-975 Warszawa, Poland; andrzej.sadurski@pgi.gov.pl



← **Fig. 1.** The location of the study site

expression of ignorance in uncertainty management (Lee, et al., 1987; Klir & Yuan, 1995). The basic assumptions of Dempster-Shafer theory are that ignorance exists in the body of knowledge, and that the belief for hypothesis is not necessary to the complement of the belief for its negation. By using the "belief functions" to represent the uncertainty of hypothesis, the theory releases some of the axioms of probability theory. The resulting system becomes a superclass of the probability theory. However, it suffers from the need for large numbers of probability assignments and from the need for independence assumptions (Malczewski, 1999). Unlike Bayesian probability analysis, Dempster-Shafer theory explicitly recognizes the possibility of ignorance in the evaluation, i.e., the incompleteness of knowledge or evidence in the hypothesis (Eastman, 1999b).

The research objective was performed with IDRISI 32 ver. 2 raster-based software. In IDRISI, the BELIEF module (Fig. 6) can be used to implement the Dempster-Shafer theory. BELIEF constructs and stores the current state of knowledge for the full hierarchy of hypotheses formed from a frame of discernment. BELIEF first requires that the basic elements in the frame of discernment be defined. As soon as the basic elements are entered, all hypotheses in the hierarchical structure will be created in the hypothesis list. For each line of evidence entered, basic probability assignment images (in the form of real number images with a 0–1 range) are required with an indication of their supported hypothesis.

The development of knowledge base

The research question guides us to define the frame of discernment — it includes two elements [present] and [absent]. The hierarchical combination of all possible hypotheses, therefore, includes [present], [absent] and [present, absent]. We are most interested in the result generated for the hypothesis [present] which here means existence of the aquifer. The final results produced for the hypothesis [present] are dependent on how all evidences are related together in the process of aggregation.

Given knowledge about existing wells and given expert knowledge about the occurrences of aquifers, each evidence is transformed into a layer representing likelihood that an aquifer exists. The aggregated evidence produces results that are used to predict the presence of an aquifer and evaluate the input of each line of evidence to the total body of knowledge.

Several bitmaps and pixel maps of elements which confirm or deny the occurrence of the unconfined aquifer were prepared for this study. At the beginning, each map included separately: point, line or area data in a dichotomous scale (0 and 1). In the next stage, the information in each map was changed due to work out membership function. As a result, a pixel map with values from 0 to 1 was obtained. Finally, all the maps (information layers) were put into the BELIEF module and a probability map was compiled.

Data input for the unconfined aquifer. There is a significant difference between analysing the extent of the unconfined versus confined aquifer with the use of the GIS methods. With the exception of the wells as the indicator of aquifers, there is far more indirect evidence of occurrences for the unconfined aquifer than the confined aquifer. For example, they are: springs, rivers, lakes, the area of extent of alluvial or outwash deposits. There is high probability that the unconfined aquifer will be close to these forms.

The authors focused here only on the unconfined aquifer. Still, GIS methods are a tool for, by and large, two-dimensional data, and there is no advanced GIS raster program for analysing three-dimensional data, required for analysing the confined aquifer.

For estimating the extent of the unconfined aquifer in a probabilistic scale the following data were selected:

a) location of wells and boreholes,

b) area of the extent of the outwash and morainic plateau,

c) course of main rivers and lakes,

d) map of depth to the water table in the area where there are permeable sediments on the terrain surface.

Creating probability maps (fuzzyfication)

The stage of fuzzyfication is a procedure, which allows for converting a discrete image (bitmap) into images with a probabilistic (nonparametric) scale. The reliability of the obtained maps depends on the applied parameters of fuzzyfication controlled by a membership function. For this study the following assumptions were used:

Probability for background. Initially, for the whole research area the background value was assumed as constant 0.5. That means there is no proof for the existence of unconfined aquifer and there is no evidence for the lack of the aquifer in the research area, either.

Membership function for wells. Wells are the point markers of an aquifer. For these features, the area in close vicinity of the wells should have obtained high likehood. However, it is very difficult to find a proper way to extrapolate hydrogeologic information from point where wells exist to close neighbourhood of the wells or boreholes; this issue needs much more work. Vector maps with locations of the wells with unconfined conditions were rasterised and all the pixels in which wells were located, obtained the value "one". The pixel values are high in the area of cone depression or in the area calculated by means of an empirical formula, and finally, the pixel value decreased down to the level of the background (Fig. 2).

The empirical formula was applied as one of the assumptions for the extent of the unconfined aquifer. That was a formula for calculating depression cone, known as the Kusakin formula (Hölting, 1996; Pazdro & Kozerski, 1990):

$$R = 2s\sqrt{k \cdot H}$$

where:

R — radial distance in [m] of the depression cone;
s — the maximum drawdown observed in a well [m];
k — hydraulic conductivity, in [m/24h] and
H — thickness of the aquifer in the well log [m].

Membership function for wells and boreholes without unconfined aquifer. The value "0" was assigned to the pixels where boreholes exist and there is no unconfined aquifer noticed. Also wells which extract water form confined aquifer were added to this set of data. In the vicinity of those pixels, probability increases from "0" to the value of background within the 300 m range (Fig. 3). The above distance was established subjectively as the optimal one after analysing the geological and hydrogeological cross-section from the research area.

Membership function for boreholes and wells with present unconfined aquifer. The value "1" was assigned for the pixels where boreholes and wells exist and there is unconfined aquifer present. In the vicinity of those pixels, probability increases from "1" to the value of background for the range 300 m.

Membership function for the area of outwash sediments and morainic plateau. In the research area the main body of the unconfined aquifer is associated with fluvioglacial outwash. The area of the outwash extent was digitised from the *Geological Map of Poland in a scale 1 : 200,000* (Butrymowicz et al., 1978). The rest of the area was classified as a logic negation, which means the area without sand sediments on the terrain surface (i.e., morainic plateau).

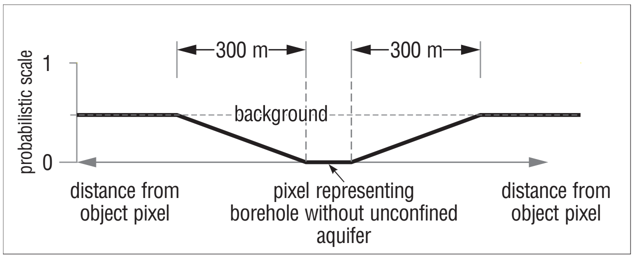


Fig. 3. Graph of the membership function for boreholes and wells without unconfined aquifer

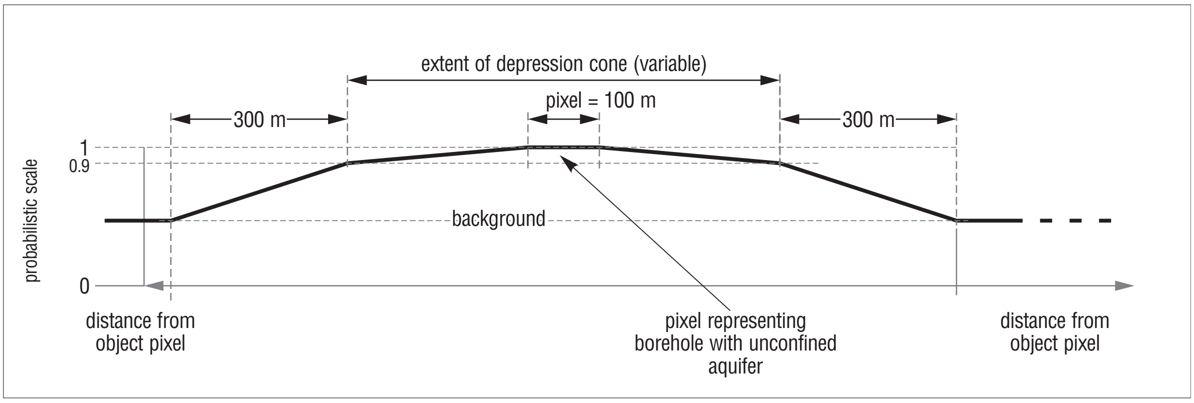


Fig. 2. Graph of the membership function for wells with unconfined aquifer

Arbitrarily the value "0.8" was assigned to all the pixels which represent the area of outwash sediments and the river valley (Fig. 4). For the remaining area a constant value "0.3" was established *a priori*.

Membership function for area in the vicinity rivers and lakes. Rivers and lakes are hydrologic objects with frequent connection to the aquifer, especially the unconfined aquifer. Close to a river or a lake there are often sand sediments with the aquifer, therefore, this vicinity to water indicates the plausibility for the aquifer. Only rivers that are longer than 5 km and lakes with the area larger than 1 ha were analysed.

Simple geostatistical methods were used in order to develop the relationship between the distance to water and the locations of the wells. After analysing the histogram of appearance of wells depending on distance from river banks or lake shores, the authors found that there should be higher likelihood (the value of 0.8) within the 200-meter-wide zone from around river banks or lake shores (Fig. 5).

Membership function for area from maps. The information about the depth to the water table and the extent of the area where there are no impermeable sediments on the terrain surface served as an additional source of data.

Data was taken from the digital *Hydrogeological Map of Poland in a scale 1 : 50,000*. These is a new map, prepared and stored in GIS system as a multisheet map. From 1994 to 2004, 1069 sheets covering the whole area of Poland were completed. The map is based on the concept of the main usable aquifer which is a productive aquifer meeting the following criteria: thickness at least 5 m, transmissivity at least 50 m^3/24h, and potential discharge of a well at least 10 m^3/h. All data is kept in 19 information vector layers, which contain among others: topographic situation, well and spring locations, type of the aquifer, water quality classes, aquifers pollution risk classes, land use and hydrodynamic information, e.g., hydraulic head, groundwater flow directions and transmissivity distribution (Paczyński et al., 1999).

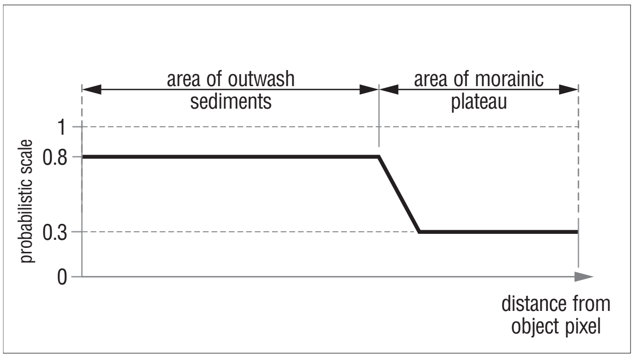


Fig. 4. Graph of the membership function for area of outwash and moraine plateou

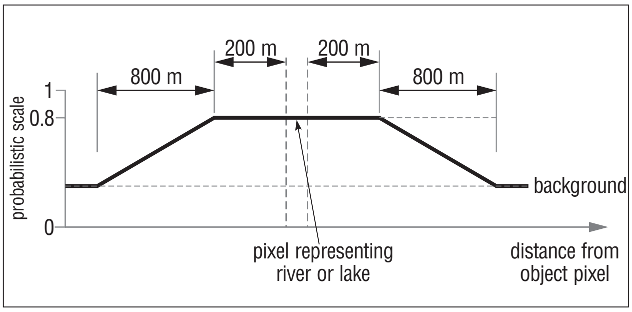


Fig. 5. Graph of the membership function for hydrological objects

Stage of calculating. All prepared pixel maps were put to the BELIEF module (Fig. 6).

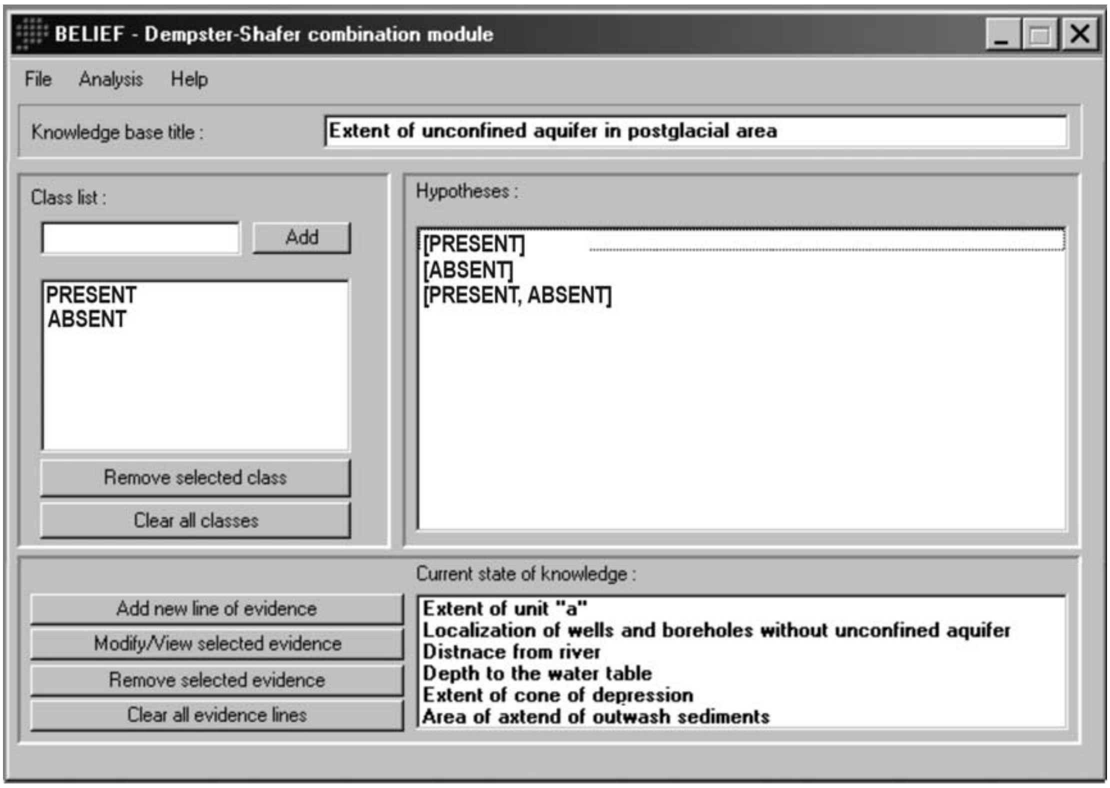


Fig. 6. The window of BELIEF module of IDRISI program

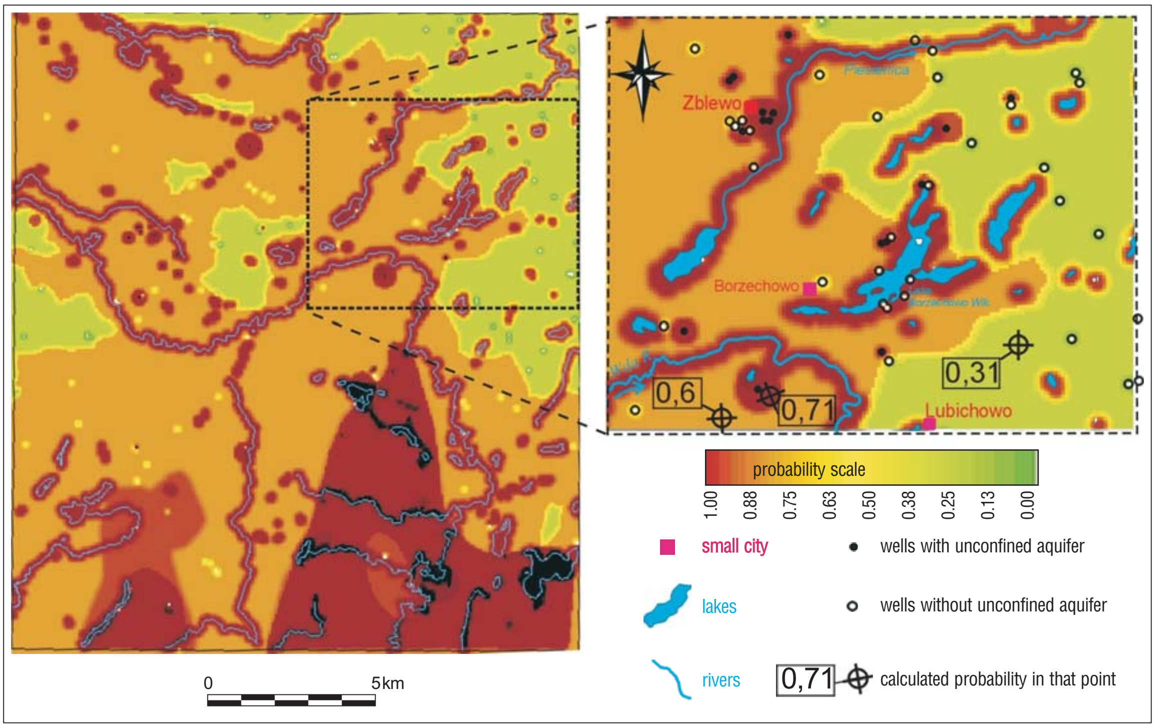


Fig. 7. The probabilistic map of the extent of the unconfined aquifer in a research area

After processing in the BELIEF module a set of maps was generated. These were maps of the degree to which evidence provided direct support for the hypothesis (belief) and the degree to which that evidence did not disprove the hypothesis (plausibility).

Summary and conclusions

The limits of geological and hydrogeological units (structures) presented in cartographic studies contain often significant errors due to poor geologic and hydrogeologic recognition. The purpose of the methodology presented here is to produce a probabilistic information layer of the extent of the unconfined aquifer in the study area (Fig. 7). It is an attempt to use Dempster-Shafer theory in hydrogeology. Taking into account the fuzzy set theory, the authors estimated the assessment of the extent of the hydrogeological unit based upon hydrogeological elements, especially boreholes and wells.

Additional information for the probabilistic map are derived from hydrological, geomorphological investigations and data from other geological reconnaissance points. The accuracy of such map is largely determined by the established membership functions.

The generated maps may be regarded as a supplement to a classic set of information concerning hydrogeology and which provides a new form of a map layer.

Acknowledgments

This research has been supported by founding provided by Nicolas Copernicus University in Toruń (UMK grant no. 378-G). The authors express their appreciation to A. Wróbel and L. Razowska for their help in translation and comments.

References

BRUNDSON C. 1995 — Estimating probability surfaces for geographical point data: an adoptive kernel algorithm. Computer and Geosciences, 21: 877–894.

BUTRYMOWICZ N., MURAWSKI T. & PASIERBSKI M. 1978 — Mapa geologiczna Polski w skali 1 : 200 000 — Arkusz Chojnice. Wyd. Geol., Warszawa.

EASTMAN J.R. 1999a — IDRISI guide to GIS Image Processing, Clark University, vol. 2: 41.

EASTMAN J.R. 1999b — IDRISI guide to GIS Image Processing, Clark University, vol. 1: 126.

FISHER P.F. 1999 — Models of uncertainty in spatial data. [In:] Longley P.A., Goodchild M.F., Maguire D.J. & Rhind D. (ed.), Geographical Information Systems: Principles and Technical Issues. John Wiley and Sons, New York, p. 191–205.

HÖLTING B. 1996 — Hydrogeologie, 3rd ed. Enke Verlag, Stuttgart.

KLIR G.J. & YUAN B. 1995 — Fuzzy sets and fuzzy logic: theory and applications. [In:] Malczewski J. 1999, GIS and multicriteria decision analysis. John Wiley and Sons, New York, p. 129.

LEE N.S., GRIZE Y.L. & DEHNAD K. 1987 — Quantitative Models for Reasoning Under Uncertainty in Knowledge-Based Expert Systems. International Journal of Intelligent Systems, no. 2: 15–38.

LEUNG Y. & LEUNG K.S. 1993 — An intelligent expert system shell for knowledge-based geographical information system. International Journal of Geographical Information Systems, 7: 189–213.

MALCZEWSKI J. 1999 — GIS and multicriteria decision analysis. John Wiley and Sons, New York, p. 129.

PACZYŃSKI B., RAZOWSKA L. & SADURSKI A. 1999 — Aquifers risk assessment by considering their vulnerability presented on the Hydrogeological map of Poland. [In:] M. Fendeková & M. Fendek (ed.), Proc. XXIX IAH Congress, Bratislava, p. 93–99.

PAZDRO Z. & KOZERSKI B. 1990 — Hydrogeologia ogólna. Wyd. Geol., Warszawa.

SHAFER G. 1976 — Mathematical theory of evidence. Princeton NJ: Princeton University Press.

ZADEH L. 1965 — Fuzzy sets. Information and Control, 8: 338–353.