Zamieszczone w Roczniku *Geoinformatica Polonica* artykuły zostały pozytywnie zaopiniowane przez Recenzentów wyznaczonych przez Komitet Redakcyjny Rocznika

Adres Redakcji

Akademia Górniczo-Hutnicza im. Stanisława Staszica
Wydział Geologii, Geofizyki i Ochrony Środowiska
30-059 Kraków, al. Mickiewicza 30
tel. (012) 617-23-68, fax. (012) 633-29-36
e-mail: lesniak@agh.edu.pl

Adres Wydawnictwa

Polska Akademia Umiejętności – Wydawnictwo
31-016 Kraków, ul. Sławkowska 17
tel. (012) 424-02-12
e-mail: wydawnictwo@pau.krakow.pl

Zamówienia przyjmuje i realizuje (łącznie z wysyłką) Wydawnictwo

POLSKA AKADEMIA UMIEJĘTNOŚCI
PRACE KOMISJI GEOINFORMATYKI

GEOINFORMATICA POLONICA

9

KRAKÓW 2009

Redaktor tomu
Jarosław Brzoskowski

Skład, łamanie
Anna Atanaziewicz

ISSN: 1642-2511

Druk i oprawa:
Oficyna Wydawniczo-Drukarska
„Secesja"
30-363 Kraków, ul. Rzemieślnicza 22

Objętość: ark. wyd. 9,00, ark. druk. 16,25, nakład: 300 egz.

SPIS TREŚCI

Artykuły

CONTENTS

Papers

PIOTR WERNER[1]

APPLICATION OF CELLULAR AUTOMATA AND MAP ALGEBRA IN STUDIES OF LAND USE CHANGES THE NEIGHBORHOOD COEFFICIENTS METHOD®[2]

Motto: Everything is related to everything else only near things are more related than distant things.
(W. R. Tobler, 1970)

Key words:

Land use, land cover, cellular automata, map algebra, neighborhood coefficient, simulation, urbanization, Warsaw

Abstract

Cellular automata are mathematical models for complex systems containing large numbers of simple identical components with local interactions.

The aim of the research is to verify the hypothesis stating that existing spatial pattern and the range of area with different classes of land use show defined tendencies for further spatial development according to the geographical model. The idea of using of the neighborhood indicator in analysis of land use changes is combination the method of map algebra and two-dimensional cellular automata. The aims of research are formulating the theoretical structure of the neighborhood indicator and checking its usability in practice.

During the process of simulation, socio-economic and physico-geographic factors are not taken into consideration. According to Foteringham (et al., 2000), one of the purposes of cartographical analysis is indirect understanding of the phenomena by analyzing spatial patterns and inferring spatio-temporal processes recognizing local hot-spots.

On the basis of a statistical analysis of the neighborhood indicators the transition rules for cellular automata were formulated to the purpose of the simulation of the state of the extent of these areas in the future.

The sets of the transition rules derived from the calculated neighborhood coefficients will be the base for simulations. The processes consist not only of computed transition rules but of all the neighborhood coefficients and simulation steps which are also (geo)visualized and statistically verified for the past.

[1] University of Warsaw, Faculty of Geography and Regional Studies, Warsaw, Poland, e-mail: peter@uw.edu.pl.

[2] Written within the framework of the research project: The Research Method and Forecast of Land Use Changes using Map Algebra and Cellular Automata.

WYKORZYSTANIE AUTOMATÓW KOMÓRKOWYCH I ALGEBRY MAP W BADANIACH ZMIAN UŻYTKOWANIA ZIEMI. METODA WSPÓŁCZYNNIKÓW SĄSIEDZTWA

Słowa kluczowe:

Użytkowanie ziemi, pokrycie terenu, automaty komórkowe, algebra map, współczynnik sąsiedztwa, symulacja, urbanizacja, Warszawa

Abstract

Obserwowana zależność dotycząca sposobu użytkowania ziemi fragmentów badanych obszarów wiąże je z rodzajem pokrycia i użytkowania terenów sąsiadujących; tzn. to w jaki sposób wykorzystywana jest gospodarczo obserwowana (badana) działka (nieruchomość) ma zdecydowanie mniejsze znaczenie dla przyszłego jej wykorzystania niż istniejące w sąsiedztwie sposoby użytkowania terenu. Sytuacja, sąsiedztwo istniejących i przeważających w danym regionie typów użytkowania fragmentów terenu (działek) jest w pewnej mierze determinantą dalszych zmian użytkowania.

W artykule przedstawiono koncepcję wykorzystania wskaźnika sąsiedztwa w analizie zmian użytkowania ziemi łączącą w sobie metodę algebry map oraz dwuwymiarowych automatów komórkowych. Przedstawione rozważania mają na celu sformułowanie konstrukcji teoretycznej wskaźnika sąsiedztwa oraz sprawdzenie jego przydatności w praktyce. Do tego celu wykorzystano mapy zasięgu obszaru zurbanizowanego aglomeracji Warszawy w latach 1969-1996. Na podstawie analizy statystycznej wskaźników sąsiedztwa sformułowano reguły przejść w celu symulacji stanu zasięgu obszaru zurbanizowanego aglomeracji Warszawy w 2023 roku.

Prezentowany model jest pierwszym autorskim opracowaniem metodyki konstrukcji wskaźników sąsiedztwa w oparciu o automaty komórkowe i próbą zastosowania funkcji zogniskowanych algebry map do symulacji zmian w modelu binarnym – prezentującym tylko dwa stany użytkowania ziemi, w celu weryfikacji i operacjonalizacji metody.

Summary

The idea of using of neighborhood coefficients in the analysis of land use changes is a combination of map algebra and two-dimensional cellular automata. The aim of the study and presented considerations is the theoretical construction of the neighborhood coefficients as well as checking their usefulness in practice. The maps of the extent of urbanized areas of Warsaw 1969 and 1996 were used as the research area to obtain this task. Next, the set of transition rules were constructed, based on the statistical analysis of neighborhood coefficients. They were applied to the process of simulation of the extent of the probable urbanized area of Warsaw in 2023.

Land use is defined as spatial distribution of forms of land cover patches, utilized or not by man within framework of the spatial and mutual relationships. The term refers to the functional character of a given terrain and is identified also with the socio-economic description of the surface (Ciołkosz & Bielecka, 2005, after Stamp, 1960, Kostrowicki, 1959).

The observed dependency of land use and neighboring land cover patches is pointed out in many publications from geography or spatial economy; i.e. the economic utilization of observed lot (or patch) has not as for reacking consequences for future as the existing land use in its neighborhood. Either the situation, existing neighborhood or predominant land use types in a given whole region are the factors stimulating of further land use changes of the observed lot (Hagoort, 2006).

The majority of land use models come from geography. They usually concern modeling the development of urbanization. But not only. The generalization of the results of the research investigations are the conclusions concerning the stability of land use. The majority of observations proved

the consolidation of the given land use forms, but there were local stability phenomena or geographical and historical inertia processes dominating over the existing forms of urbanization (Hagoort 2006, after Wilder, 1985, Bourne, 1965, Lee, 1974, Batty, 1996). The collision of different types of land use with existing ones in neighborhood can be the modifying factor. Other factors which can be taken into account are spatial and economic externalities. It was observed that the existing patterns of lot and neighborhood land use consolidate through the succession.

The state of art: research methods and models of land use changes

The wide set of tools and methods has been used to studies of land use changes. Between others the statistical and econometrics models were used (determined usually by the domain of the scientist). Geographers often used spatial interactions models, spatial planners – models of (non)linear programming and economists either utility maximization models or often multi-criteria and multipurpose decision support systems. The research field of land use changes is interdisciplinary and forces the application of integrated tools based on simulations; The input – output models (Hagoort, 2006) as well as fuzzy logic (Bilozor, 2003) were also used. GIS software obviously has became the tool of simulations and geovisualizations of the models' results.

The two categories of factors influencing land use changes (both within urban and rural areas) were confronted. The core of the studies was the evaluation of spatial changes arising from affecting neighborhood and succession for each form of land use for a given area. The definition of basic spatial unit was an inseparable and indispensable element of the studies. This problem was connected with the scale of spatial research; either the real estates lots, administrative units or even the (ir)regular geometric grids of cells (e.g. squares, Voronoi's polygons) could be used. The spatial units were categorized according to the predefined typology of land use classes.

Obviously the number and the types of classes of land use differentiated dependent on the aim, the scale and the area of geographical study.

One kind of these studies is CORINE Land Cover (CLC) inventory database worked out in the framework of the EEA – European Environment Agency project. According to the metadata file, all features on the map were digitized from an interpretation of Landsat TM satellite image printouts to scale 1:100000. The positional accuracy (according to CLC specifications) was 100 m and 25 ha minimum mapping unit. The format of map files is BIL – band interleaved by line. The longitude that defines the center of a projection was 09°00'00"E and the latitude that defines the center of a projection – 48°00'00"N. The reference system is Lambert Azimuthal. The scale of digital maps is defined as 1:25000. The legend of maps contain 50 (numbered) classes of land use (Ciołkosz & Bielecka, 2005).

The dynamics of changes of land use forces monitoring this process, especially within transitional zones between cities and villages, but also in such cases applies to the described above modeling tools and spatial units. The practice of studies reduces into cartographical analysis of maps, the satellite or aerial images for a given configuration of spatial units.

Another aspect of the research of land use changes takes into account spatial externalities (Hagoort, 2006). They are defined as unexpected indirect results of leverage of men causing changes of costs or the values for other ones – without visible operations (Mishan, 1969). Spatial externalities of land use changes disclose themselves usually throughout intensity (rate) and spatial extent. Both negative and positive spatial externalities can be distinguished. Higher spatial externalities related to given spatial land use class (form) implicit higher probability of the change of land use class for a spatial unit (Tobler, 1979). Summarized influence of land cover spatial externalities (net externalities) can be analyzed using multi-variables models, spatial or conditional spatial models of land use changes (according to Tobler – described detailed further). On the other hand the land use map (land cover map) does not contain important information as, e.g. environment protected areas, coverage of spatial organizations plans or urban antique areas which are also factors influencing the way and rate of land use changes. The are numerous examples taking into consideration individual geographical regions.

Models of land use changes

The main changes of land use were recognized and the set of models was distinguished (Hagoort 2006, after Tobler 1979):

1. Independent model of changes: each land use form is not related to the previous one in the same place; this is a random variation;
2. Dependent model of changes: each land use form is strictly related to the previous one in the same place; future land use depends on the present situation in the same place;
3. Historical model of changes: newly appearing land use forms depend directly on at least two chronological previous ones in the same place;
4. Multivariate model of changes: when the future land use depends on many different features, recognized at present in this area;
5. Geographical model – spatial model of changes – the future land use in the place depends on the neighborhood (at present).
6. Future geographical model – conditional spatial model – the future land use in the place depends on planned land use forms in the neighborhood.

The regions of urbanization are the most interesting for geographers. The studies of urban land use changes are the larger part of numerous research projects. Their conclusions were different because the studies concerned specific and different geographical locations. However some common findings were recognized. The different elements of landscape change with different speed (Whitehand, 1967) but some of them are more susceptible to changes and additionally the different rates of land use changes take place relative to cities centers (Wilder, 1985). Different temporary profiles of urban land use changes became longer than the scale of man's life and even longer than the perspective of spatial planning (Wegener, 1995). On the other hand, the courses of main streets are usually the most durable elements (Whitehand, 1967) and some types of land use forms are not subject to changes or the successions do not take place (Wilder, 1985).

Cellular automata as the research tool of land use changes

The changes of land use can be treated as a complex and (at least partly) random process. The complexity of interactions implicates that conventional deductive models are constrained especially in the case of spatio-temporal phenomena (Fulong Wu, 1999). Probabilistic models are also effectively applied. They make use of concepts of random development of reality, and the notions of statistical events, probability, random variables and empirical distributions of probabilities (Ratajczak, 1999). According to Foteringham (et al., 2000), one of the purposes of spatial and visual analysis of maps is indirect understanding of the phenomena by analyzing spatial patterns and inferring about spatio-temporal processes. Wolfram (2002:299) stated three possible mechanisms that can be responsible for complexity and randomness: "(…) there is random input from the environment at every step (…) there is random input only in initial conditions. And in the third case, there is effectively no random input at all (…)" The first mechanism is captured in the so-called stochastic models. The second is suggested by chaos theory. The third mechanism is suggested by the behavior of simple programs (op. cit.). This approach is related to use of cellular automata in research of land use changes (Werner, 2005).

One can observe a more and more frequent application of cellular automata to geographical research (Takeyama, Couclelis, 1997). "Cellular automata may be considered as (parallel-processing) computers, in which the initial configuration encodes the program and input data, and time evolution yields the final output"[3] (Wolfram, 1982). "Cellular automata are mathematical models for complex natural systems containing large numbers of simple identical components with local interactions. They consist of a lattice of sites, each with a finite set of possible values. The value of the sites evolves synchronously in discrete time steps according to identical rules. The value of a particular site is determined by the previous values of a neighborhood of sites around it."[3]

[3] http://www.brunel.ac.uk/depts/AI/alife/al-wolf2.htm

The cellular automaton is represented as follows[4]:

$$A \equiv (\alpha, S, N, f)$$

where:

α – *regular grid comprised of cells*
S – *state of cell*
N – *neighborhood*
f – *transition rule of automaton*

Usually the anamorphic plane (the map) covered with the grid of cells is considered as the base of the study of land use changes with cellular automata. The ranges of different land cover forms should be disjoint but they should cover the whole area (without undefined regions). It means that each cell should have only one state and no one cell in the grid should be undefined. This is the way of quantification of the space. There appears the model of two dimensional cellular automaton with appropriate, countable states. Usually one the two types of neighborhood is defined: von Neumann or Moore. The tenet depends on the connections of edges and vertices.

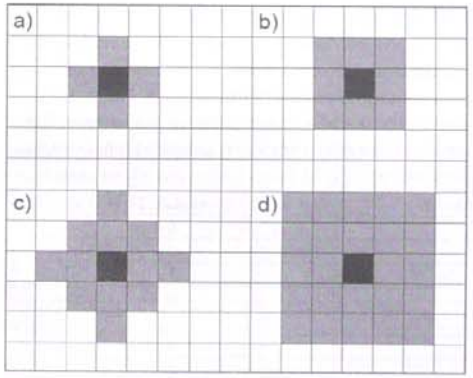


Fig. 1. Different neighborhoods of cellular automaton (a) von Neumann, (b) Moore, (c) von Neumann-Moore, (d) extended Moore
Rys. 1. Otoczenia automatów komórkowych: (a) von Neumann'a, (b) Moore'a, (c) von Neumann-Moore'a, (d) rozszerzone Moore'a

[4] The theory of cellular automata was by John von Neumann with the help of Polish mathematician Stanisław Ulam.

The transition rule defines the evolution of cellular automaton: the state of every cell in the centre changes in every step (the time t + 1) depending on its initial state as well as on the initial state of cells in the neighborhood (in time t). The transition rules can be described by an algorithm, with a table or with a set of established rules. Also shore conditions before the experiment (simulation) are important.

The initial conditions of the study of land use changes with cellular automata are pre-defined by the existing land cover.

Stephen Wolfram classified all 256 one-dimensional, two-states cellular automata, distinguishing four main classes of their behavior (2002:231):

- Class 1: the behavior is simple, and almost all initial conditions lead to exactly the same uniform final state;
- Class 2: there are many different possible final states, but all of them consist just of a certain set of simple structures that either remain the same for ever or repeat every few steps;
- Class 3: the behavior is more complicated, and seems in many respects random, although triangles and other small-scale structures are generally seen at some level;
- Class 4: involves a mixture of order and randomness: localized structures are produced which on their own are fairy simple, but these structures move around and interact with each other in very complicated ways.

Additionally, one should distinguish between deterministic and probabilistic cellular automata. The latter are based on the averages of neighborhood cells, which in turn define the probability of the final states.

The notations of cellular automata

The attempts to introduce the notations of cellular automata involving a description of transitions rules did not succeed, apart from the notation by Wolfram in 1983 for one-dimensional, two-states ones. Wolfram notation as follows: the numbers describe all possible configurations of the states of a central cell and neighborhood ones (two states, three cells: $2^3 = 8$ numbers). For each configuration the value is interpreted as a binary number.

The two-dimensional, two-states cellular automaton is sometimes described with letter H (fr. *huite* – eight ones around the center). For example three cells with ones are described as H3 (for Moore neighborhood).

The neighborhood coefficient is proposed as the notation for two-dimensional cellular automata either for two – or more states. The coefficient in fact is the unique number and each fits for a definite configuration of call values. It can be applied to any neighborhood and defined (not numerous) number of states of cells. The neighborhood coefficient can accept very large values, growing exponentially near larger number of states and the larger number of cells. For a given Moore neighborhood, two-dimensional, two-states cellular automata, the algorithm of neighborhood coefficient is the following:

Table 1. Symbolic, decimal and graphical (geographical) notation
Tabela 1. Zapis symboliczny, dziesiętny i graficzny konfiguracji komórek w otoczeniu Moore'a

Symbolic notation

1	1	0
1	1	0
0	0	0

traditional H3, proposed H^{54}

$1*2^2=5$	$1*2^1=0$	$0*2^0=1$
$1*2^5=32$	$1*2^4=16$	$0*2^3=0$
$0*2^8=0$	$0*2^7=0$	$0*2^6=0$

Decimal notation (54)
Binary notation 000110110

Graphical & geographical notation

NW	N	NE
W	C	E
SW	S	SE

{C1,N1,NE0,E0,SE0,S0,SW0,W1,NW1}

The proposed notations for the configuration of value of the cells in table 1 are as follows:

– Symbolic notation H_2^{54}

– Binary notation $H_2^{000110110}$

– Text notation {C1, N1, NE0, E0, SE0, S0, SW0, W, NW1}

This simple notation allow to tie, mutually, with one-to-one relation, the configuration of values of the cells and its description (symbolic notation). The algorithm enabling its transformation into a binary record and next into a graphic notation is break-down number 54 on factors, dividing modulo 2. The letter H defines that there is the Moore neighborhood, and the binary record is stable, i.e. the positions of binary digits are always occupied by the same cells. The rests of division (the division of modulo with the number of states) define the spatial distribution of states of individual cells (see table).

Cell		dividend (divider = 2)	modulo 2
NE	$\rightarrow 0*2^0$	54	0
N	$\rightarrow 1*2^1$	27	1
NW	$\rightarrow 1*2^2$	13	1
E	$\rightarrow 0*2^3$	6	0
C	$\rightarrow 1*2^4$	3	1
W	$\rightarrow 1*2^5$	1	1
SE	$\rightarrow 0*2^6$	0	0
S	$\rightarrow 0*2^7$	0	0
SW	$\rightarrow 0*2^8$	0	0
Binary notation from the rests			000110110
checking		$0*2^8+0*2^7+0*2^6+1*2^5+1*2^4+0*2^3+1*2^2+1*2^1+0*2^0=54$	

This algorithm can be proved with mathematical induction for any (but limited) neighborhood and any (also limited) states but the author leaves it for mathematicians and only some more examples will be presented further.

In case of two-dimensional, two-state cellular automata for Moore neighborhood the finished number of configurations of values is equal $\overline{W}_2^9 = 2^9 = 512$. For the neighborhood coefficient H_2^{488} the appropriate text notation

is the following {C0, N0, NE0, E1, SE1, S1, SW1, W1, NW0} and the related graphical notation:

NW	N	NE
W	C	E
SW	S	SE

=

0	0	0
1	0	1
1	1	1

The $H_2^{113} \rightarrow H_2^{488}$ transition rule $f: S^N \rightarrow S$ can be denoted as the following , for the passage of the configuration 001110001 into 111101000.

For larger number of states – the number of possible neighborhood coefficients (Moore neighborhood) grows. Taking six-states, two-dimensional cellular automata for Moore neighborhood – finished number of all configurations carries out $\overline{W}_2^9 = 6^9 = 10\,077\,696$. For all values of cells equal to 0 the neighborhood coefficient takes H_6^0 ($H_6^{000000000}$). For all cells equal to value 5 – $H_6^{10077695}$ ($H_6^{555555555}$). The algorithm is presented in the table below.

Cell		Dividend (divider = 6)	Floor result	Rest (modulo 6)
NE	$\rightarrow 5*6^0$	10077695	1679615	5
N	$\rightarrow 5*6^1$	1679615	279935	5
NW	$\rightarrow 5*6^2$	279935	46655	5
E	$\rightarrow 5*6^3$	46655	7775	5
C	$\rightarrow 5*6^4$	7775	1295	5
W	$\rightarrow 5*6^5$	1295	215	5
SE	$\rightarrow 5*6^6$	215	35	5
S	$\rightarrow 5*6^7$	35	5	5
SW	$\rightarrow 5*6^8$	5	0	5
			Hexagonal notation	555555555
checking		$5*6^8+5*6^7+5*6^6+5*6^5+5*6^4+5*6^3+5*6^2+5*6^1+5*6^0=10077695$		

Next examples describes 16 – state, two-dimensional cellular automaton. Maximum value, when all cells for Moore neighborhood are equal 15 (F – hexadecimal) carry out 6,87 E + 10, but the number of possible configurations one more. If random number from the range is 24852515340, then the reconstruction of the configuration of values of the cells is the following:

Cell		Dividend (divider =16)	Floor result	Rest (modulo 16)
NE	$\rightarrow 12*16^0$	24852515340	1553282208	12 (C)
N	$\rightarrow 0*16^1$	1553282208	97080138	0 (0)
NW	$\rightarrow 10*16^2$	97080138	6067508	10 (A)
E	$\rightarrow 4*16^3$	6067508	379219	4 (4)
C	$\rightarrow 3*16^4$	379219	23701	3 (3)
W	$\rightarrow 5*16^5$	23701	1481	5 (5)
SE	$\rightarrow 9*16^6$	1481	92	9 (9)
S	$\rightarrow 2*16^7$	92	5	12 (C)
SW	$\rightarrow 5*16^8$	5	0	5 (5)
			Hexadecimal notation	5C9534A0C
checking		$5*16^8+12*16^7+9*16^6+5*16^5+3*16^4+4*16^3+10*16^2+0*16^1+12*16^0=24852515340$		

And the graphical notation for Moore neighborhood:

$H_{16}^{24852515340}$

A	0	C
5	3	4
5	C	9

For (2D50S) – 50-states, 2– dimensional cellular automaton the reconstruction of neighborhood (Moore) coefficient equals $H_6^{1,61E+14}$ ($H_6^{46400000}$) and is the following:

Dividend (divider =50)	Floor result	Rest (modulo 50)
1,61E+14	3,22E+12	0
3,22E+12	6,44E+10	0
6,44E+10	1,29E+09	0
1,29E+09	2,58E+07	0
2,58E+07	5,15E+05	0
5,15E+05	1,03E+04	0
1,03E+04	2,06E+02	4
2,06E+02	4,00E+00	6
4,00E+00	0,00E+00	4
	notation	46400000
checking	$4*50^8+6*50^7+4*50^6+0*50^5+0*50^4+0*50^3+0*50^2+0*50^1+0*50^0=1,61E+14$	

And graphical notation:

0	0	0
0	0	0
4	6	4

Geovisualization of land cover using cellular automata and map algebra

The cellular automata are similar to map algebra models in GISc. There have been numerous approaches developed to create a of uniform methodology joining both ideas (Takeyma & Couclelis, 1997). Map algebra is the scheme of cartographical modeling. It describes all the possible transformations of raster maps assembling them in four basic classes. They are the following (Takeyma & Couclelis 1997, Longley et al. 2006, after Tomlin, 1983):

• local operations examine rasters cell by cell, comparing the value in each cell in one layer with the values in the same cell in other layers;

• focal operations compare the value in each cell with the values in its neighboring cells – most often eight neighbors;

• global operations produce results that are true of the entire layer, such as its mean value;

• zonal operations compute results for blocks of contiguous cells that share the same value, such as the calculation of shape for contiguous areas of the same land use, and attach their results to all of the cells in each contiguous block.

The assumption of the analysis of land use is that there is one-to-one relation between land cover class description and its unique number.

Cellular automata could be perceived as an element of map algebra. The neighborhood coefficients described above concern focal operations. Yet not only. Cellular automata transform layers or zones into quite new maps – neighborhood coefficients maps. Their statistical analysis, comparison of two situations for different years lets us define main tendencies and carry out a simulation of the future land use (see figure):

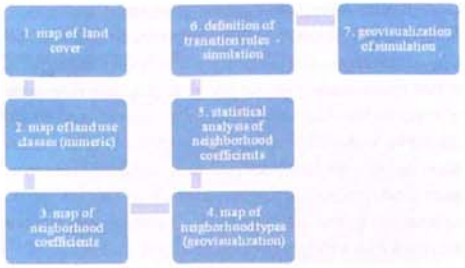


Fig. 2. Steps in analysis of neighborhood
Rys. 2. Kroki w analizie sąsiedztwa

Theoretical example of land cover analysis

For a given theoretical representation of land cover map with six classes described as 0,1,2,3,4 and 5 the number of all possible configurations in Moore neighborhood carries out 6^9=10077696. Taking into account the shore conditions it is assumed that all cells out of border will be equal to 0.

0	0	0	0	0	0	0	0	0	0	0	0
0	0	3	4	2	4	4	5	1	1	0	0
0	3	3	4	2	4	5	2	1	0	1	0
0	5	3	4	2	4	2	4	0	5	1	0
0	2	5	0	5	5	5	5	2	0	1	0
0	4	4	4	4	5	5	5	4	4	1	0
0	4	4	4	4	5	5	5	5	5	1	0
0	3	4	4	4	3	5	5	5	5	1	0
0	3	1	1	5	5	5	5	5	5	1	0
0	3	2	2	2	2	2	2	4	4	1	0
0	3	2	2	2	2	2	0	3	3	1	0
0	0	0	0	0	0	0	0	0	0	0	0

Fig. 3. Random generated map of land cover. The concrete classes of land cover are described with numbers between 0 and 5 (see description in text)
Rys. 3. Losowo wygenerowana mapa pokrycia terenu. Poszczególnym klasom pokrycia terenu przypisano numery od 0 do 5 (por. opis w tekście)

The counted neighborhood coefficients can also be presented on the map:

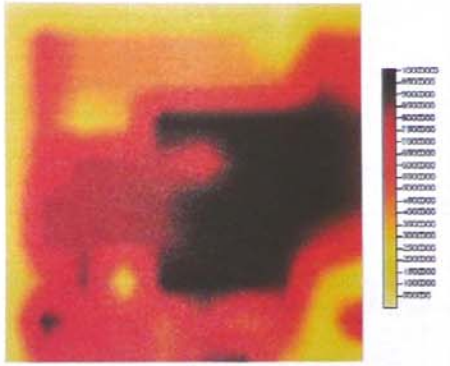


Fig. 4. The map of neighborhood coefficients
Rys. 4. Mapa numeryczna współczynników sąsiedztwa

The statistical analysis of histogram proved that the map is multimodal. There are two modes. These are coefficients H_6^{18662} and $H_6^{10077695}$. The median is equal to $H_6^{6175440}$ and the average to $H_6^{5600742}$. There are the following configurations of Moore neighborhood:

Table 2. Modes, median and average of the neighborhood coefficients map
Tabela 2. Wartości mody, mediany i średniej dla mapy wskaźników sąsiedztwa

2	2	2	H_6^{18662} – mode			5	5	5	$H_6^{10077695}$ – mode		
2	2	2				5	5	5			
0	0	0				5	5	5			
0	0	0	$H_6^{6175440}$ – median			2	1	0	$H_6^{5600742}$ – average		
2	1	0				0	1	3			
3	4	0				3	2	0			

Frequency of particular neighborhood coefficients are presented in the table below:

Table 3. Frequency of neighborhood coefficients
Tabela 3. Częstość występowania wskaźników sąsiedztwa

Neighborhood coefficients	Number	%	Cumulative %
below 18662	4	4	4
equal 18662	3	3	7
between 24774 and 10077694	90	90	97
equal 10077695	3	3	100
total	100	100	x

Empirical case study: analysis of land use changes for 2D&2S cellular automata

The statistical analysis of land use changes and the definition of the successions rules can be done for given two maps of neighborhood for different years. In fact, if the order of appearance of neighborhood coefficients is defined this means that the transition rules of cellular automaton are defined. These are focal and local operations according to map algebra.

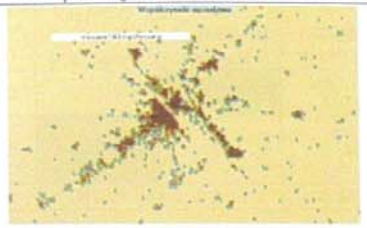
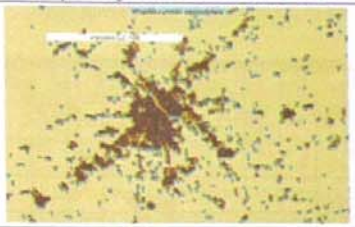
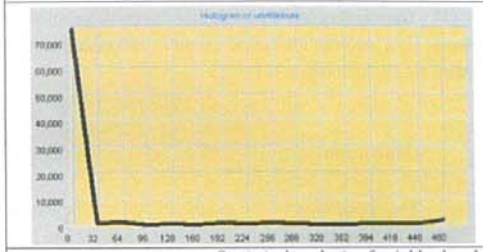
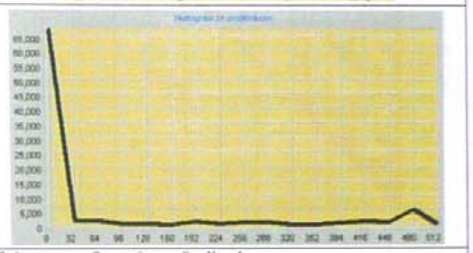
The starting point of analysis are maps of the extent of the urbanized area of Warsaw in 1969 and 1996. The maps can be treated as a model of 2-dimensional and 2-states cellular automaton.

The first analysis made earher (Werner, 2005) used a probabilistic model and the so called outer totalistic rule. The state of the central cell was evaluated by applying the state of surrounding ones and multiplied by appropriate weights taken from the cellular automaton rule. The cellular automaton is also normalized so that they add up to 1. In fact it was random search for an appropriate rule – one of 512 possible. The verification of the cellular automaton operations was done through the comparisons of the simulation results with the real extent. Such research gives sometimes incidental transitional rules, may be of relative worth, because one cannot definitely prove that there is no another transition rule which better maps the real situation.

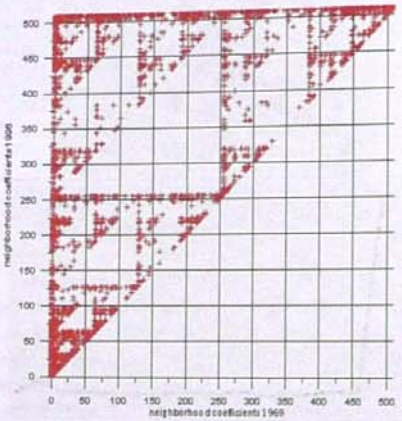
The proposed method is an application of neighborhood coefficients to the transition rules definition. In this case the study concerns only the deterministic cellular automata. There where two raster maps used in the simulation: starting (1969) and verification (1996) ones. Uhorczak's map (1969) of land use in Poland and CLC coverage was worked out in 1996. The maps present a spatial pattern using two states: 1 (black) for the urban and suburban areas, 0 (white) for the rest of cells.

The visual analysis of maps proves that the extent of urbanized areas is expanding. The statistical analysis of coefficients shows that the predominant neighborhood configuration in 1969 year consists of empty cells around a single central urbanized one within Moore neighborhood. If there are more urbanized cells they are present most frequently in SW cells. In 1996 the most frequent configurations are either for all cells marked as urbanized or just the opposite, till 1969, cells marked as urbanized are present elsewhere except of south-west part of Moore neighborhood.

Table 4. Statistical analysis, maps and histograms of neighborhood coefficients for the extent of Warsaw urbanized area in 1969 and 1996
Tabela 4. Opis statystyczny i mapy oraz histogramy wskaźników sąsiedztwa na podstawie map zasięgu obszarów urbanizacji aglomeracji Warszawy w 1969 i 1996 roku

1.Map of neighborhood coefficients – 1969	2. Map of neighborhood coefficients – 1996
	
	
Statistical analysis of neighborhood coefficients – configurations of cell values	
I quartile = 9	I quartile = 180
0 0 1 0 0 1 0 0 0	1 0 0 1 1 0 0 1 0
Median = 72	Median = 366
0 0 0 0 0 1 0 0 1	1 1 0 1 0 1 1 0 1
III quartile = 256	III quartile = 500
0 0 0 0 0 0 1 0 0	1 0 0 1 1 0 1 1 1
Mode = 0	Mode = 511
0 0 0 0 0 0 0 0 0	1 1 1 1 1 1 1 1 1
Rejecting the extreme values 0 and 511	
0 0 0 0 0 0 1 0 0 Mode = 256	1 1 1 1 1 1 0 1 1 Mode =255

Analyzing the statistical graph of correlations between the neighborhood coefficients in 1969 and 1996 one can observe visual similarity to a classic fractal (Sierpinski triangle). The aim of the visual graph analysis is the recognition of the occurrence of transition rules.



Then the transition rules were established using cross-tabulation. 75% of observations were pairs of coefficients occurrthet one or twice. Only 748 pairs were over 3rd quartile (excluding identical transitions 0→0, 1→1, ...). There was no transition to configuration with all empty cells (H_2^0). The observed transitions are always from lower to equal, or higher values of neighborhood coefficients. Within the 4th quartile there were excluded those with coefficients equal to 0 in 1969 (203 cases). They mean the that the urbanized areas appear in an empty space and the transition rule cannot be clearly established. Then 545 remained. Taking into account this set, a similar analysis was done. There 128 transitions of coefficients stayed at the end of analysis. Over half of them (69) were such pair transitions which ended with coefficients equal to 511 (covering all Moore neighborhood).

The effect was 79 rules got of unique transitions. The simulation for the period of 1969–1996 let us compare the consistence of map of the simulated and real range of urbanized area (from CLC map in 1996 – almost 93%).

Table 5. Simulated and real extent of urbanized areas
Tabela 5. Udział procentowy powierzchni obszarów zasięgu symulowanego i rzeczywistego obszaru zurbanizowanego aglomeracji warszawskiej dla 1996 roku

Share	Area of simulated range (%)	Area out of simulated range (%)	total %
Area of real range (%)	7.53%	3.84%	11.37%
Area out of real range (%)	3.22%	85.40%	88.63%
Total %	10.76%	89.24%	100.00%

The verification of recognized transition rules let made simulation of the future extent of urbanized area of Warsaw agglomeration till 2023.

Conclusion

The application of the neighborhood coefficients to land use analysis links the method of map algebra and two-dimensional cellular automata. Presented considerations aimed at the theoretical construction of neighborhood coefficient and its usability in simple empirical studies. The results are promising but the procedure should be verified for different geographic areas and simultaneously for more number of states than two – representing real land use classes.

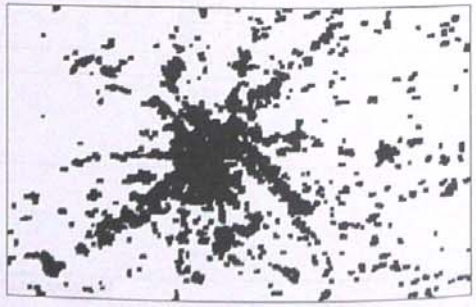


Fig. 5. Simulated extent of urbanized area of Warsaw agglomeration in 2023
Rys. 5. Symulowany zasięg obszaru zurbanizowanego aglomeracji Warszawy w 2023 roku.

Bibliography

Beim M., 2007, *Modelowanie procesu suburbanizacji w aglomeracji poznańskiej z wykorzystaniem sztucznych sieci neuronowych i automatów komórkowych*, rozprawa doktorska, Uniwersytet im. Adama Mickiewicza, Instytut Geografii Społeczno-Ekonomicznej i Gospodarki Przestrzennej, Poznań.

Biłozor A., 2003, *Zastosowanie logiki rozmytej do delimitacji strefy przejściowej miasta i wsi*, [w:] H. Rogacki, *Problemy interpretacji wyników metod badawczych stosowanych w geografii społeczno-ekonomicznej i gospodarce przestrzennej* (pp. 215–223), Poznań: Bogucki Wydawnictwo Naukowe.

CiołkoszA., & Bielecka E., 2005, *Pokrycie terenu w Polsce. Bazy danych CORINE Land Cover*. (wyd. Biblioteka Monitoringu Środowiska), Warszawa: Inspekcja Ochrony Środowiska.

Fotheringham A. S., Brundson C., & Charlton M., 2000, *Quantitative Geography. Perspectives on Spatial Data Analysis*, London: Sage Publications.

Fulong W., 1999, *GIS-based simulation as an exploratory analysis for space-time processes*, Journal of Geographical Systems, vol. 1., 199–218.

Hagoort M. J., 2006, *The Neigbourhood Rules. Land use interactions, urban dynamics & cellular automata* (vol.(3)34), Utrecht: Netherlands Geographical Studies.

Longley P. A., Goodchild M. L., Maguire D. J., & Rhind D. W., 2006, *GIS. Teoria i praktyka*, Warszawa: PWN.

Mishan E. J., 1969, *Welfare economics: Ten introductory essays*. New York.

Ratajczak W., 1999, *Modelowanie sieci transportowych*, Poznań: Wydawnictwo Naukowe UAM.

Takeyma M., & Couclelis H., 1997, *Map dynamics integrating cellular automata and GIS through Geo-Algebra*, Int. J. Geographical Information Science, 11 (1), pp. 73–91.

Tobler W. R., 1979, *Cellular Geography*, [in:] S. Gale, & G. Olsson, Philosophy in Geography (s. 379–386), Dordrecht, The Netherlands: Reidel.

Wegener M., 1995, *Current and Future Land Use Models. Travel Model Improvment Program Land Use Model*. Dallas, TX.: U.S. Department of Transportation.

Werner P., 2005, *Symulacja zmian zasięgu obszaru zurbanizowanego aglomeracji Warszawy. Eksperyment zastosowania automatów komórkowych*, [w:] T. Czyż, & H. Rogacki, *Współczesne problemy i koncepcje teoretyczne badań przestrzenno-ekonomicznych*. (vol. 219, 212–219), Warszawa: Biuletyn KPZK PAN.

Whitehand J. W., 1967, *Fringe Belt: a neglected aspect of urban geography*, Transactions, Institute.

Wilder M. G., 1985, *Site and situation determinats of land use change: an empirical exampl*, Economic Geography, 61(4), 332–344.

Wolfram S., 2002, *A New Kind of Science*, Champaign: Wolfram Media Inc.

Streszczenie

W wielu publikacjach z zakresu geografii i gospodarki przestrzennej wskazuje się na obserwowaną zależność dotyczącą sposobu użytkowania ziemi fragmentów badanych obszarów od rodzajów pokrycia i użytkowania terenów sąsiadujących; tzn. to w jaki sposób wykorzystywana jest gospodarczo obserwowana (badana) działka (nieruchomość) ma zdecydowanie mniejsze znaczenie dla przyszłego jej wykorzystania niż istniejące w sąsiedztwie sposoby użytkowania terenu. Sytuacja, sąsiedztwo istniejących i przeważających w danym regionie typów użytkowania fragmentów terenu (działek) jest w pewnej mierze determinantą dalszych zmian użytkowania ziemi na danej działce.

Większość modeli zmian użytkowania terenu to modele stworzone na gruncie geografii. Duża część z nich dotyczy modelowania rozwoju urbanizacji. Ale nie tylko. Uogólnieniem rezultatów badań są wnioski dotyczące stabilności form użytkowania ziemi wydzieleń na danym terenie. W większości obserwuje się w dłuższym okresie utrwalenie danej formy użytkowania ziemi na badanych fragmentach terenu, a w przypadku urbanizacji pojawia się zjawisko stabilności lokalnej bądź występują procesy geograficzno-historycznej inercji, które mogą zdominować istniejące formy urbanizacji. Czynnikiem modyfikującym może być kolizja określonego typu użytkowania ziemi danego wydzielenia z istniejącym pokryciem terenu w sąsiedztwie oraz czynniki, które można określić mianem procesów zewnętrznych. Obserwuje się, że istniejące wzory użytkowania ziemi i otoczenia utrwalają się poprzez sukcesję.

W badaniach zmian użytkowania ziemi wykorzystywano szeroki wachlarz narzędzi i metod. Wraz z rozwojem oprogramowania systemów informacji geograficznej – stały się oczywistym narzędziem symulacji i wizualizacji wyników modelowania. W ostatnich latach coraz częściej stosuje się modele oparte na wykorzystaniu automatów komórkowych.

Inny aspekt badań zmian użytkowania ziemi to uwzględnienie przestrzennych efektów zewnętrznych – spatial externalities. Sumaryczne oddziaływanie efektów zewnętrznych pokrycia terenu (net externalities) można analizować przyjmując modele wielo-zmienne, przestrzenne lub warunkowe przestrzenne zmian klas użytkowania ziemi. Na podstawie analizy czasowo-przestrzennej zmian pokrycia terenu w badaniach rozpoznano szereg modelowych zmian użytkowania ziemi. Zmiany użytkowania ziemi można traktować jako złożony i częściowo losowy proces. W przypadku analizy kartograficznej (kartograficznej metody badań w geografii) jednym z celów jest zrozumienie zjawisk pośrednio, poprzez analizę rozmieszczenia (wzorów przestrzennych) z których można wnioskować o procesach czasowo-przestrzennych.

To ostatnie podejście ściśle wiąże się z wykorzystaniem automatów komórkowych do badań zmian użytkowania ziemi (Werner, 2005).

Obserwując dwuwymiarowe automaty komórkowe, zaproponowano oryginalną autorską metodę notacji automatów komórkowych, którą można nazwać wskaźnikiem sąsiedztwa. Ten zapis pozwala w bardzo prosty sposób wzajemnie jednoznacznie związać wartości i rozmieszczenie komórek w otoczeniu Moore'a oraz proponowany zapis. Automaty komórkowe mają wiele wspólnego z wypracowaną w geoinformatyce algebrą map.

Metodę zweryfikowano prezentując symulację zmian obszaru zurbanizowanego aglomeracji Warszawy w latach 1969–1996–2023.

GEOINFORMATICA POLONICA
9: 2009

EWA KRZYWICKA-BLUM, ADAM MICHALSKI[1]

ANAMORPHIC MAPS – AN OVERVIEW

Key words:

anamorphic maps and models, cartograms, map classification

Abstract

In the paper an attempt at the classification of mathematical foundations of maps according to practical functions has been presented. This classification refers to the choice of the metric and the principles of placing the elements of the content of maps: A – maps with a scale whose fluctuations (changeability) a user can determine analytically (Euclidean metric), B – variable-scale, anamorphic maps (elements placing depends on its density or importance), C – anamorphic models with a functional metric (operational scale). Since the 19th century all maps have had to be treated as variable-scale models of reality. Among the solutions known in historical development of cartography mostly two of three groups of proposed classification have been used. The new methodological proposition are maps in the third group: collecting anamorphic maps with an operational, functional scale. The examples of different problems that may be expressly shown using anamorphoses were presented. Comparing a new form of notation with cartodiagrams or choropleth maps (which were also called "cartograms" in the beginning of thematic cartography) one can observe that the proposed kind of modelling is better from the logical, semiotic and semantic point of view, first of all because the same visual variable "value" is not used for two different objects.

MAPY ANAMORFICZNE – PRZEGLĄD

Słowa kluczowe:

mapy i modele anamorficzne, klasyfikacja map

Abstrakt

W artykule zaprezentowano próbę klasyfikacji podstaw matematycznych map ze względu na ich funkcje praktyczne. Klasyfikacja ta odnosi się do wyboru metryki oraz zasad umiejscowienia elementów treści map i wygląda następująco: A – mapy opatrzone skalą, której wahania użytkownik może analitycznie wyznaczyć (metryka euklidesowa), B – mapy zmiennoskalowe anamorficzne (umiejscowienie elementów zależne od ich gęstości lub znaczenia), C – modele anamorficzne o metryce funkcjonalnej (skali operacyjnej). Od XIX wieku mapy można (dzięki teorii odwzorowań) traktować jako

[1] Institute of Geodesy and Geoinformatics, The Faculty of Environmental Engineering and Geodesy, Wrocław University of Environmental and Life Sciences, Wrocław, Poland.

zmiennoskalowe modele rzeczywistości. W procesie rozwoju kartografii głownie dwie grupy map spośród trzech były rozwijane i używane. Nowa propozycją metodologiczną są mapy trzeciej grupy: gromadzące mapy (modele) o metryce funkcjonalnej (skali operacyjnej). Wskazano na problemy wynikające z użytkowania anamorfoz. Porównanie nowego sposobu notacji z klasycznymi metodami jak kartogram i kartodiagram pozwala zauważyć, że zaproponowany sposób modelowania może być lepszy z punktu widzenia logicznego, semiotycznego i semantycznego, szczególnie dlatego, że zmienna wizualna (wartość) jest przedstawiana na powiązanym tematycznie tle.

Introduction

It seems to be necessary to mention that in the USA and the United Kingdom studied kind of maps are called "cartograms". In European countries (France, Poland, Russia) a term "anamorphic map" exists. In accodance to Kretschmer's lexicon of cartography (Kretschmer et al. 1986 after Tobler 2004) – "Kartogramm" is known as a choropleth or a statistical map, but also a "verzerre" (distorted) map. In this paper the authors deliberately used terms: "anamorphic map" or "anamorphic model", "anamorphose".

The crucial component of a map definition is determining the relation between the prototype of a map and its presentation in a notation. The features such as a reduced size scale, simplification and mathematical basis indicate how to show symbols on a map.

If a map is considered to be a model of spatial relations, the scope of map functions is established. A graphic code, viz. a combination of a few visual variables and a few cartographic methods as well as using the principles of a system geometrization, which derive from the Euclidean metric applied in the Mediterranean culture since 300BC, make it possible to reveal abstract characteristics, which are very important in human cognitive processes and activities. These characteristics include shapes and sizes of objects, location of objects belonging to the same and different categories, randomness and well-ordering of the whole arrangement of objects, isotropy, orientation, rhythm, etc.

A geometrized picture notation in which the Euclidean metric or any other metric or semi-metric chosen because of its relevance for a specific function of the model determines the way of showing symbols on the map is called an anamorphic map.

Because of the wide range of objects and different principles for assigning symbols to them, geometrized picture notations may be divided into three groups:

A – maps with a scale whose fluctuations (changeability) a user can determine analytically,

B – variable-scale, anamorphic maps

C – anamorphic models with a functional metric (operational scale)

All three groups of notation meet topological requirements of models. It is also assumed that all groups may be used according to the Euclidean geometry code.

Maps with a scale whose fluctuations a user can determine analytically

Constant-scale maps covering base maps and thematic maps constitute models of spatial relations understood in terms of the Euclidean metrics.

Base maps account for models of spatial distribution and object shapes depending on sizes of objects being presented. They are more or less in accordance with reality perceived from the Euclidean perspective. This group of maps covers a wide range of functions including education, science, planning, management, military study and social communication. Large-scale maps with constant scale became official documents in many fields of life. The example of such a map may be a map of property and political relations (cadastral maps, political maps).

Thematic maps can be distinguished by a selection hierarchy and presentation of the content, depending on whether they belong to the background or form the thematic substance. The projection of the background must ensure correctness of the spatial features of symbols which are essential for the description of thematic objects such as shape – for morphological elements, areas – for political division and land use.

Maps presenting a spatial distribution of a distinct quantitative attribute of a real or an abstract object which

is linked to points, lines and areas on a map only by relation, not by meaning, have played an important role in thematic cartography since the 19th century. It means that a map user is not interested in the area of a given region but a population of people who lives there or an unemployment rate which is a percentage share of the number of unemployed persons in a total number of economically active people. The existence of two scale dividers for such maps makes it difficult to perceive information that is crucial to the user. What is even a greater limitation in the perception process is the fact that this type of modelling does not meet requirements of logical conformity. The graphic variable "quantity" refers not only to two different functional layers (background and thematic substance), but also to two different objects (area of a region and area as the number of citizens).

Variable-scale anamorphic maps

The essence of variable-scale maps is their local resemblance to constant-scale maps. It applies to the presentation of the content which is characteristic of constant-scale maps. Scale modulation is aimed at:

a – adjustment of the graphic weight of notation, when the density of objects in a given area is much diversified,

b – highlighting certain areas, which is explicitly adopted to operational functions.

An example of a variable-scale anamorphic map is a geodetic field sketch. The lower density of terrain objects corresponds to a relatively shorter survey line. Analogically, the survey line is longer for higher density of objects.

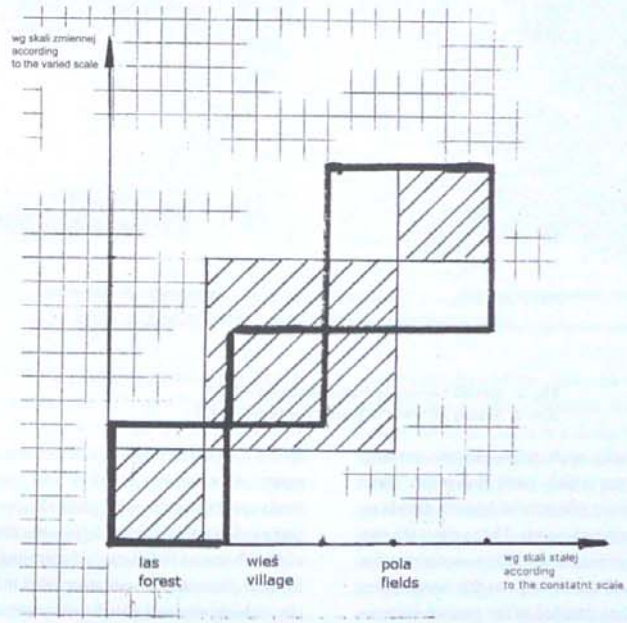


Fig. 1. Relation between the scale and the density of elements
Rys. 1. Relacja między skalą a gęstością elementów

When drawing up a constant-scale map, the scale must be adjusted. This means that symbol sizes and distances between symbols must be amended (Boutoura 1993, 1994). A demonstrative example was shown in figure 1 – geometric figures drawn with a wide black line represent "forest, village and fields" in the original size, however patterned figures represent transformation conditioned density of map elements. An opposite process depends on the adjustment of the graphical weight of pictorial notation, which is clearly unbalanced on the constant-scale map (Fig. 2). As a result, the user has a variable-scale map at his disposal, which is used, however, as a constant-scale, multi-functional model of spatial relations.

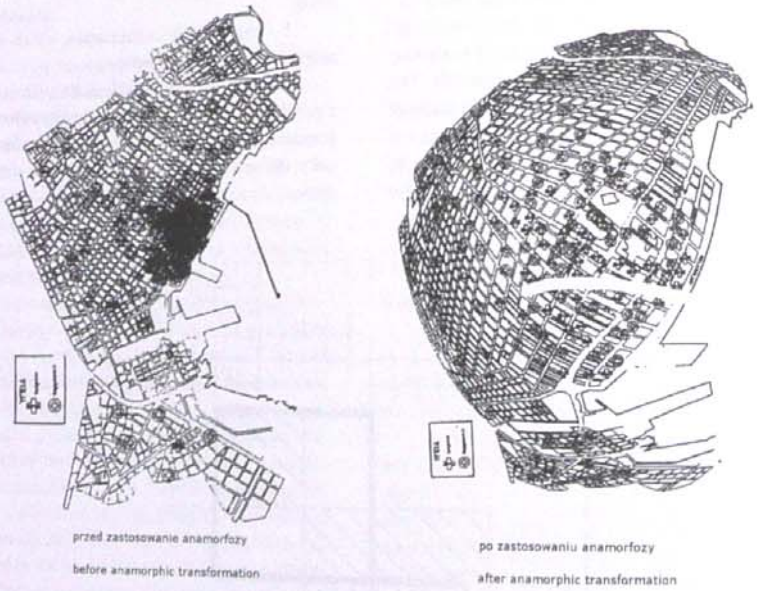


Fig. 2. Health Service in Volos (Boutoura 1993)
Rys. 2. Służba zdrowia w Volos (Boutoura 1993)

This type of variable-scale maps, despite its long tradition and merits, is not widely used at present. Their sole application is for town plans, which were drawn as views from ground observation points. These plans are one of the oldest examples of width distortion, occurring for streets the observer could not spot from the observation point, or drawing churches situated in the greater distance from the observer higher than they actually were.

The variable-scale map of Moscow from 1611 (Fig. 3) elaborated by a Polish spy for operational purposes deserves a special recognition in the history of Polish cartography (Alexandrowicz 1989). The map covers buildings, roads and city walls with gates. Wider streets going from gates to Kremlin, indicate a possible direction of the march of Polish troops in the case of storming the city. A detailed form of presentation was structured in a functional hierarchy, which allowed for the reconstruction of Kitajgorod after the fire. The Polish map was essential in this reconstruction and it became a basis for other maps, which became first detailed town plans.

MOSCOW PLAN FROM NIEŚWIEŻ (1611)

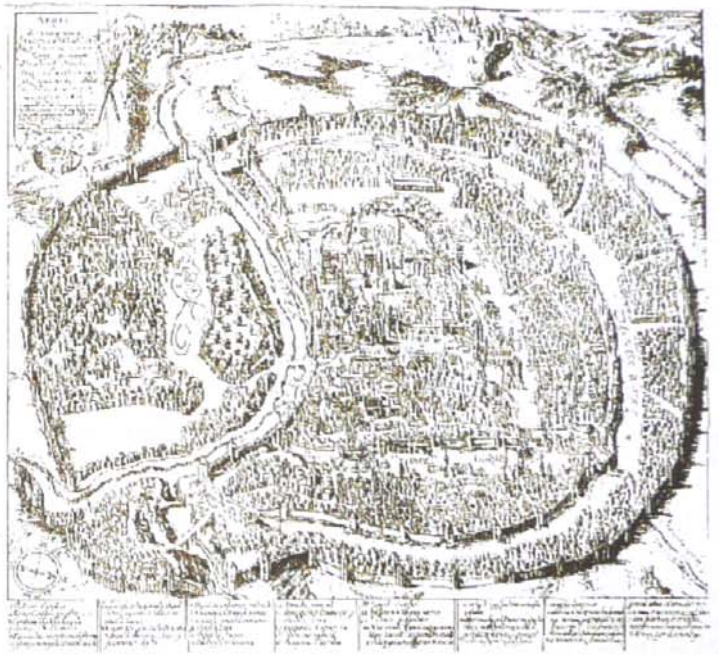


Fig. 3. Town plan of Moscow from 1611 (Alexandrowicz 1989)
Rys. 3. Plan Moskwy z roku 1611 (Alexandrowicz 1989)

The project by Naftal Kadmon, put to life by Falk's company, had a practical role as well (Kadmon 1975, 1978). Its primary aim was to decrease a plan format and make it more comfortable to use by tourists; the secondary aim was to put sub-regions in a hierarchical order. Plans are created as isotropic anamorphic maps with a central part in historic centres. Aerial photos taken with cameras that remove aberration are used as source materials. Hyperbolic, 1-kilometre grids provide the user with Euclidean characteristics. However, the aim of designers is to make the change of the scale unnoticeable.

The adjustment of the scale change to a few sub-areas located in different places and of greater importance to the user than other sub-areas is far more difficult to achieve than a partly automatic, isotropic, one-pole reduction of the scale. There is a series of projects on the transformation of constant-scale plans of Wrocław (Krzywicka-Blum 1993) into plans with a scale that is modulated in accordance with the distribution of zones crucial to the user. This series include a variable-scale map for restorers, an academic map, a map created for the need of Eucharistic Congress (Fig. 4) and a map of sport objects. The software for the conversion of coordinates was used to generate this map and formulas were derived on the strength of constructional assumptions on the scale of division zones.

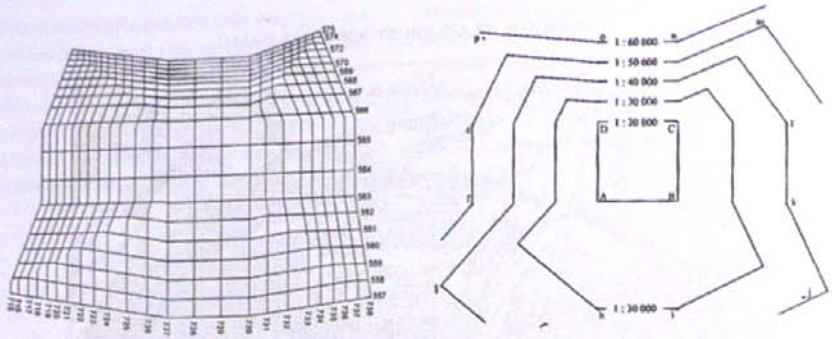


Fig. 4. 1-km grid and variability of scale on the map created for Eucharistic Congress in Wroclaw (Krzywicka-Blum 1997)
Rys. 4. Siatka kilometrowa oraz zmienność skali na mapie opracowanej na potrzeby Kongresu Eucharystycznego we Wrocławiu (Krzywicka-Blum 1997)

Anamorphic models with a functional metric (operational scale)

Time metric was used as one of the metrics in the old cartography.

The most famous anamorphic map is the Peutinger Table from the 3rd century AD. The emperor Diocletian requested to place the prototype of the Table in the Field of Mars (Sirko 2000). The table as such is a parchment scroll, 0.34m high and 37m long. The map shows 3,500 functional place symbols connected with marching routes of Roman forces. The map was designed for the need of the conquest of the area stretching from the Rhine to the Don river. The map does not provide the outline of lands and correct orientation. It only gives topological features of ordering and adjacency of elements.

"Pilgrims' maps" (Sirko 2000) constituting of many sketches are also linear anamorphic maps. They were created till the 6th century AD and depicted transcontinental routes, for example a route from London to Jerusalem. The linear transformation of length using the time metric is counted among achievements of Polish cartography. Stanisław August commissioned a group of pollsters directed by a royal cartographer named Perthes to make an inventory of a diocese in the North-East of Poland (Fig. 5). Apart from maps, the work included a descriptive database with numerical characteristics of such terms as "summer and winter availability".

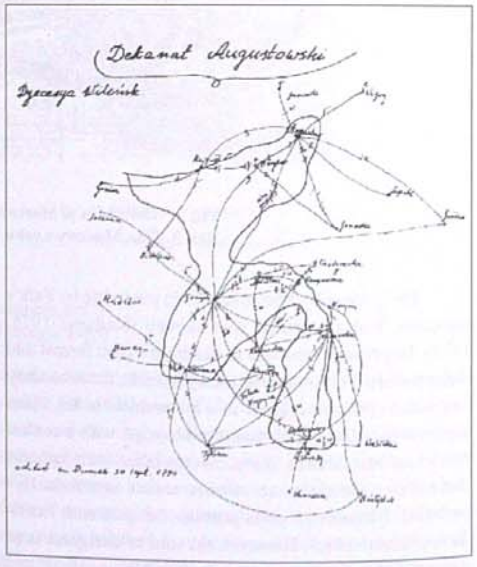


Fig. 5. Augustow Diocese (Wernerowa 1996)
Rys. 5. Mapa diecezji augustowskiej (Wernerowa 1996)

As opposed to variable scale anamorphic maps, functional roles of anamorphic models (with a functional metric) restrict themselves to illustrate a distribution of one selected quantitative feature of a real or an abstract object, e.g. population or unemployment. It can be well seen in anamorphic transformations for maps with larger scales. An example map from figure 6 is a demographic anamorphose of Opole voivodship areal units. This map has been used as a background for pancreatic cancer clustering analysis (Michalski 2003).

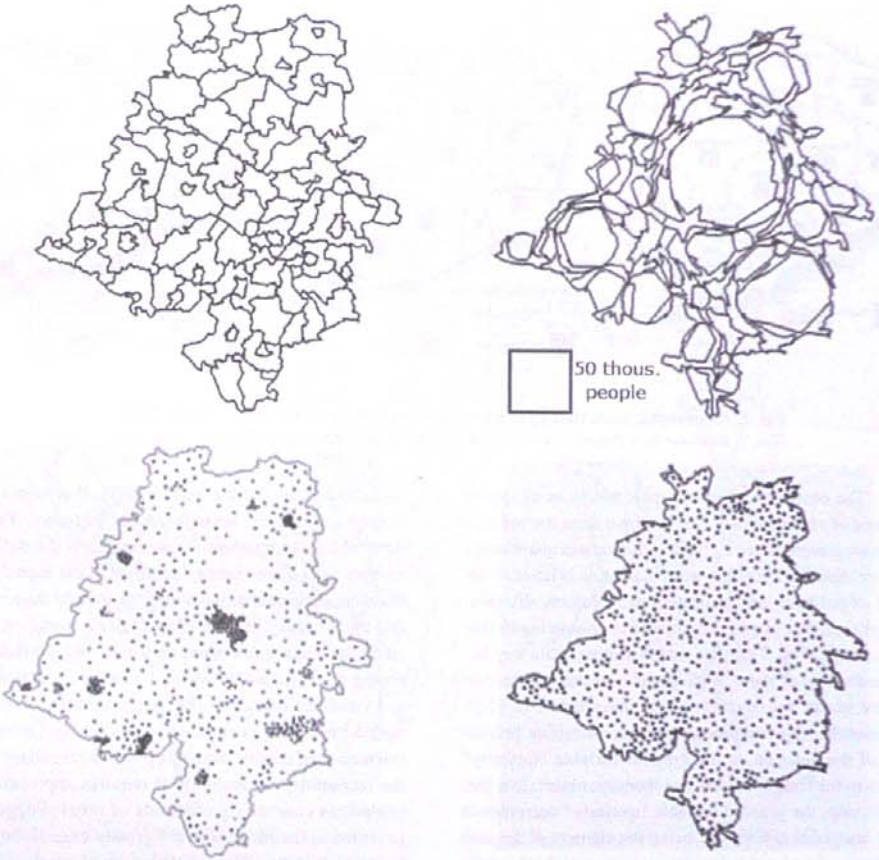


Fig. 6. Administrative borders and population anamorphic model of Opole voivodship – upper maps, pancreatic cancer cases – bottom maps (Michalski 2003)
Rys. 6. Mapa administracyjna oraz anamorfoza ludnościowa województwa opolskiego – u góry, wystąpienia nowotworu trzustki – u dołu (Michalski 2003)

Hägerstrand (1957) paid attention to the usefulness of linear notations with transformed geometry to the Euclidean distance. He used the transformation of linear scale using a logarithmic function as a transformation corresponding to transport expenses. The other well-known solutions encompass French anamorphic maps of birds' flights (Cauven 1986), American maps adjusted to a patchwork of roads, metrics that are the sum of legs lenghts, not the length of hypotenuse, and finally, functional transformation of the layout of German towns, which illustrate the temporal accessibility to the airports that can be reached by trains, buses and cars (Muller 1982, 1983).

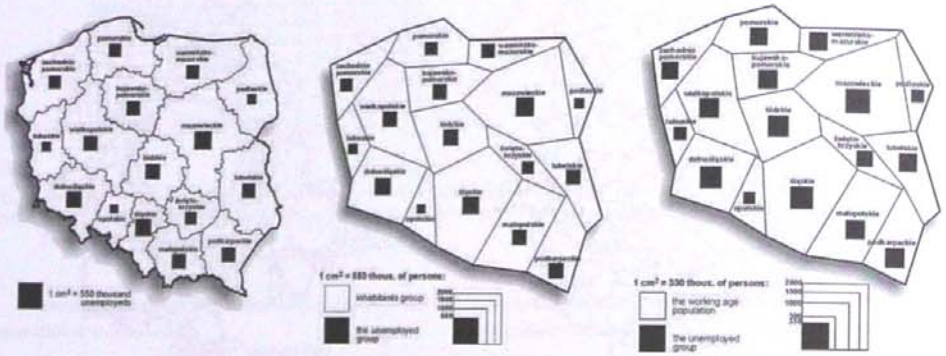


Fig. 7. Anamorphic presentations of unemployment in Poland (Krzywicka-Blum 2001)
Rys. 7. Bezrobocie w Polsce – model anamorficzny (Krzywicka-Blum 2001)

The other type of anamorphic notations of spatial features of objects which is not known from the old cartography is represented by models of a certain quantitative feature distribution. This presentation is relational for units of political, administration and property division. The anamorphic model is competitive comparing to thematic maps (Fig. 7). It has one operating scale key because the size of the area gives the user the information he/she needs, for example about investments in PLN. Anamorphic representation increases modelling properties of the notation as the graphic variable "quantity" refers to the feature of only one thematic object. In a thematic map, the graphic variable "quantity" corresponds to at least two objects: one being the element of the map background and one belonging to a layer of the thematic map. Reading information on this map is more difficult than in the case of traditional maps and may be an important limitation in perception. The user of the anamorphic model identifies a reference unit on the basis of topological conformity of the layout of fields that form a transformed and original arrangement of a mosaic. The element which is important for perception is the outline of borders of the area being identified. This aspect is reflected in suggestions concerning the transformation process. the accuracy of areas is certainly a condition to accredit the anamorphic notation to a model. Modular reasoning is expressed in terms of relations between areas and visual estimation of the area is more accurate when figures have basic, geometric shapes (Fig. 8). The retained orientation of mosaic elements plays an important role in the recognition of location. It requires appropriate assumptions concerning centroids of areas. Suggestions provided in the literature vary greatly even if they refer to the same topic. If the distribution of given reference units is not fitted to create an anamorphic map with "geometrical" units, shapes most similar to real borders are desired. It can be resolved by using the authors' method with the a genetic algorithm (Michalski 2004).

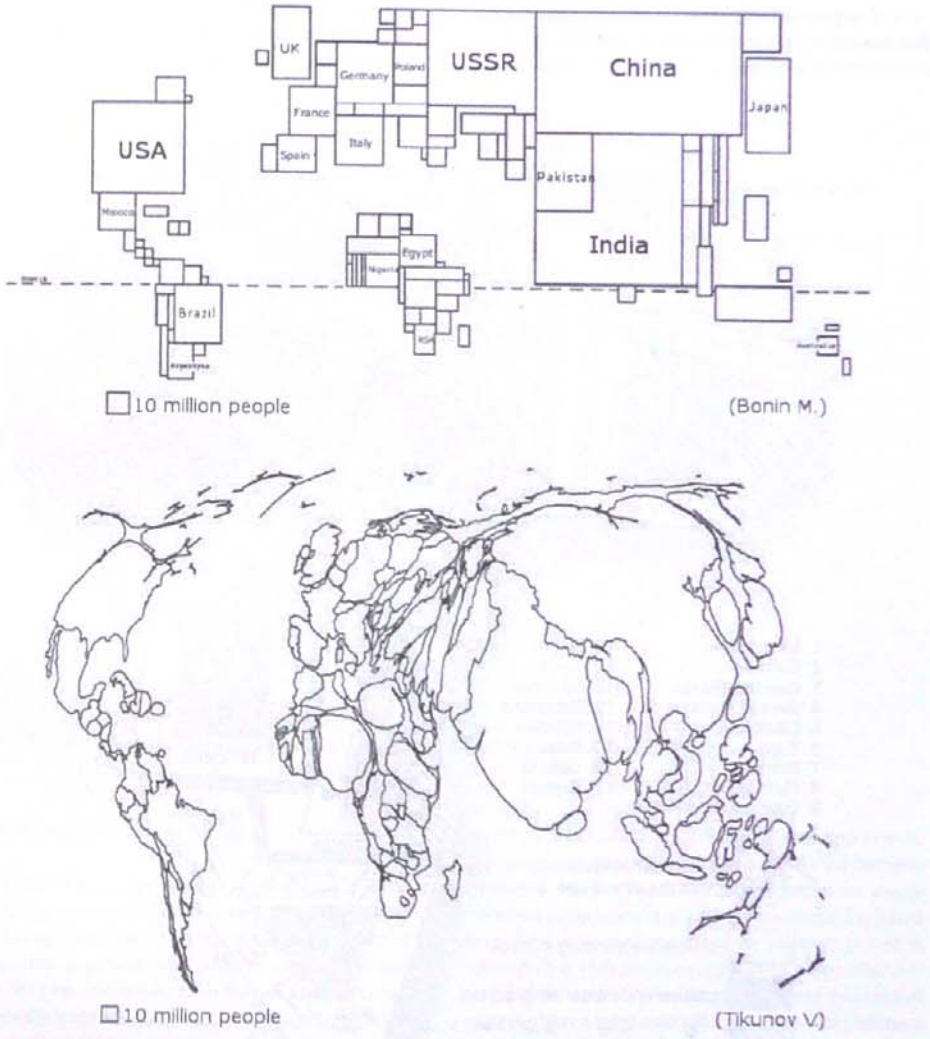


Fig. 8. Anamorphose of world's countries based on their population numbers (Bonin, unknown publication date, a map from the Internet based on Tikunov's algorithm: Tikunov 1994)
Rys. 8. Mapa anamorficzne ludności krajów całego świata (Bonin, nieznana data wydania, mapa z internetu zbudowana algorytmem Tikunova: Tikunov 1994)

Complex anamorphic notations constitute models that present two features the user is interested in or two distributions of a feature corresponding to states observed subsequently. Graphic variable "quantity" refers to a number of children at school age and a number of students (Fig.9). For two time cross-sections.

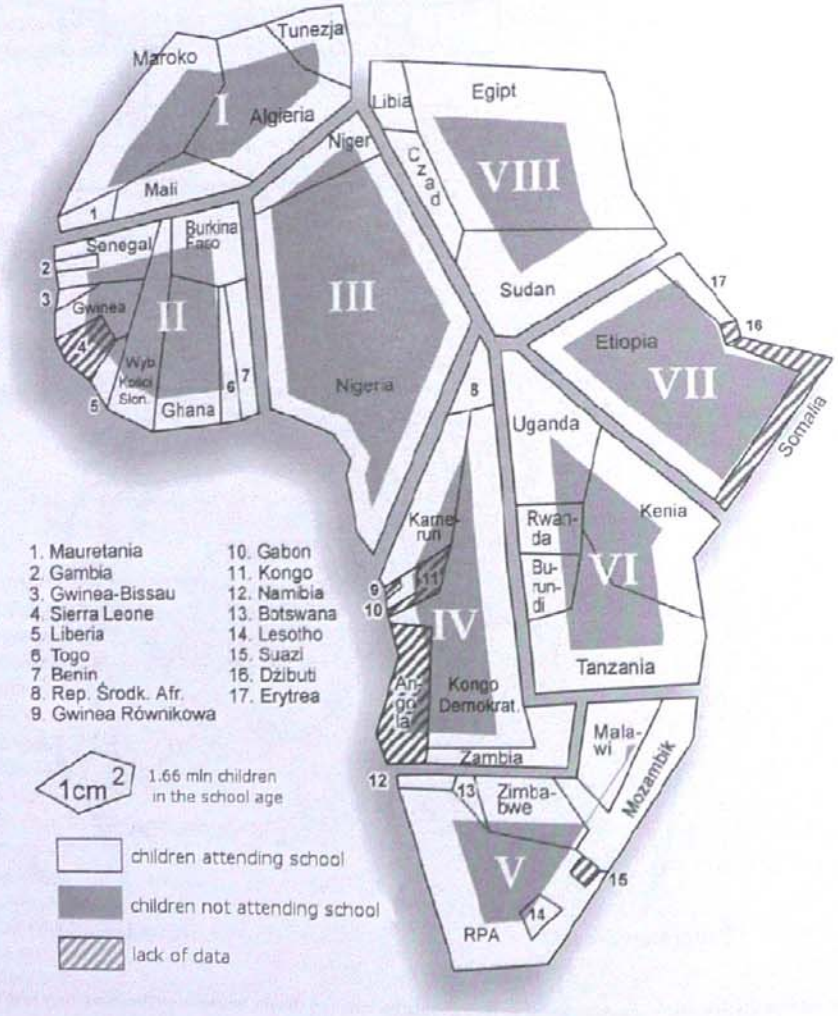


Fig. 9. Anamorphic map of children attending schools in Africa (Krzywicka-Blum)
Rys. 9. Mapa anamorficzna dzieci uczęszczających do szkół w Afryce (Krzywicka-Blum 1999)

enriching a pictorial notation not only with "quantity" but also with a second variable called "hue" enables synthetic evaluation of immigration rates in the course of two years and the indication of areas where changes (increase or decrease of the phenomenon in comparison to the past) are recorded (Fig. 10).

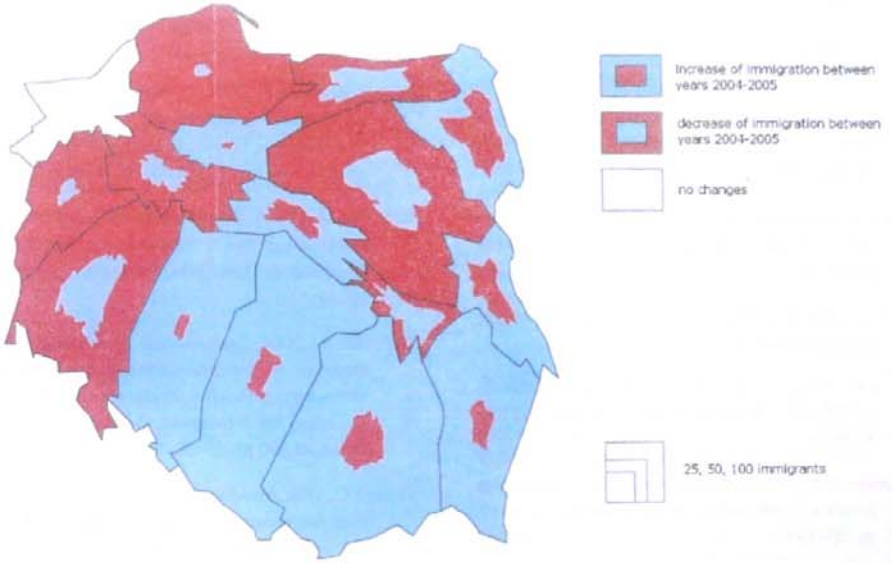


Fig. 10. Map of the changes of immigration level in years 2004–2005 (Kmiecik 2007)
Rys. 10. Mapa zmian poziomu imigracji w latach 2004–2005 (Kmiecik 2007)

Conclusion

Anamorphic maps may be regarded as credible documents in research and official regulations only if their theoretical assumptions are complemented with results of investigations on perception of different models, as it happens in the case of constant-scale thematic maps. A relevant, broadly treated education programme included in the school curriculum is essential to avoid creation of many incorrect maps with contemporary software. Features of models are often attributed to works of this type. A very useful programme for statistical methods was developed for public schools in India by a group of researchers directed by a famous mathematician named Rao (1994). Characterising anamorphic models perceptual barriers, connected with fast localisation of reference units, can be overcome using labels or attaching a traditional Euclidean map of the same territory. American scientists proved in research (Sui, Holt 2008) on anamorphic maps effectiveness that the reader's prior knowledge about that kind of maps can have a significant influence on how they interpret this unconventional form of thematic maps. The methods of making anamorphic maps were developed within the research project in which authors participated. The result of this project, GIS scripts, will be available on the web site (GIS web).

Bibliography

Alexandrowicz S., 1989, *Rozwój kartografii Wielkiego Księstwa Litewskiego od XV do połowy XVIII wieku*, UAM Poznań.

Bonin M., unknown publication date, *Les anamorphoses: population du monde, La graphique dans la presse*.

Boutora C., 1993, *A treatment of dense thematic symbolism in LIS environment*, ELIS'93, Aristotle University of Thessaloniki.

Boutora C.,1994, *Logarithmic urban thematic mapping in MIS environment*, Cartographica, vol. 31, no. 1, pp. 41–53.

Cauven C., Reymond H., 1986, *Nouvelles methods en cartographie*. GIP RECULUS, Montpellier.

GIS web, *GIS Laboratory web site: www.gislab.ar.wroc.pl*, Wrocław University of Environmental and Life Sciences.

Hägerstrand T., 1957, *Migration and area*, Migration in Sweden, Lund Studies in Geography, Ser. B, no 13, pp. 27–158.

Kadmon N., 1975, *Data Bank Derived Hyperbolic-Scale Equitemporal Town Maps*, International Yearbook of Cartography, Kirchbaum Verlag.

Kadmon N., 1978, *A Polyfocal Projection for Statistical Surfaces*, The Cartographic Journal, vol. 15, no.1, pp. 36–41.

Kmiecik A., 2007, *Modelowanie procesów demograficznych na powierzchniowych mapach anamorficznych*, Graduate work – unpublished, Instytut Geodezji i Geoinformatyki Uniwersytetu Przyrodniczego we Wrocławiu.

Kretschmer, I., Dorflinger J.,Wawrik F., eds. 1986, *Lexikon zur geschichte der Kartographie*. 2 vols. Wien: Deuticke.

Krzywicka-Blum E., 1993, *New types of city maps using operational scale*, 16th ICC Conference Proceedings 2, Cologne.

Krzywicka-Blum E., 1999, *Thematic maps at changing scales and the exemplary applications thereof*, The Polish Cartography.

Krzywicka-Blum E., 2001, *Cartographic visualization of two-component quantitative features recognized in sampled form*, Electronic Journal of Polish Agricultural Universities (EJPAU), vol. 4 issue 2.

Michalski A., 2003, *Mapy anamorficzne w badaniach epidemiologicznych*, Acta Scientiarum Polonorum: Geodesia et Descriptio Terrarum, vol. 2, no. 1, pp. 43–51.

Michalski A., 2004, *Zastosowanie anamorficznych przekształceń pól odniesienia w tematycznym modelowaniu kartograficznym*, Zeszyty Naukowe Akademii Rolniczej we Wrocławiu, Geodezja i Urządzenia Rolne, nr 500 pp. 33–52, 2004.

Muller J.C., 1982, *Non-Euclidean geographic spaces: mapping functional distances*, Geographical Analysis, vol. 14, no. 3, pp. 189–203.

Muller J.C. 1983, *La cartographie des espaces fonctionnels*, L'Espace Geographique, No. 2, pp. 142–153.

Rao C.R., 1994, *Statystyka i prawda*, PWN, Warszawa.

Sirko M., 2000, *Zarys historii kartografii*, UMCS Lublin.

Sui D., Holt J., 2008, *Visualizing and Analysing Public-Health Data Using Value-by-Area Cartograms: Toward a New Synthetic Framework*, Cartographica, vol. 43, issue 1, pp. 3–20.

Tikunov V., 1994, *Numerical methods in the compilation of transformed images*, Mapping Science and Remoting Sensing, vol. 31, no. 1, pp. 66–85.

Tobler W.R., 2004, *Thirty Five Years of Computer Cartograms*, Annals of the Association of American Geographers, March 2004, vol. 94, no. 1, pp. 58–73.

Wernerowa W. (editor), 1996, *Rękopiśmienne opisy parafii litewskich z 1784 roku*, Instytut Historii Nauki PAN, Warszawa.

Praca naukowa finansowana ze środków na naukę w latach **2007–2009** jako projekt badawczy.

The scientific work funded from financial means to support science within the years **2007–2009** as a research project.

Streszczenie

Niezbędnym elementem każdej definicji mapy jest określenie relacji między oryginałem a jego ujęciem w notacji. Mówi się więc o skali pomniejszenia, podobieństwie, uproszczeniu, wreszcie – matematycznych podstawach, jako kluczu umiejscowienia oznaczeń reprezentujących obiekty. Zgeometryzowana notacja obrazowa, w której podstawą umiejscowienia oznaczeń obiektów jest celowe przekształcenie metryki euklidesowej bądź – inna metryka lub semimetryka,wybrana w dostosowaniu do określonej funkcji użytkowej modelu, nosi nazwę anamorfozy. W artykule zaprezentowano próbę klasyfikacji podstaw matematycznych map ze względu na ich funkcje praktyczne. Klasyfikacja ta odnosi się do wyboru metryki oraz zasad umiejscowienia elementów treści map i wygląda następująco: A – mapy opatrzone skalą, której wahania użytkownik może analitycznie wyznaczyć (metryka euklidesowa), B – mapy zmiennoskalowe anamorficzne (umiejscowienie elementów zależne od ich gęstości lub znaczenia), C – modele anamorficzne o metryce funkcjonalnej (skali operacyjnej). Zawierające się w pierwszej grupie mapy, w podziale na ogólnogeograficzne i tematyczne, stanowią modele euklidesowo rozumianych relacji przestrzennych. Ogólnogeograficzne, stanowią modele rozkładu przestrzennego i kształtu obiektów, zależnie od rozmiarów przedstawianych obszarów, mniej lub bardziej zgodne z rzeczywistością rozumianą euklidesowo. Mapy tematyczne charakteryzuje hierarchiczność wyboru i ujęcia elementów treści, zależnie czy należą one do tła ogólnogeograficznego, czy stanowią treść tematyczną. Odwzorowanie tła musi zapewniać poprawność tych przestrzennych cech oznaczeń, które są priorytetowe dla charakterystyki obiektów tematycznych, a więc: kształtu – dla elementów morfologicznych, czy pól powierzchni – dla podziału politycznego. Drugą grupę stanowią mapy zmiennoskalowe. Istotą takich map jest ich lokalne podobieństwo do map o stałej skali. Dotyczy to sposobu ujęcia treści charakterystycznego dla map stałoskalowych. Modulowanie skali ma na celu: a – wyrównanie ciężaru graficznego notacji, w przypadku bardzo zróżnicowanego zagęszczenia obiektów w przedstawianym obszarze, b – wyraźnie dostosowane do operacyjnych funkcji – uwypuklenie pewnych podobszarów. Mapy zmiennoskalowe znalazły swoje za-

stosowanie między innymi w serii map (planów) miast oficyny Falka. Trzecia grupa zawiera mapy (modele) anamorficzne, gdzie następuje zmiana funkcji odległości lub powierzchni w stosunku do tradycyjnego ujęcia euklidesowego. W kilku ostatnich dziesięcioleciach zauważa się zwłaszcza rozwój algorytmicznych metod budowy powierzchniowych map anamorficznych. W stosunku do kartodiagramów powierzchniowy model anamorficzny jest konkurencyjny. Ma jeden, praktycznie sprawny klucz skalowy, gdyż wielkość pola powierzchni wyraża wprost informację ważną dla użytkownika, a więc, dla przykładu, nakłady inwestycyjne w złotych. Ujęcie anamorficzne podnosi też własności modelowe notacji, bowiem zmienna graficzna „wielkość" dotyczy charakteryzowanej cechy tylko jednego obiektu tematycznego, a nie, jak w kartodiagramie – co najmniej dwóch: jednego – należącego do tła i co najmniej jednego – do warstwy tematycznej mapy. Istotną barierę percepcyjną stanowi utrudnione, w stosunku do map tradycyjnych, odczytanie umiejscowienia informacji. W modelu anamorficznym użytkownik identyfikuje jednostkę odniesienia na podstawie topologicznej zgodności układu pól tworzących przekształconą i źródłową kompozycję mozaiki. Ważny percepcyjnie element stanowi kształt granic identyfikowanego obszaru i ten aspekt znalazł swe odbicie w propozycjach różnie przyjmowanych założeń dotyczących procesu przekształceń. Na to, aby mapy i modele anamorficzne zyskały rangę dokumentów wiarygodnych w pracach poznawczych i regulacjach urzędowych należy prowadzić dalsze badania percepcji takich opracowań, podobnie, jak miało to miejsce w odniesieniu do tradycyjnych dziś już postaci tematycznych map o stałej skali.

JACEK STRZELCZYK, TOMASZ DANEK[1]

CREATION AND APPLICATION OF HETEROGENEOUS GRID ENVIRONMENT FOR SEISMIC MODELING

Key words:

grid, heterogeneous computing, seismic modeling

Abstract

Grid computing is a simultaneous application of several computers to one, single, computationally intensive problem. A computer grid can be created from special, dedicated hardware, but also unused CPU cycles of available desktop machines can be exploited. This article reports on a successful effort of creating a heterogeneous grid infrastructure. The installation consisted of two Blade servers and one PC machine that were available at the time. Heterogeneity was present on both levels: hardware platform and operating system. Blade machines ran Linux Fedora Core, while PC ran Windows 2000 SP4. Using only free grid software, the authors created working infrastructure and used it in seismic modeling calculations.

OPIS UTWORZENIA I ZASTOSOWANIA HETEROGENICZNEGO ŚRODOWISKA GRIDOWEGO DO MODELOWAŃ SEJSMICZNYCH

Słowa kluczowe:

grid, obliczenia heterogeniczne, modelowanie sejsmiczne

Abstract

Termin *grid computing* oznacza jednoczesne wykorzystanie kilku komputerów do rozwiązywania jednego, złożonego problemu obliczeniowego. Komputerowy *grid* może zostać stworzony ze specjalnych urządzeń dedykowanych, przeznaczonych do tego celu, bądź także z nieużywanych procesorów dostępnych w biurowych komputerach PC. W tym artykule opisano udaną próbę stworzenia heterogenicznej infrastruktury gridowej. System składał się z dwóch serwerów typu IBM Blade oraz z jednej maszyny PC, która była aktualnie dostępna. Heterogeniczność była obecna na obu poziomach: platformy sprzętowej oraz systemu operacyjnego. Na maszynach typu IBM Blade zainstalowany był system Linux Fedora Core, a komputer PC działał pod kontrolą systemu Windows 2000 SP4. Używając jedynie darmowego oprogramowania gridowego, autorzy stworzyli działającą infrastrukturę i użyli jej w procesie modelowania sejsmicznego pola falowego.

[1] AGH University of Science and Technology, Department of Geoinformatics and Applied Computer Science, Krakow, Poland.

1. Introduction

1.1 The Grid Technology

Grid technology helps to utilize resources of many, even geographically dispersed computers that can cooperate in common tasks. Back in 1998 Ian Foster and Carl Kesselman defined grid as "a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities."[1]

According to this definition, a grid could be easily created from computers running various operating systems on different hardware platforms. Then, its feature *pervasive* could be fully fulfilled. Unfortunately, free solutions for creating such a heterogeneous infrastructure are not common and easy to find. The authors thought it is a big gap in a grid computation field and decided to investigate the possibility of creating a grid network consisting of different hardware platforms and operating systems using only free grid software.

The authors decided to use machines available at our university and created a grid of two IBM Blade units running Linux Fedora Core and PC running Windows 2000. No publications could be found on this topic before, therefore the authors believe it is an unprecedented installation.

1.2 Seismic Wave Modeling

The computational power was needed to solve a real geophysical problem which was seismic wave modeling. Acoustic or elastic wave field modeling is a tool which is common in modern seismic or seismology. It can be used for oil, gas and other natural resources exploration [2], and for various analyses of natural earthquakes [3]. In the past few years growing popularity of this method caused a rapid development of new algorithms and techniques which are usually computationally very expensive. Fortunately, this kind of problems are easy to parallelize effectively [4], but especially for large scale problems, resources needs are almost unlimited. In this kind of modeling even a small additional computational power counts.

2. Created Heterogeneous Grid Environment

In this project a simple *cluster grid* was created. The open source version of the SGE, called The Grid Engine v6.2, was used. The environment was heterogeneous. For test purposes it consisted of two blade units and one PC.

2.1. Hardware Configuration

Blade units were of HS21 series. They had two Dual-Core Intel Xeon 2.8 GHz processors with 6 MB L2 shared cache and 3.5 GB of memory the network interface was Dual Gigabit Ethernet.

A personal computer that was used had Intel Pentium 4 2.8 GHz processor and 1 GB of memory, the network card was Realtek RTL 8139.

2.2. Software used

Two blade servers ran Linux Fedora Core 8 (Warewolf) system, kernel 2.6.26.6–49.fc8, and PC ran Windows 2000 SP4. On the Windows system Services For Unix 3.5 were used to allow the execution of the Grid Engine version 6.2. The Grid Engine project "provides enabling distributed resource management software for wide ranging requirements from compute farms to grid computing." [5] MPICH2 1.0.7 was installed to make communication between parallel processes possible. Developing software to compile programs executed on the grid was gcc 3.4.4 (on Fedora) and Dev-Cpp 4.9.9.2 (on Windows).

2.3. Software Installation

NIS

NIS is a service that provides information, which has to be known to all machines on the network. It can be very helpful in maintaining a coherent user structure on all nodes in the grid. Full NIS HOWTO can be found on the NIS web page [6]. For purposes of this installation one user account was needed. It was named *sgeadmin* and added to the NIS database. The *HOME* variable was set to /usr/grid.

NFS

A common filesystem is necessary to install and run SGE in a simple and flexible manner. It can be achieved in many ways, but in this case, the Network File System was chosen. Full NFS HOWTO can be found on the NFS web page [7]. In this installation, directory /usr/grid was shared by all machines.

Windows Services for UNIX

„Windows Services for UNIX version provides a full range of cross-platform services for integrating Windows into existing UNIX-based environments."[8] It includes the Interix subsystem, which "provides a fully POSIX--compilant environment that runs a native subsystem in the Windows kernel." [9] This environment provides two shells – *Korn* and *C*. Beside that, SFU includes the User Name Mapping service. It is "a component of Microsoft Windows 2000 Services for UNIX. The User Name Mapping service is used by UNIX-based Server for NFS, Client for NFS, and Gateway for NFS, as well as Remote Shell Service (rshsvc), for mapping UNIX-based network user names to Windows-based network user names and vice versa." [10] Keeping coherent user system is vital for proper working of heterogeneous installation, because users both from Linux and Windows machines have to have access to the same directory with GridEngine and MPICH2 installation.

MPICH2

MPICH2 was installed to provide communication between processes running in one task. It is "a high-performance and widely portable implementation of the Message Passing Interface (MPI) standard" [11]. At first, a/the Linux version of MPICH2 was compiled and installed on one of the Blade servers – the *master host* – under the /usr/grid/mpich2 directory. The installation was shared with the second Linux node via NFS. Then, the MPICH2 library was installed independently on the PC from the dedicated Windows version. It was started as a system service. To enable communication between Linux and Windows versions, the MPICH2 daemon owners on all machines have to be the same users – in this case *sgeadmin*.

Grid Engine

The resource management was done by the Grid Engine software. It was installed on all machines that established the grid. One of the Blade machines was a *master host* responsible for managing tasks and resources. The same machine was also running a Grid Engine execution daemon which made it also a *slave node*. The second Blade machine and PC were only *slaves*. The full installation process is described in detail on the Sun Microsystems Documentation web page. [12] For the Windows host, specific platform files of Grid Engine were installed (available after registration on the Sun page). The same code for seismic modeling was executed on all machines, but it had to be compiled for the a specific platform. A source code was compiled separately for the Linux machines and for the Windows PC, and put in the execution environment. Additionally, to enable MPI programs on the grid, MPICH2 parallel environment was added to the Grid Engine configuration. The installation procedure is described in detail on the Grid Engine wiki page [13].

The installation procedure is described in detail on the Grid Engine wiki page [13].

3. Seismic Modeling

For test runs of our installation the authors used classic Alford algorithm [14] for acoustic wave field modeling. In this algorithm acoustic wave equation is as follows:

$$\frac{\partial^2 p}{\partial t^2} - c^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2}\right) = f(x, z, t) \qquad (1)$$

where $p(x,z)$ is pressure, $c(x,z)$ is velocity of acoustic wave, t is time and f denotes function which describes pressure change in source.

The above equation is approximated by the second order finite difference scheme:

$$p(i,j,k) = 2(1-2\gamma^2)\,p(i,j,k) - p(i,j,k-1) + \gamma^2 (p(i+1,j,k) + p(i-1,j,k) + p(i,j+1,k) + p(i,j-1,k)) \qquad (2)$$

where $\gamma = c\Delta t/\Delta h$, Δt is a time sampling interval, Δh is a distance between grid points in x and z directions. The stability criterion for the above scheme is: $\gamma \leq 1/\sqrt{2}$.

The authors decided to use typical absorbent boundaries condition by Reynolds [15] as border conditions t.

4. Results

4.1. Heterogeneous Environment Effectiveness

In the experiment, the seismic wave modeling for 20 shot points (SP) was done. Results were calculated for one up to nine slave processors' cores and execution times were measured. For each number of processors' cores, the program was executed ten times and the average execution time was calculated to level temporary operating system influences on processors. The results of the measurement are shown in Figure 1.

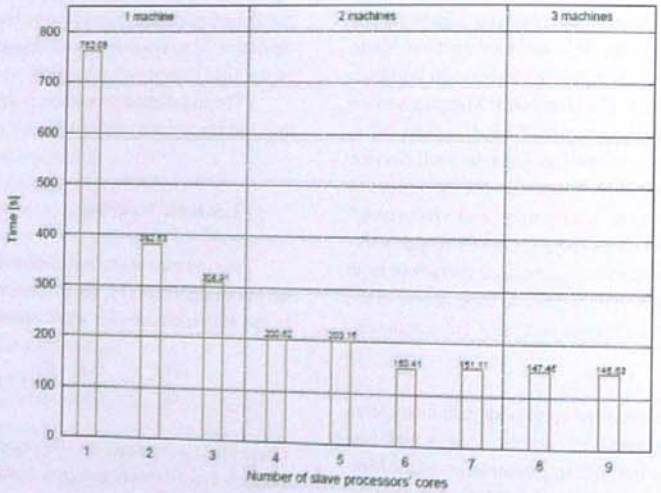


Fig. 1. Seismic wave modeling program execution times for 20 shot points
Rys. 1. Czasy działania programu do modelowań fali sejsmicznej dla 20 punktów strzałowych

5. Discussion

A deeper analysis of the experiment's results showed few interesting phenomena. Primarily, adding processor's second core to the second machine did not accelerate the program execution. Adding the fourth core, in this case, even slowed it down. To investigate this effect the authors decided to see how the Linux operating system allocates CPU resources.

Since first two blade machines have both four processors' cores, adding one PC with just one 2-core CPU cannot considerably decrease the execution time. Additionally, according to Amdahl's law program acceleration does not increase linearly with the number of processors. However, the difference after adding the Windows host to the grid is noticeable. Seismic wave modeling executes even 4% (6 seconds) faster with the Windows PC in the grid.

Despite that adding the machine with different operating system to the grid did not decrease the calculation time considerably, it was proved that the creation of heterogeneous grid environment is possible. One task can be dispersed on different software and hardware platforms.

In the first phase, the number of shot points calculated by each slave was checked. It proved that the slave process number two (that is the third process on the first machine – the first process is a/the master) calculates more shot points than the others. Then, calculation times for each SP were measured in each slave process. The results are shown in Table 1. Slave 2 calculates one SP on average almost 15% (7 seconds) faster than other slave processes. This was the reason why it calculated more shot points in one task.

Afterwards, the CPU cores load analysis were done on the first Linux node while executing master process and one, two, and three slaves. It proved that the master process, which did not make any calculations, used CPU in nearly 0%. Slaves used CPU in 100%, so they were redeployed between cores constantly by the CPU load balancing algorithm. In the next step, multi-core processors efficiency tests were made to discover why the second slave process was faster than the others. The authors suspected that it was due to the fact, that slave 2 most of the time executed on the same CPU as a master process. To test multi-core processors efficiency, different number of slave processes were executed on the second Blade machine, and the master on the first. The results are shown in Table 2.

Table 1. Calculation times for one shot point on each slave
Tabela 1. Czasy obliczeń na każdym procesorze dla jednego punktu strzałowego

	Slave 1 [s]	Slave 2 [s]	Slave 3 [s]	Slave 4 [s]	Slave 5 [s]
	50	38	52	50	50
	50	39	51	49	50
	48	38	50	49	48
	50	39	51	25	25
	49	38	51	50	50
	48	39	50	48	49
	52	38	50	51	51
	49	38	51	51	49
	50	39	48	49	48
	51	39	39	49	51
	51	38	39	51	50
	50	38	38	49	48
	39	50	51	49	48
	49	50	49	50	50
	38	48	47	49	50
	38	42	48	49	48
Average:	**47.63**	**40.75**	**47.86**	**48**	**47.81**

Table 2. 20 shot points calculation times for different number of slaves on one machine
Tabela 2. Czasy obliczeń 20 punktów strzałowych dla różnej ilości procesów niewolniczych na jednej maszynie

	One slave [s]	Two slaves [s]	Three slaves [s]	Four slaves [s]
	472.73	492.84	250.51	248.34
	473.08	492.24	245.34	248.76
	473.10	492.74	245.74	247.96
	472.83	481.36	245.80	248.25
	473.00	491.13	245.56	249.27
	473.22	495.38	245.63	248.10
	473.80	490.10	245.45	247.15
Average:	**473.11**	**492.26**	**446.29**	**448.26**
Acceleration:	**1**	**0.96**	**2**	**1.91**

It can be seen that two processes cooperating in one task work slower than one process calculating the same problem individually. It probably means that for the first two slaves two cores of the same CPU were assigned, and for the next two slaves two cores of the second CPU were used. So, if four processes use 100% of the cores, the sequence of assigning CPU is as follows: first core of a first CPU, second core of a first CPU, first core of the second CPU and second core of the second CPU. It also explains the effect visible on Figure 1, where the calculation times for 4 and 5 slave cores are almost equal, and the same for 6 and 7 cores. As a result of those experiments the authors concluded that, in this case, using multi-core processors is not beneficial and can even slow down the calculation process.

According to those conclusions, slave's 2 shorter execution time probably comes from the fact, that the load balancing algorithm most often assigned it to the same processor as a master process. Since the master did not make any calculations and did not load processor, slave 2 could used more CPU resources than the other slaves. The authors plan to investigate this phenomenon more deeply in the future.

6. Conclusions

The main goal of this project was achieved. A heterogeneous grid infrastructure was created successfully. It was based on different hardware and software platforms.

It showed that even a temporarily available and not fully exploited machines can constitute a grid. This can be considered as a step towards a better utilization of CPU cycles of many different machines in a common goal.

Since Linux and Windows are two of the most popular operating systems, they were chosen for this installation. The authors believe that the combination of servers and desktop machines running various systems can be worthwhile and give positive effects by considerably improving calculation effectiveness and reducing its time. That kind of connection can also improve resources utilization and may lead to many cost cuts.

There were also some unexpected results in this research. The authors believe they found a problem with proper working of dual-core processors in Blade machines running Linux Fedora Core, kernel 2.6.26.6–49.fc8. One process, on one CPU core, accomplishing a task, works faster than two processes cooperating in the same task on two cores of the same CPU. The authors plan to investigate this topic in the future.

Acknowledgments. This work was financed by the AGH – University of Science and Technology, Faculty of Geology, Geophysics and Environmental Protection as a part of statutory project number 11.11.140.561.

Reference

Foster, I., Kesselman, C., *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers Inc., San Francisco, CA, 1998

Pietsch K., Marzec P., Kobylarski M., Danek T., Lesniak A., Tatarata A., Gruszczyk E., *Identification of seismic anomalies caused by gas saturation on the basis of theoretical P and PS wavefield in the Carpathian Foredeep*, SE Poland //Acta Geophysica; ISSN 1895–6572. 2007 vol. 55 no. 2

Danek T., Pieta A., Lesniak A., *Simulation of seismic waveforms from "Rudna" copper mine*, Poland, using staggered grid // W: 31st General assembly of the European Seismological Commission: Hersonissos, Crete, Greece, 7–12 September, 2008

Danek T., Franczyk A., *Parallel and distributed seismic wave-field modelling // TASK Quarterly: scientific bulletin of Academic Computer Centre in Gdansk*; ISSN 1428–6394. – 2004 vol. 8 no. 4

Grid Engine Project, http://gridengine.sunsource.net/

Network Information Service, http://www.linux-nis.org/nis-howto/HOWTO

Network File System, http://nfs.sourceforge.net/nfs-howto/

MPICH2: *High-performance and Widely Portable MPI*, http://www.mcs.anl.gov/research/projects/mpich2/

Windows Services for UNIX, http://technet.microsoft.com/en-us/interopmigration/bb380242.aspx

Microsoft Support, http://support.microsoft.com/kb/324081

Microsoft TechNet: *User Name Mapping Service*, http://technet.microsoft.com/en-us/library/bb463221.aspx

Sun Microsystems Documentation, http://docs.sun.com/app/docs/doc/817–6118/emrar?q=N1GE&a=view

Strzelczyk J., *Install and configure Grid Engine in heterogeneous environment on Linux and Windows with MPICH2*, The Grid Engine wiki, http://wiki.gridengine.info/wiki/index.php

Alford R. M., Kelly K. R., Boore D. M., 1974, *Accuracy of finite-difference modeling of the acoustic wave equation*, Geophysics, 39(6).

Reynolds A. C., 1978, *Boundary Conditions for the Numerical Solution of Wave Propagation Problems*, Geophysics, 43.

Streszczenie

Termin *grid* został zdefiniowany w roku 1998 przez Iana Fostera i Carla Kesselmana jako „infrastruktura sprzętowa i software'owa dostarczająca niezawodny, spójny, rozpowszechniony i niedrogi dostęp do zasobów obliczeniowych". W tej pracy autorzy postanowili ułatwić rozpowszechnianie systemu *grid* opracowując metodę tworzenia heterogenicznej infrastruktury, która pozwoli połączyć ze sobą maszyny obliczeniowe działające pod kontrolą różnych systemów operacyjnych.

Głównym zastosowaniem utworzonego systemu było przeprowadzanie złożonych obliczeniowo modelowań fali sejsmicznej. Modelowania akustycznych bądź elastycznych pól falowych są popularnym narzędziem w nowoczesnej sejsmice i sejsmologii, a przy tym są zadaniem wymagającym dużych zasobów obliczeniowych.

Utworzona infrastruktura składała się z dwóch serwerów typu IBM Blade z działającym na nich systemem Linux Fedora Core 8, oraz z jednego komputera PC z systemem Windows 2000 SP4. Za pomocą darmowego oprogramowania Sun Grid Engine utworzono system typu *cluster grid*, w którym wiele maszyn zlokalizowanych w jednym miejscu może pracować wspólnie nad jednym zagadnieniem. Dla zapewnienia spójnego dostępu do danych dla wszystkich maszyn, oraz w celu umożliwienia komunikacji i zarządzania zadaniami użyto następujących pakietów oprogramowania: NIS, NFS, Windows SFU, MPICH2 oraz Grid Engine.

Po utworzeniu i przetestowaniu heterogenicznej infrastruktury dokonano modelowań fal sejsmicznych z jej wykorzystaniem. Stwierdzenie poprawności działania instalacji łączącej różne systemy było pierwszym sukcesem pracy. Badania pokazały, że współpraca różnych systemów operacyjnych w obrębie *cluster grid* nie stwarza znaczących problemów związanych z wydajnością takiej infrastruktury. Umożliwia natomiast lepsze wykorzystanie dostępnych zasobów obliczeniowych w postaci biurowych komputerów PC, które najczęściej pracują w systemie Windows, a przez to szybsze otrzymanie wyników modelowania sejsmicznego.

ZBIGNIEW KASINA[1]

MODEL STUDY OF VARIATIONS OF SPECTRUM CENTROIDAL FREQUENCY OF TURNING WAVE PROPAGATING IN NEAR-SURFACE LAYER[2]

Key words:

geophysics, seismic methods, turning waves, low velocity layer, attenuation, centroidal frequency

Abstract

In the presented paper the results of seismic ray trajectories and traveltimes calculations have been described for the case of turning wave propagating in the discrete gradient medium of the near-surface layer. These results were the base of defining the medium attenuation characteristics of each ray and spectra of assumed source signals in the receiver points with their spectrum centroidal frequencies. The analysis of the trajectories distributions, variations of the spectra and their centroidal frequencies was conducted for different source signal parameters and different parameters of the sejsmogeological model of the near-surface layer. In the calculations the scheme of surface seismic aquisition was taken into consideration.

STUDIUM MODELOWE ZMIAN CZĘSTOTLIWOŚCI CENTROIDALNEJ WIDMA FALI REFRAGOWANEJ PROPAGUJĄCEJ W WARSTWIE PRZYPOWIERZCHNIOWEJ

Słowa kluczowe:

geofizyka, metody sejsmiczne, fale refragowane, strefa małych prędkości, tłumienie, częstotliwość centroidalna

Abstrakt

W pracy przedstawiono wyniki obliczeń trajektorii i czasów przebiegu promieni sejsmicznych fali refragowanej w dyskretnym ośrodku gradientowym strefy przypowierzchniowej. Wyniki tych obliczeń były podstawą określenia charakterystyk tłumienia ośrodka odpowiadających poszczególnym promieniom oraz widm założonych sygnałów źródła w punktach odbioru wraz z częstotliwościami centroidalnymi tych widm. Dokonano analizy rozkładu trajektorii, zmian widm i ich częstotliwości centroidalnych dla różnych parametrów sygnału źródła i różnych parametrów modelu sejsmogeologicznego przypowierzchniowej strefy ośrodka. W obliczeniach uwzględniono schemat obserwacji sejsmiki powierzchniowej.

[1] University of Science and Technology AGH, Faculty of Geology, Geophysics & Environment Protection, Department of Geophysics, Cracow, Poland.

[2] The paper was prepared within the project of MNiSW no. N N525 168735 and statutory works of MNiSW in the Department of Geophysics contract AGH no. 1111.140.06.

Introduction

The attenuation of seismic waves propagating in the geological medium plays an essential role in imaging geological information in the wave pattern of the recorded seismic traces. The variations of the pore fluid (especially the occurrence of a gas), the variations of the medium consolidation (especially the occurrence of the fractured zones and the zones of the anomalous stress) are a source of essential variations of the medium attenuation coefficient. These variations cause considerable changes of the seismic signal amplitude as well as its frequency band. The interpretation of the attenuation coefficient distributions is an important source of additional geological information, essentially broadening the base of inference in comparison with traditionally interpreted velocity distributions. The development of this base plays a special role in the reservoir seismics (seismic monitoring, enhanced oil recovery) as well as in the engineering seismics (estimation of the elastic parameters of a medium) and in the mining seismics (estimation of the medium heterogeneity, predicting the zones of anomalous stress). An increased interest in the attenuation problem was expressed in the last years in the publications aimed at the theory and application of that parameter in deep seismics (Carcione 2000, Parra 2000, Chichinina *et al.* 2006, Rossi *et al.* 2007, Liu *et al.* 2007, Payne *et al.* 2007, Zhu 2007) as well as in the investigation of the near-surface layer (Best *et al.* 2007, Yadari *et al.* 2008). The investigation of the attenuation in the shallow part of the seismogeological model plays an important role in the estimation of filtering characteristics of this zone. These characteristics are also used in the estimation of frequency variations of propagating signal in the process of the design of 3-D acquisition.

One of the information sources about the attenuation coefficient (or quality factor Q) distribution may be the results of the tomographic inversion realized in the framework of amplitude tomography based on the refracted (turning) waves. The effectiveness of that inversion depends on the range of spectrum variations of propagating signal. These variations are determined by the source signal parameters as well as by the velocity and attenuation distributions in the near-surface zone. Therefore in the presented paper an analysis of the variations of the spectrum centroidal frequency of the turning wave signal has been undertaken. That frequency is the basic parameter one of the most effective method of solving the inverse problem of the amplitude tomography (Quan, Harris 1997, Kasina 2008a,b).

1. The methodology of numerical modeling of the spectrum centroidal frequency variations of the turning wave signal

In the program used for the generation of the frequency characteristics of the medium attenuation the following relation was applied:

$$H(f) = \exp\left[-f^p \int \alpha_0 \, dl\right] = \exp\left(-f^p \cdot QL\right) \qquad (1)$$

where the curvilinear integral QL is defined along the ray trajectory L from the source to the receiver in the process of curvilinear seismic ray tracing, and the coeffcients p and α_0 define the relation between the attenuation coefficient and frequency:

$$\alpha = \alpha_0 f^p$$

In the case of consolidated media most often the assumption is made – as in the applied program – that p = 1. In that case the coefficient α_0 is defined by the following relation:

$$\alpha_0 = \frac{\pi}{Qv}$$

where Q is the so called quality factor and v is the wave propagation velocity.

The value of the curvilinear integral QL in the relation (1) was calculated – for the assumed discrete distributions of velocity and Q factor – in the process of ray tracing, realized by means of solving the set of differential equations resulting from the Fermat principle (Kasina 2001):

$$d\alpha = \frac{1}{v}\left[\frac{\partial v}{\partial x}\sin\alpha - \frac{\partial \alpha}{\partial z}\cos\alpha\right] ds$$

$$dx = ds\cos\alpha$$

$$dz = ds\sin\alpha$$

$$dt = \frac{ds}{v}$$

where angle α defines the slope of the tangent to the ray trajectory, ds is the element of the ray trajectory, v – velocity defined along the ray trajectory. In the process of ray tracing we solve the above set of differential equations using numerical Runge-Kutta method of the fourth order.

The calulated – along each ray trajectory – attenuation characteristicts of the medium were used to define the variations of the signal spectrum on the way from a source to receivers. The source signal spectrum (most often in the form of Gauss or rectangular spectrum) was multiplied by the appriopriate attenuation characteristic of the medium. The spectrum of Gauss signal had the well known form:

$$S(f) = e^{\frac{-(f-f_0)}{2\sigma_S^2}} \qquad (2)$$

where f is a frequency, σ^2_S is the spectrum variance, f_0 is the dominating spectrum frequency.

The discrete signal spectra were used to calculate the centroidal frequencies f_G and f_S of the signal spectra in the receiver point G and the source point S in accordance with the relation:

$$fc = \frac{\int_0^\infty f\,W(f)\,df}{\int_0^\infty W(f)\,df} \qquad (3)$$

where $W(f) = G(f)$, $f_C = f_G$ in the case of the signal with spectrum G(f) recorderd in the receiver point and $W(f) = S(f)$, $f_C = f$ in the case of the signal with spectrum S(f) defined in the source point. The integrals in the above relations were calculated by means of the procedure DSIMP1 from the library of scientific programs Fujitsu SSL II. The procedure realizes the integration of discrete function with the help of Simpson quadrature.

The input discrete models with velocity and Q factor distributions were interactively prepared using the *Interactive Velocity Editor* procedure from the ProMAX processing system. After exporting the data from the ProMAX system in ASCII format the conversion to binary files was realized. These binary files were used in the prepared Fortran programs.

2. The model parameters and results of spectrum centroidal frequency estimation

In the model calculations the near surface zone was taken into account with the maximum thickness of 300 m with a constant vertical velocity gradient. The gradient models with the anomalous zone with decreased velocity and Q factor values were considered too. Different values of the velocity gradient were taken into account. The value of Q factor was varying in relation to velocity: for the velocity of 1800 m/s the value of Q = 21 was assigned in accordance with publication data (Yadari *et al.* 2008). Additional values of Q = 24, 30, 40 were considered too. The values of Q for other values of velocities were linearly interpolated. The source signals with Gauss and rectangular spectrum were considered. The calculations were realized for the centroidal frequencies of source signal spectrum from the range 40 – 120 Hz.

The typical split spread of land seismics was used in the modeling: 120 receivers, receiver interval – 50 m, distance of shot point from the nearest receiver – 50 m. The calculations were performed for the left part of the spread, varying the position of the anomalous zone in relation to the shot point. The maximum offset – for the first receiver no. 1 – was 3000 m, minimum offset (receiver no. 60) – 50 m.

The discrete models of the velocity and factor Q fields were defined through the values defined in the nodes of the grid with dimensions 200 × 31 (200 nodes in the „x" direction, 31 nodes in the „z" direction) for the sampling intervals Δx = 20 m and Δz = 10 m. Some of the calculations for gradient models with an anomalous zone were realized for Δz = 1 m. The shot point and the receiver points were located at the level of $2\Delta z$ in accordance with the requirements of the estimation of velocity and factor Q values (and their spatial derivatives) by means of the finite differences.

2.1. The results of calculations for the gradient models without anomalous zones

The calculations for the gradient models without anomalous zones were performed changing the following model and source signal parameters:

– the range of variations of velocity and Q factor in the models with constant vertical gradient: model 1: V = 1,7 – 1,8 km/s, Q = 20 – 21; model 2: V = 1,7 – 1,75 km/s, Q = 20 – 20,4; model 3: V = 1,5 – 1,7 km/s, Q = 17,5 – 19,83; model 4: V = 1,7 – 1,9 km/s, Q = 20 – 22,2;

– different levels of Q factor variations: range 1: Q = 20 – 21, range 2: Q = 22,7 – 24, range 3: Q = 28,3 – 30, range 4: Q = 37,8 – 40,

– different types of the source signals: signals with Gauss or rectangular spectrum,

– different centroidal frequencies of the signal spectrum from the range: 40 – 100 Hz (in the case of rectangular spectrum $0 - f_{max}$ centroidal frequency $f_{centr} = f_{max} / 2$, in the case of Gauss signal $f_{centr} = f_0$).

The graphs of seismic ray trajectories of turning wave for the four discussed gradient models without anomalous zones are presented – with the location of shot (XSP) and receivers points – in the Figures 1 – 4. The analysis of these graphs lets us state that the maximum depth of a ray penetration considerably exceeds the most often occurring depths of the low velocity layer bottom reaching commonly the values from a dozen or so to tens of meters. That means that in the real conditions only part of the receivers located nearer to the shot point will record the turning wave in first breaks. The others will record the head wave originating at the bottom of low velocity layer. In the Figures 1–4 two horizontal lines are drawn at the depths of 50 and 100 m in relation to the level of shot/receivers. Taking into account these lines we can state that only for the model from Figure 2 (V = 1,7 – 1,75 km/s) the maximum depths of a ray penetration are within the range of 100 m. For other cases that condition is fulfilled for about 2/3 to 3/4 of the considered left part of the spread. Therefore in the analysis of the variations of the spectrum centroidal frequencies the most important are the receivers nearer to the shot in the range of their numbers from 30 to 60.

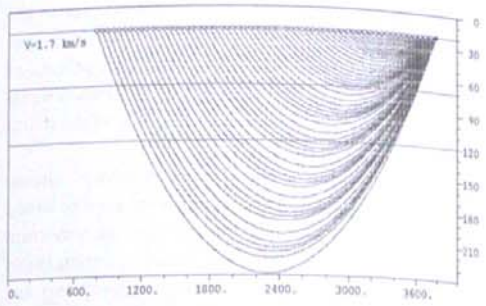


Fig. 1. The model of gradient medium with ray trajectories of turning wave, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21, vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m

Rys. 1. Model ośrodka gradientowego z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m

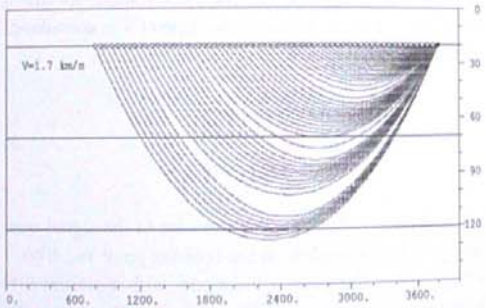


Fig. 2. The model of gradient medium with ray trajectories of turning wave, VMIN = 1,7 km/s, VMAX = 1,75 km/s, QMIN = 20, QMAX = 20,4; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m

Rys. 2. Model ośrodka gradientowego z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,75 km/s, QMIN = 20, QMAX = 20,4; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m

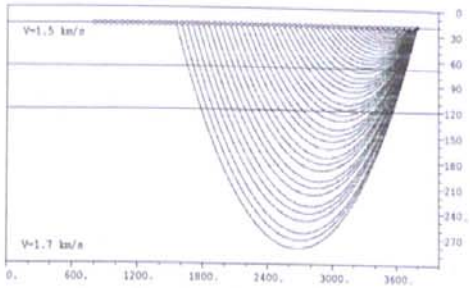


Fig. 3. The model of gradient medium with ray trajectories of turning wave, VMIN = 1,5 km/s, VMAX = 1,7 km/s, QMIN = 17.5, QMAX = 19.83; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m

Rys. 3. Model ośrodka gradientowego z trajektoriami promieni fali refragowanej, VMIN = 1,5 km/s, VMAX = 1,7 km/s, QMIN = 17.5, QMAX = 19.83; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m

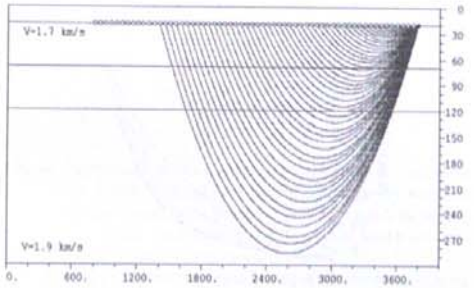


Fig. 4. The model of gradient medium with ray trajectories of turning wave, VMIN = 1,7 km/s, VMAX = 1,9 km/s, QMIN = 20, QMAX = 22.2; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m

Rys. 4. Model ośrodka gradientowego z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,9 km/s, QMIN = 20, QMAX = 22.2; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m

Figure 5 illustrates the rectangular spectrum of the source signal and the signal spectra in the receiver points no. 10 – 60 for the model from Figure 1. For the rectangular spectrum of the source signal the spectra of signals in the receiver points may be identified with the frequency characteristics of the medium in the frequency band of the source signal. Applications of the rectangular spectrum of the source signal means the simulation of the vibrator source. The variations of the spectra with distance are significant only in the range of receivers with numbers 30–60.

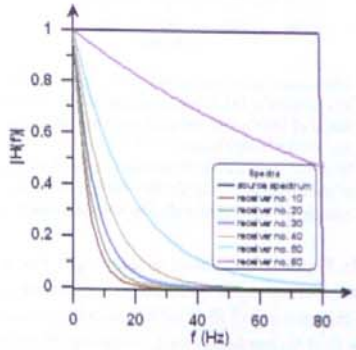


Fig. 5. The rectangular spectrum of the source signal and the signal spectra in the receisver points no. 10–60 (the distance of shot point from receiver no. 60 is 50 m) for the model from Figure 1 (VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21)

Rys. 5. Widmo prostokątne sygnału źródła oraz widma sygnałów w punktach odbioru nr 10–60 (punkt wzbudzenia odległy 50 m od odbiornika nr 60) dla modelu z fig. 1 (VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21)

Figure 6 illustrates the Gauss spectrum of the source signal and the signal spectra in the receiver points no. 30–60 for the model from Figure 1. The variations of the signal spectra with a distance from the shot point are very significant in this case. We can observe very strong decreasing of the spectrum amplitude and distinct decreasing of the frequency of the spectrum maximum.

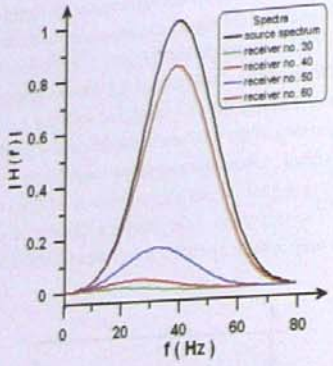


Fig. 6. The Gauss spectrum of the source signal and the signal spectra in the receiver points no. 30 – 60 (the distance of shot point from receiver no. 60 is 50 m) for the model from Figure 1

Rys. 6. Widmo Gaussa sygnału źródła oraz widma sygnałów w punktach odbioru nr 30 – 60 (punkt wzbudzenia odległy 50 m od odbiornika nr 60) dla modelu z fig. 1

In Figure 7 the graphs of centroidal frequencies f_{centr} of the receiver signal spectrum are presented as a function of the receiver number for the medium models from Figures 1 – 4 for the frequency f_{max} = 80 Hz of rectangular source signal spectrum. The significant variations of f_{centr} – in the range from 40 Hz to about 5 Hz – can be observed only for the receiver numbers from 30 to 60. So strong decreasing of f_{centr} values induced the author to introduce much greater values of Q factor (lower attenuation). The results of calculations for these new values of Q illustrates Figure 8.

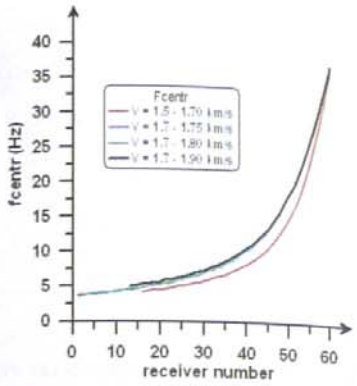


Fig. 7. The graphs of the spectrum centroidal frequencies of the receiver signal as a function of receiver number for the medium models from Figures 1 – 4 for the frequency f_{max} = 80 Hz of the rectangular source signal spectrum

Rys. 7. Wykresy częstotliwości centroidalnych widma sygnału odbiornika w funkcji nr odbiornika dla modeli ośrodka z figur 1 – 4 dla częstotliwości f_{max} = 80 Hz prostokątnego widma sygnału źródła

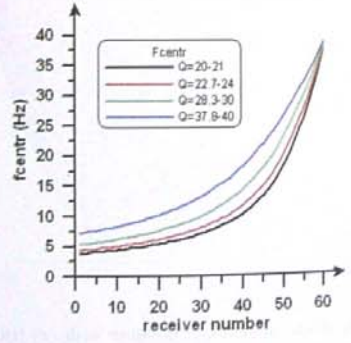


Fig. 8. The graphs of the spectrum centroidal frequencies of the receiver signal as a function of receiver number for the model from Figure 1 for the frequency f_{max} = 80 Hz of the rectangular source signal spectrum and for different ranges of Q factor values

Rys. 8. Wykres częstotliwości centroidalnych widma sygnału odbiornika w funkcji nr odbiornika dla modelu z fig. 1 dla częstotliwości f_{max} = 80 Hz prostokątnego widma sygnału źródła i dla różnych zakresów wartości współczynnika Q

The analysis of f_{centr} variations confirms a relatively not large change of f_{centr} decreasing with distance from the shot point. In the range of receiver numbers from 60 to 30 f_{centr} is decreasing from 40 Hz to about 5 – 10 hz for the values of Q factor from the range 20 – 40.

Figure 9 presents the graph of centroidal frequencies of the receiver signal spectrum as a function of the receiver number for the model from Figure 1 for different frequencies f_{max} rectangular spectrum (80 – 120 Hz) of the source signal for modified range of Q values (Q = 22,67 – 24). A significant influence of f_{max} frequency on the f_{centr} variations is visible only in the range of receiver numbers 50 – 60.

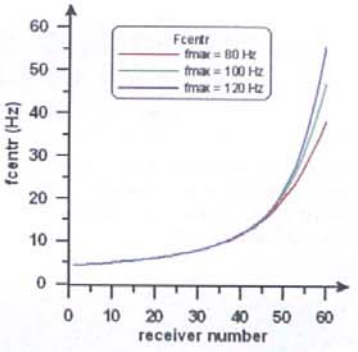


Fig. 9. The graphs of the spectrum centroidal frequencies of the receiver signal as a function of receiver number for the model from Figure 1 for different frequencies f_{max} of the rectangular source signal spectrum and modified range of Q factor values

Rys. 9. Wykres częstotliwości centroidalnych widma sygnału odbiornika w funkcji nr odbiornika dla modelu z fig. 1 dla różnych częstotliwości f_{max} prostokątnego widma sygnału źródła dla zmodyfikowanego zakresu wartości współczynnika Q

The succeeding calculations were performed for the source signal spectrum in the form of the Gauss function. That signal is commonly used in the amplitude tomography based on the inversion of f_{centr} variations. The spectra of the signals for values of the f_0 parameter (equation 2) from the range 40 – 100 Hz are shown in Figure 10, and the results of f_{centr} calculations are illustrated in Figure 11.

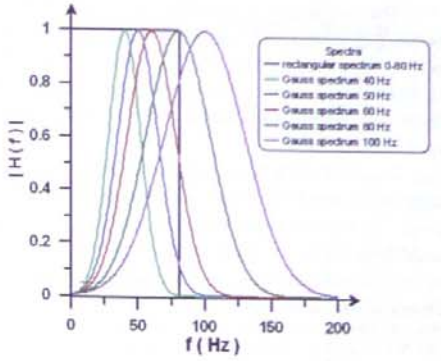


Fig. 10. The graphs of the source Gauss signal spectra for different peak frequencies f_0 against a background of the rectangular spectrum 0 – 80 Hz

Rys. 10. Wykresy widm sygnału Gaussa źródła dla różnych częstotliwości dominujących f_0 na tle widma prostokątnego 0 – 80 Hz

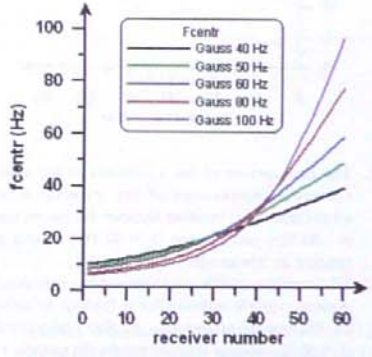


Fig. 11. The graphs of the spectrum centroidal frequencies of the receiver signal as a function of receiver number for the model from Figure 1 for different peak frequencies f_0 of the Gauss source signal spectrum and for a range of the quality factor values Q = 20 – 21

Rys. 11. Wykres częstotliwości centroidalnych widma sygnału odbiornika w funkcji nr odbiornika dla modelu z fig. 1 dla różnych częstotliwości dominujących f_0 widma Gaussa sygnału źródła dla zakresu wartości współczynnika Q = 20 – 21

The most significant variations of f_{centr} are visible in the range of the receiver numbers from 40 to 60. For the receiver no. 40 the value of f_{centr} is about 25 Hz. For the receivers much more distant from the source f_{centr} is slowly decreasing to the value of about 5 – 10 Hz for the receiver no. 1.

Figure 12 presents the comparison of the f_{centr} variations for the case of the rectangular source signal spectrum (f = 0 – 80 Hz) and Gauss spectrum (f_0 = 40 Hz) for the model from Figure 1. The differencies in f_{centr} variations are very distinct and much more regular (nearly linear) in the case of the Gauss signal.

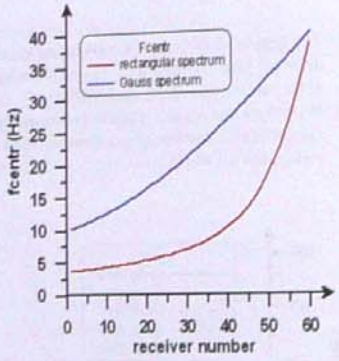


Fig. 12. The comparison of the variations of the spectrum centroidal frequencies of the receiver signal as a function of the receiver number for the rectangular (0 – 80 Hz) and Gauss (f_0 = 40 Hz) source signal spectra for the model from Figure 1

Rys. 12. Porównanie zmian częstotliwości centroidalnych widma sygnału odbiornika w funkcji nr odbiornika dla prostokątnego (0 – 80 Hz) i gaussowskiego (f_0 = 40 Hz) widma sygnału źródła dla modelu z fig. 1

2.2. The results of calculations for the gradient models with anomalous zone

In the case of calculations made for the gradient models with heterogeneities the anomalous zone with the decreased velocity ((V = 1200, 1400 m/s) and decreased Q factor (Q = 14, 16.33) has been taken into account. The anomalous zone had the shape of a rectangle of the dimensions: width – 500 m, height – 30 m, and it was located at the depth of 30m under the level of shot/receivers. The variation of the shot point location in relation to the anomalous zone was taken into calculations. Figure 13 illustrates the velocity model and Figure 14 – the model of the Q factor.

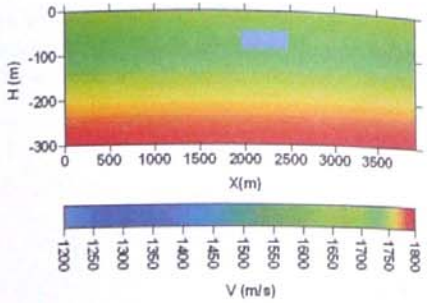


Fig. 13. The velocity model of a medium for a variant of the anomalous zone with a velocity V = 1200 m/s

Rys. 13. Model prędkościowy ośrodka dla wariantu anomalii z prędkością V = 1200 m/s

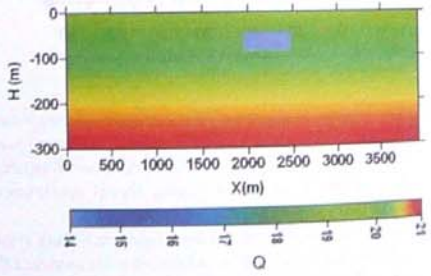


Fig. 14. The Q model of a medium for a variant of the anomalous zone with a velocity V = 1200 m/s

Rys. 14. Model Q ośrodka dla wariantu strefy anomalnej z prędkością V = 1200 m/s

Figures 15 – 18 present the graphs of ray trajectories of turning wave for the discussed models for vertical velocity sampling Δz = 10 m (Figures 15 – 16) and for decreased vertical velocity sampling Δz = 1 m (figures 17 – 18). The analysis of the seismic ray trajectories confirms their complicated shape resulting in great part from the specificity of the modeling using discrete velocity models. Their feature is a lack of sharp boundaries of heterogeneity zone which were drawn in figures only for the purpose of the zone localization. In discrete models the anomalous zone is surrounded with the band (its width is equal to the distance between the nodes of the grid) of strong velocity (and Q factor) gradient. This is the vertical gradient under and above the anomalous zone and horizontal gradient on both sides of that zone. We must remember that the ray trajectory is also determined by the second spatial derivatives of the velocity. As a result, the effect of „boundaries" of the anomalous zone ranges over the distance of double sampling interval of velocity causing the ray deviation at some distance from the anomalous zone. Such false trajectories should be excluded from the calculations of frequency characteristics of the medium. It seems that the best criterion of the exclusion may be a comparison of the traveltimes obtained in the process of ray tracing with the results of the first break picking applied to the theoretical records resulting from wave equation modeling.

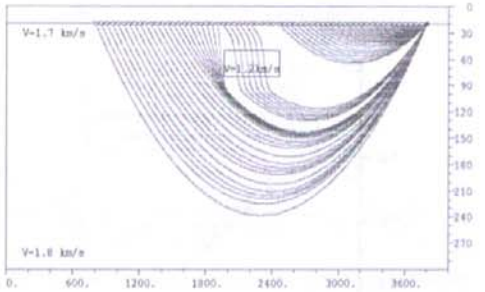


Fig. 15. The ray trajectories of the turning wave in a model of the gradient medium with anomalous zone V = 1,2 km/s, Q = 14; VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m, Δz = 10m

Rys. 15. Model ośrodka gradientowego z niejednorodnością V = 1,2 km/s, Q = 14 z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21, oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach; współrzędna źródła XPS = 3800 m, Δz = 10m

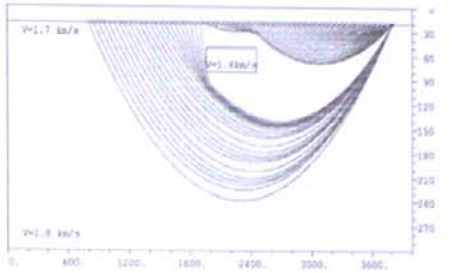


Fig. 16. The ray trajectories of the turning wave in a model of the gradient medium with anomalous zone V = 1,4 km/s, Q = 16,33; VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m, Δz = 10m

Rys. 16. Model ośrodka gradientowego z niejednorodnością V = 1,4 km/s, Q = 16,33 z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m, Δz = 10m

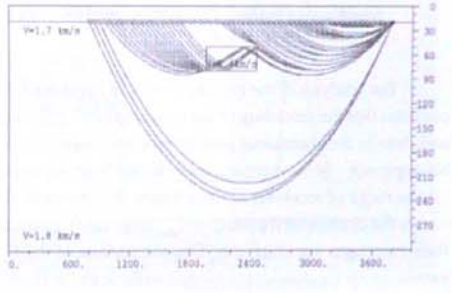


Fig. 17. The ray trajectories of the turning wave in a model of the gradient medium with anomalous zone V = 1,4 km/s, Q = 16,33; VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 3800 m, decreased sampling interval Δz = 1 m

Rys. 17. Model ośrodka gradientowego z niejednorodnością V = 1,4 km/s, Q = 16,33 z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 3800 m, zmniejszony krok próbkowania Δz = 1 m

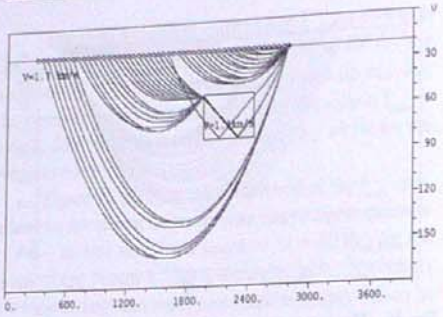


Fig. 18. The ray trajectories of the turning wave in a model of the gradient medium with anomalous zone V=1,4 km/s, Q = 16,33; VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; vertical axis – depth in meters, horizontal axis – distance in meters, source coordinate XSP = 2800 m, decreased sampling interval Δz = 1 m

Rys. 18. Model ośrodka gradientowego z niejednorodnością V = 1,4 km/s, Q = 16,33 z trajektoriami promieni fali refragowanej, VMIN = 1,7 km/s, VMAX = 1,8 km/s, QMIN = 20, QMAX = 21; oś pionowa – głębokość w metrach, oś pozioma – odległość w metrach, współrzędna źródła XPS = 2800 m, zmniejszony krok próbkowania Δz = 1 m

The analysis of the ray trajectories in Figures 15–18 confirms that the recording of the rays propagating through and close to the anomalous zone is expected – applying the ray approach – in the receivers more distant from the source (in the range of receivers number below 30). At these distances the centroidal frequencies f_{centr} have low values and their variations are small too. The comparison of the f_{centr} variations for the considered models without and with a/the anomalous zone is presented for a smaller distance of the source point from that zone in Figure 19. As we could expect the f_{centr} variations appear at long distance receivers and they are relatively small (several Hz). Additionally, the regular behavior of the hodograph in the case of the anomalous zone (Figure 20) does not reflect the complicated ray trajectories from Figure 18.

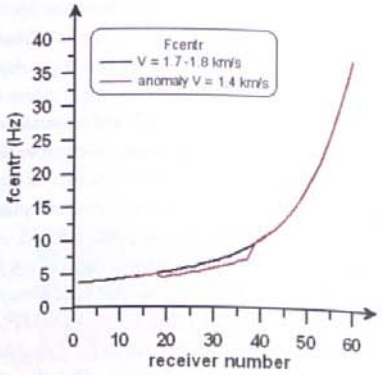


Fig. 19. The graphs of f_{centr} for a model from Figure 18 in the case of the medium without anomalous zone (black line) and with anomalous velocity zone V = 1.4 km/s (red line) for the source coordinate XSP = 2800 m, decreased sampling interval Δz = 1 m

Rys. 19. Wykres f_{centr} dla modelu z rys. 18 dla wariantu ośrodka bez anomalii (linia czarna) i z anomalią prędkościową V = 1,4 km/s (linia czerwona) przy współrzędnej źródła XPS = 2800 m, zmniejszony krok próbkowania Δz = 1 m

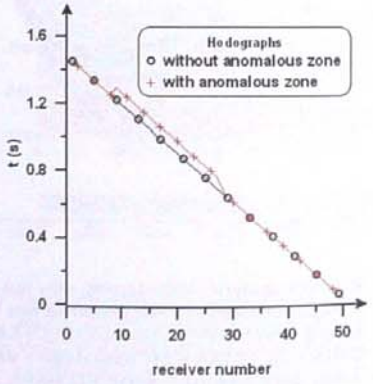


Fig. 20. Comparison of hodographs for the case of a model from Figure 19 with and without anomalous zone

Rys. 20. Porównanie hodografów dla przypadku modelu z rys. 19 ze strefą anomalną lub bez niej

From the above analysis it can be expected that only a very shallow anomalous zone located near the source may bring an essential change of centroidal frequencies. Therefore the last examined model consisted of a shallow gradient layer (thickness 10 m) with very low velocities (varying linearly from 800 to 1700 m/s) and small values of Q factor (varying from 9,33 to 19,83) introduced to the gradient model from Figure 1. The anomalous zone with decreased velocity and Q factor (V = 800, 1000, 1400 m/s; Q = 9,33; 11,66; 16.33) was located at a depth of 15 m (width 500 m). The ray trajectories for the discussed model and anomalous zone with a velocity of 1.4 km/s are presented in Figure 21. In Figure 22 the graphs of the centroidal frequences are shown for three variants of anomalous zone velocities. The differences between f_{centr} for models with and without anomaly are illustrated in Figure 23. As we can see from these figures the anomaly of f_{centr} is located in the range of strong variations of centroidal frequencies but the anomaly of f_{centr} does not exceed a value of 1 Hz.

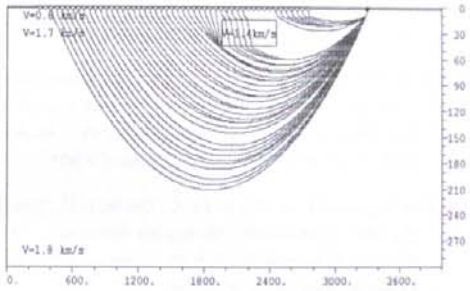


Fig. 21. The ray trajectories of the turning wave for the model from Figure 18 after decreasing the depth of anomalous zone and introducing additional near-surface low velocity thin gradient layer, XSP = 3300 m/s

Rys. 21. Trajektorie promieni fali refragowanej dla modelu z rys. 19 po zmniejszeniu głębokości strefy anomalnej i wprowadzeniu dodatkowej nisko-prędkościowej przypowierzchniowej cienkiej warstwy gradientowej, XSP = 3300 m/s

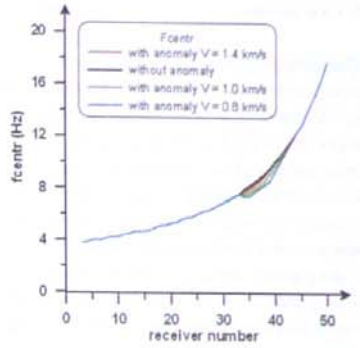


Fig. 22. The f_{centr} graphs for the model from Figure 21 for three variants of anomalous zone velocities

Rys. 22. Wykresy f_{centr} dla modelu z rys. 21 dla trzech wariantów prędkości strefy anomalnej

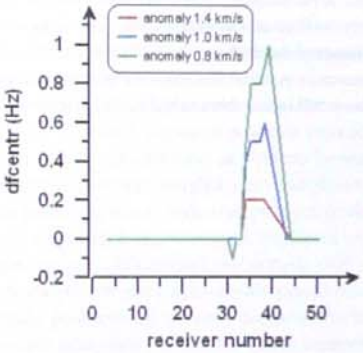


Fig. 23. The differences between f_{centr} graph for the model from Figure 21 without anomalous zone and f_{centr} graphs from Figure 22 for three variants of anomalous zone velocity and Q factor

Rys. 23. Różnice między wykresem f_{centr} dla modelu z fig. 21 bez strefy anomalnej i wykresami f_{centr} z rys. 22 dla trzech wariantów prędkości i współczynika Q strefy anomalnej

3. Conclusions

The analysis of the graphs of seismic ray trajectories and the variations of the spectrum centroidal frequencies with distance from shot point lets us formulate several conclusions defining the possibilities of the application – for the considered models – the variations of f_{centr} in near surface zone for the inversion in amplitude tomography:

– the strong variations of f_{centr} are observed only to the distance of about 1000 m from the source,

– complicated and very often oscillating character of seismic ray trajectories, observed in the case of gradient models with the/a anomalous zone, makes it necessary to select – before determining the frequency characteristics of the medium attenuation – the rays taking into account the traveltimes resulting from modeling based on wave approach,

– the variations of f_{centr} of broadband signals in medium with lower values of velocity and factor Q may be expected in the range of first breaks only on the receiver positions located at a distance from several hundred to about thousand meters from the shot point for shallow and broad anomalous zones but these variations are expected to be too small to use them in tomographic inversion.

Bibliography

Best A. I., Sothcott J., McCann C., 2007, *A laboratory study of seismic velocity and attenuation anisotropy in near-surface sedimentary rocks*. Geophysical Prospecting, 55, 5, 609–625.

Carcione, J. M., 2000, *A model for seismic velocity and attenuation in petroleum source rocks*. Geophysics, 65, 1080–1092.

Chichinina T., Sabinin V., Ronquillo-Jarillo G., 2006, *QVOA analysis: P-wave attenuation anisotropy for fracture characterization*.Geophysics, 71, C37–C48.

Kasina Z., 2001, *Seismic Tomography*. Ed. Inst. GSMiE PAN, Cracow (*in Polish*).

Kasina Z., 2008a, *The analysis of the influence of seismogeological model and signal parameters on the variations of centroidal signal spectrum frequencies in well to well measurements – model study*. Geologia, t. 34, 2, 231–254 *(in Polish)*.

Kasina Z., 2008b, *The analysis of amplitude tomography based on the variations of centroidal signal spectrum frequency in well to well measurements – model study*. Geologia, t. 34, 2, 255–284 (*in Polish*).

Liu E., Chapman M., Varela I., Li X., Queen J. H., Lynn H., 2007, *Velocity and attenuation anisotropy: Implication of seismic fracture characterizations*. The Leading Edge, 26, 1170–1174.

Parra J. O., 2000, *Poroelastic model to relate seismic wave attenuation and dispersion to permeability anisotropy*. Geophysics, 65, 202–210.

Payne S., Worthington M., Odling N., West L., 2007, *Estimating permeability from field measurements of seismic attenuation in fractured chalk*. Geophysical Prospecting, 55, 643–653.

Quan Y., Harris J. M., 1997, *Seismic attenuation tomography*. Geophysics, 62, 895–905.

Rossi G., Gei D., Bohm G., Madrussani G., Carcione J., 2007, *Attenuation tomography: An application to gas-hydrate and free-gas detection*. Geophysical Prospecting, 2007, 55, 655–669.

Yadari N. E., Ernst F., Mulder W., 2008, *Near-surface attenuation estimation using wave-propagation modeling*. Geophysics, 73, no. 6, U27–U37.

Zhu Y., Tsvankin I., Dewangan P., and Wijk K. V., 2007, *Physical modeling and analysis of P-wave attenuation anisotropy in transversely isotropic media*. Geophysics, 72, no. 1, D1–D7.

Streszczenie

W pracy przedstawiono wyniki obliczeń trajektorii i czasów przebiegu promieni sejsmicznych fali refragowanej w dyskretnym ośrodku gradientowym strefy przypowierzchniowej. Trajektorie promieni (figury 1–4) określane były na drodze rozwiązywania układu równań różniczkowych, wynikających z zasady Fermata. Wyniki tych obliczeń były podstawą określenia charakterystyk tłumienia ośrodka, odpowiadających poszczególnym promieniom oraz widm założonych sygnałów źródła w punktach odbioru wraz z częstotliwościami centroidalnymi tych widm (figury 5–12). W obliczeniach uwzględniono również modele gradientowe z anomalnymi strefami o obniżonej wartości prędkości i tłumienia (figury 13–23). Dokonano analizy rozkładu trajektorii, zmian widm i ich częstotliwości centroidalnych dla różnych parametrów sygnału źródła i różnych parametrów modelu sejsmogeologicznego przypowierzchniowej strefy ośrodka (figury 19–20, 22–23). W obliczeniach uwzględniono schemat obserwacji sejsmiki powierzchniowej. Wyniki obliczeń pozwoliły stwierdzić, że silne zmiany częstotliwości centroidalnej sygnału propagującego w strefie przypowierzchniowej są obserwowane jedynie w zakresie odległości do około 1000 m od źródła. Złożony charakter trajektorii promieni sejsmicznych w strefie przypowierzchniowej ze strefami anomalnymi potwierdził konieczność uwzględnienia podejścia falowego przy selekcji promieni, wykorzystywanych do obliczeń częstotliwościowych charakterystyk tłumienia ośrodka. Uwzględnienie w obliczeniach anomalnych stref o obniżonych wartościach prędkości i współczynnika dobroci Q w obrębie gradientowej warstwy przypowierzchniowej dostarczyło relatywnie niewielkich zmian częstotliwości centroidalnej szerokopasmowego sygnału sejsmicznego.

ZBIGNIEW KASINA[1]

THE DIRECTIONS OF CURRENT APPLICATIONS OF TRANSFORMS TO SEISMIC DATA PROCESSING AND ANALYSIS[2]

Key words:

geophysics, seismic methods, transforms, seismic data processing and analysis

Abstract

In the presented paper the directions of current applications of the mathematical transforms to seismic data processing and analysis have been described. In the analysis the following transforms were taken into consideration: Gabor transform, Karhunen-Loeve transform (K-L Transform), wavelet transform (WT transform), curvelet transform. The theoretical background of the analyzed transforms were presented and their main applications relating to coherent and random noise attenuation in the seismic data processing. Additionally, the applications were described relating to seismic sections analysis based on the instantaneous spectral analysis (ISA) as well as the applications used in the image analysis. The comparison of the results of different kinds of spectral decompositions were presented using selected model and field data.

KIERUNKI NAJNOWSZYCH ZASTOSOWAŃ TRANSFORMAT W PRZETWARZANIU I ANALIZIE DANYCH SEJSMICZNYCH

Słowa kluczowe:

geofizyka, metody sejsmiczne, transformaty, przetwarzanie i analiza danych sejsmicznych

Abstrakt

W pracy przedstawiono kierunki najnowszych zastosowań transformat matematycznych w przetwarzaniu i analizie danych sejsmicznych. W analizie uwzględniono transformację Gabora, transformację Karhunena-Loevego (*K-L Transform*), transformację falkową (*Wavelet Transform*) oraz transformację krzywkową (*Curvelet Transform*). Przedstawiono podstawy teoretyczne tych transformacji oraz ich główne zastosowania, związane z usuwaniem zakłóceń koherentnych i przypadko-

[1] University of Science and Technology AGH, Faculty of Geology, Geophysics & Environment Protection, Department of Geophysics, Cracow, Poland.

[2] The paper was prepared within the project of MNiSW no. N N525 168735 and statutory works of MNiSW in the Department of Geophysics contract AGH no. 1111.140.06.

wych w procesie przetwarzania danych sejsmicznych. Opisano także zastosowania związane z analizą sekcji sejsmicznych opartą na chwilowej analizie spektralnej (*ISA*) oraz zastosowania związane z analizą obrazów. Przedstawiono porównanie wyników różnych typów dekompozycji spektralnej na wybranych danych modelowych i polowych.

Introduction

Mathematical transforms are the most important tools of seismic data processing creating the theoretical background of many processes used in seismic processing systems. The most popular and commonly used ones are the Fast Fourier Transform applied in 1-D processing and in 2-D processing as the so-called F-K filtering, Radon Transforms used in the linear version (the so-called Slant Stack) or in hiperbolic (parabolic) versions (Kasina 1998a, b). Generally, the mentioned transforms are used mainly to improve the signal to random noise ratio or to attenuate the coherent noise (surface waves, multiples).

In the last years the new transforms appeared in the theory of seismic data processing creating a background for new system processing procedures. Among them the most promising are the wavelet transform and the curvelet transform. The theory of the wavelet transform is based on the well known Gabor transform and the theory of curvelet transform is quite new. The way of the application of wavelet and curvelet transform is very similar to Karhunen-Loeve Transform (K-L Transform), well known for many years but still used now in new applications (Bitri & Grandjean 2004, Kritski et al. 2007).

1. Gabor Transform

Gabor stated in 1946 (Gabor 1946) that any signal $\varphi(t)$ can be expressed as a superposition of shifted and modulated versions of the elementary signal $g(t)$:

$$\varphi(t) = \sum_m \sum_k a_{mk}\, g(t - m\alpha T) e^{jk\beta\Omega t} \qquad (1)$$

where m and k are integer values, a_{mk} are properly chosen coefficients, $g(t-m\alpha T)$ is a shifted version of $g(t)$ with a time shift αT, $\beta\Omega$ is the frequency shift. Time shift and frequency shift satisfy the relationships:

$$\Omega T = 2\pi \qquad \alpha\beta \le 1$$

In the special case $\alpha\beta = 1$ it can be showed (Bastiaans 1980) that the expansion coefficients a_{mk} may be defined using so-called Gabor transform:

$$a_{mk} = \int \varphi(t) w^*(t - m\alpha T) e^{-jk\beta\Omega t} dt \qquad (2)$$

where $w(t)$ is the window function. The discrete Gabor transform may be expressed in the form:

$$s(i) = \sum_{m=0}^{\infty} \sum_{n=0}^{N-1} C_{m,n} h(i - m\Delta M) \exp(j2\pi in/N)$$
$$C_{m,n} = \sum_{i=0}^{\infty} s(i) \gamma^*(i - m\Delta M) \exp(-j2\pi in/N) \qquad (3)$$

where $s(i)$ is time series, $C_{m,n}$ are Gabor coefficients, m is time index, n is frequency index, $h(i)$ is synthesis window, $\gamma^*(i)$ is analysis window, N is a number of frequency indices, ΔM and ΔN are sampling intervals in time and frequency domains.

The Gabor transform was applied to the theory of seismic signal propagation as well as to the theory and practice of seismic data processing. Morlet et al. (1982a, b) used the zerophase wavelets modulated by the Gaussian envelope and corresponding complex wavelets as the basic wavelets in the Gabor expansion of arbitrary signal in 2-D time and frequency domain. The complex functions defined in this way were used to describe the process of seismic signal propagation and to process the seismic signals. The Gabor series may be treated as the tool to define the discrete spectra of instantaneous frequency and to high resolution processing with preservation an information about the signal phase.

Margrave and Lamoureux (2002) presented a new approach to nonstationary seismic deconvolution based on Gabor transform. The proposed deconvolution made it possible to define the Gabor transform of seismic reflectivity by the division of the Gabor spectrum of nonstationary seismic trace with the estimate of source signal and Q characteristic of the absorbing medium. In the continuous Gabor transform of a signal s(t) defined in the form:

$$\bar{s}(\tau, f) = \int_{-\infty}^{\infty} s(t) g(t - \tau) e^{-2\pi i f t} dt \qquad (4)$$

where $g(t)$ is the Gabor analysis window and parameter τ defines the location of a window center, the function $g(t)$ had the form of Gaussian function.

The next application of Gabor transform in the process of prestack depth migration was presented by Yongwang i Margrave (2006). The proposed Gabor depth imaging algorithm approximates wavefield extrapolation in the variant of generalized phase-shift-plus-interpolation (GPSPI).

The application of Gabor transform together with neural networks for the purpose of an automatic phase correlation was described by Glinsky et al. (2001). The wavelet Gabor transform was used here to provide magnitude and phase information about the events at a variety of resolutions (scales), orientations (rotational angles), and frequencies. A variety of elliptical Gabor kernels were applied to/in the form:

$$K(t, \chi, t', \chi') = \exp\left[\frac{-T^2}{2\delta_T^2} - \frac{S^2}{2\delta_S^2}\right] [\cos 2\pi f T + i \sin 2\pi f T]$$

where

$$T \equiv (t - t')\cos\theta + \frac{s - s'}{\Delta s/\Delta t}\sin\theta$$

$$S \equiv (s - s')\cos\theta - \frac{t - t'}{\Delta t/\Delta s}\sin\theta$$

$$tg\theta \equiv \frac{dt/ds}{\Delta t/\Delta s}$$

where f is the frequency of the Gabor kernel, σ_τ is the time width, σ_s is the offset width, ds/dt is the orientation, Δt is the time sampling interval, Δs is the offset spacing, s is the offset, t is time, and x is the subpoint. The amplitude of the Gabor transform indicates where the events should be looked for, the phase indicates where they should be picked.

2. Wavelet Transform

In the process of Fourier transform the signal is transformed from the amplitude-time domain to the amplitude-frequency domain breaking down a signal into constituent sinusoids of different frequencies (Fig. 1).

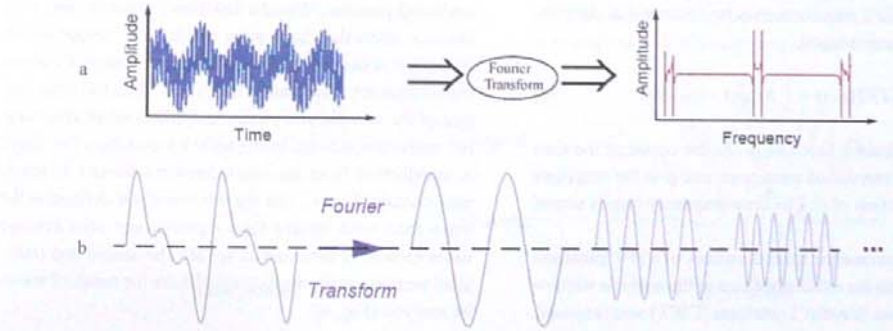


Fig. 1. The most important features of Fourier transform: (a) transformation from the amplitude-time domain to amplitude-frequency domain, (b) breaking down signal into constituents
Rys. 1. Najważniejsze cechy transformacji Fouriera: a) przejście z domeny amplituda-czas do domeny amplituda-częstotliwość, b) rozkład sygnału na składowe harmoniki

The most important drawback of Fourier analysis is loosing the time information during the transformation to the frequency domain. As a result we cannot tell *when* a particular event took place. To overcome this problem – a very serious one in the case of the nonstationary signals – Dennis Gabor (1946) adapted the Fourier transform to analyze only a small section of the signal at a time (Fig. 2).

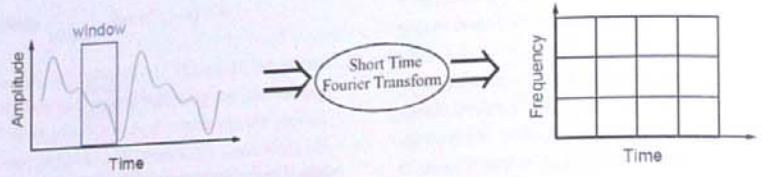


Fig. 2. The transformation from amplitude-time domain to frequency-time domain in STFT transform (Short-Time Fourier Transform)
Rys. 2. Przejście z domeny amplituda-czas do domeny częstotliwość-czas w transformacji STFT (*Short-Time Fourier Transform*)

This technique is called windowing the signal. Gabor's adaptation, named the Short-Time Fourier Transform (STFT), maps a signal into a two-dimensional function of time and frequency. STFT provides some information about both when and at what frequencies a signal event occurs. However, the information is obtained with limited precision determined by the size of the sliding analysis window. The drawback of STFT is that the analysis window is the same for all frequencies.

The STFT transform may be described in the form of the following integral:

$$STFT(\omega,\tau) = \int_{-\infty}^{\infty} f(t)\bar{\phi}(t-\tau)e^{-i\omega t}dt \qquad (5)$$

where the window function ϕ has the center at the time $t = \tau$, τ is the translation parameter, and $\bar{\phi}$ is the conjugate complex function of ϕ. The time frequency map is named spectrogram.

To overcome the main drawback of STFT transform connected with the established size of the analysis window the Continuous Wavelet Transform (CWT) was proposed. The CWT is defined as the sum over all time of the signal multiplied by scaled, shifted versions of the oryginal (or *mother*) wavelet function:

$$C(\text{scale, position}) = \int_{-\infty}^{\infty} f(t)\,\psi(\text{scale, position, } t)dt$$

where $\psi(t)$ is the *mother* wavelet playing the role ofan analysis window of varying size. The results of the CWT are many wavelet coefficients C, which are a function of scale and position. Wavelet analysis applies to long time analysis intervals where more precise low-frequency information is needed, and shorter analysis intervals where high-frequency information is desired. Then the main feature of the wavelet analysis is an application of size varying analysis windows. In the wavelet transform the signal is transformed from the amplitude-time domain to scale-time domain (Fig. 3). The *mother wavelet* is defined as the wave form with limited time duration and zero average value located in time and in space. The scaled and translated versions of the mother wavelet are the result of wavelet analysis (Fig. 4).

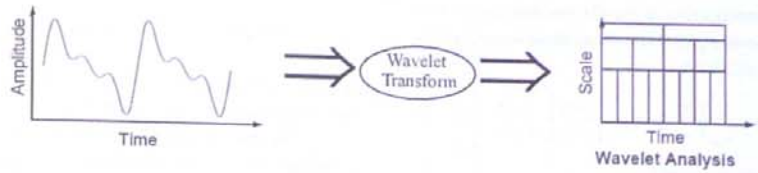


Fig. 3. The transformation from amplitude-time domain to scale-time domain in wavelet transform (WT)
Rys. 3. Przejście z domeny amplituda-czas do domeny skala-czas w transformacji falkowej (WT)

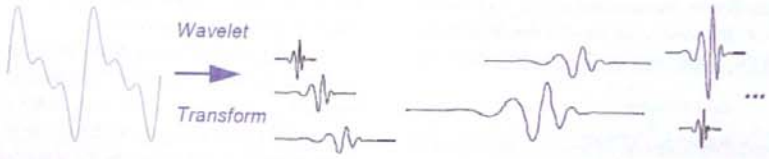


Fig. 4. Properly scaled and translated wavelets as the results of wavelet analysis
Rys. 4. Właściwie przeskalowane i przesunięte fale elementarne jako wynik analizy falkowej

The set of wavelets is generated by stretching and the translation of wavelet $\psi(t)$ (Sinha et al. 2005):

$$\psi_{\sigma,\tau}(t) = \frac{1}{\sqrt{\sigma}}\,\psi\left[\frac{t-\tau}{\sigma}\right]$$

where $\psi(t) \in L^2(\Re)$, σ i $\tau \in \Re$, parameter σ is called dilatation parametr or scale. Then CWT transform may be presented in the form:

$$F_w(\sigma,\tau) = \int_{-\infty}^{\infty} f(t)\,\frac{1}{\sqrt{\sigma}}\,\bar{\psi}\left[\frac{t-\tau}{\sigma}\right]dt \qquad (6)$$

where $F_w(\sigma, \tau)$ is the time-scale map (*scalogram*). There is a correspondence between wavelet scales and frequency (Fig. 5). The low scale and compressed wavelet is used to analyze high frequency rapidly changing details of the signals. The high scale and stretched wavelet is used to analyze low frequency slowly changing details of the signals.

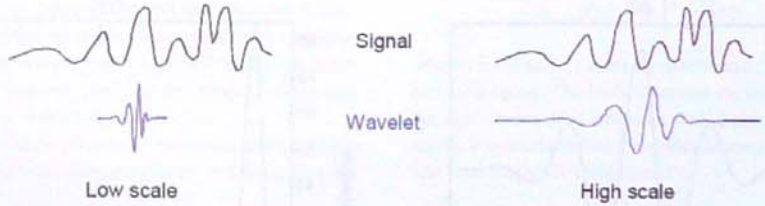


Fig. 5. The relation between the degree of wavelet stretching and scale value in CWT transform
Rys. 5. Relacja między stopniem rozciągnięcia waweletu i wartością skali w transformacji CWT

The reconstruction of the $f(t)$ function based on its wavelet transform $F_w(\sigma, \tau)$ may be realized using the relation (Sinha et al. 2005):

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_w(\sigma,\tau) \psi\left[\frac{t-\tau}{\sigma}\right] \frac{d\sigma\, d\tau}{\sigma^2 \sqrt{\sigma}} \quad (7)$$

where C_ψ satisfy the condition:

$$C_\psi = 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty$$

and $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(t)$, C_ψ is constant for wavelet ψ. In the continuous wavelet transform commonly used is the *Morlet wavelet* (Fig. 6) defined by the relation:

$$\psi_0(t) = \pi^{-1/4} e^{i\omega_0 t} e^{-t^2/2} \quad (8)$$

where ω_0 is the frequency. The center frequency of the Morlet wavelet is inversely proportional to the scale.

Sinha et al. (2005) proposed new approach named Time Frequency Continuous Wavelet Transform (TFCWT). Using the Fourier transform of $f(t)$ function in the form:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

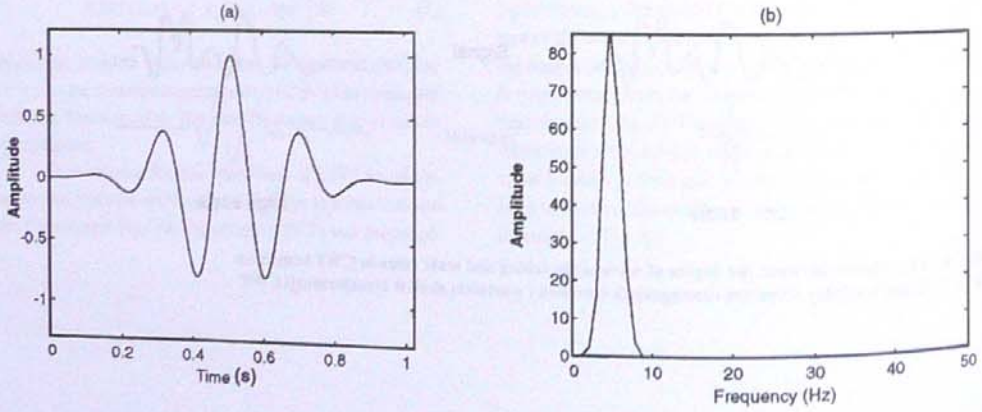
they introduced function $f(t)$ from relation (7) obtaining after some transforms the basic relation of TFCWT transform:

$$\hat{f}(\omega,\tau) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} F_w(\sigma,\tau) \hat{\psi}(\sigma\omega) e^{-i\omega\tau} \frac{d\sigma}{\sigma^{3/2}} \quad (9)$$

where $\hat{\psi}(\omega)$ is the complex conjugate of ψ and $F_w(\sigma, \tau)$ is the time-scale map (i.e. the scalogram).

The main feature of TFCWT is high-frequency resolution and low-time resolution at low frequencies and low-frequency resolution and high time resolution at high frequencies. Thus the discussed transform can improve the resolution of a nonstationary signal avoiding the subjective choice of window length necessary for the case of STFT (Short Time Fourier Transform).

In the CWT (7) the window function $\psi(t)$ is called the kernel wavelet, σ is called the scale and in practice is replaced by 2^j, where j is the scale index. For the small values of the scale index the width of the frequency band of the function is not large. As the scale index increases the function frequency band increases and the time width of the window decreases making it possible to identify the individual spikes. The function $\psi(t)$ should be absolutely integrable and square integrable being band limited and having zero mean (Chakraborty, Okaya 1995). An example of such a function is a commonly used modulated Gaussian defined as Morlet wavelet (relation (8), Fig. 6). The set of wavelets used in wavelet transform is very broad. It contains Daubechies wavelets (notation dbN where N is the wavelet order), biortogonal wavelets (notation bior), Coifman wavelets (called Coiflets), Symlets, Meyer wavelet.



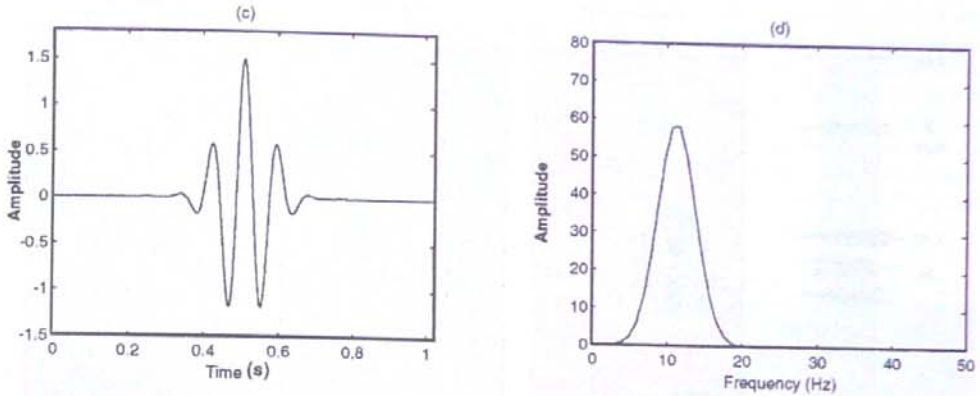


Fig. 6. Morlet wavelet: for scale index 0 (a) and its spectrum (b), for scale index 2 (c) and its spectrum (d) (after Chakraborty & Okaya 1995)

Rys. 6. Fala elementarna Morleta: dla indeksu skali 0 (a) oraz jej widmo (b), dla indeksu skali 2 (c) oraz jej widmo (wg Chakraborty & Okaya 1995)

In the continuous wavelet transform the function is decomposed using band-pass filtering the signal at different bandwidths while in the discrete wavelet transform (Discrete Wavelet Transform – DWT) 2-D scale-transformation space is implemented using *quadrature mirror filter* (*QMF*) (Chakraborty, Okaya 1995). QMF consists of two filters: high cut and low cut. The original signal is filtered by a halfband low-pass and a half-band high-pass filter followed by a down sampling by a factor of two. In the next step the output of the low-pass filter is then filtered again using the same two filters. The process of filtering is continued until the desired level of decomposition is obtained. Such a process is equivalent to a CWT with scale index of the kernel wavelet changing by integral values (the so-called diadic scales and positions).

The resolution of wavelet transform is not uniform across the entire time-frequency plane. WT has good time resolution for high frequencies (and therefore poor frequency resolution) and good frequency resolution for low frequencies. The problem of good resolution at the intermediate frequencies is still unresolved. To solve this problem Mallat and Zhang (1993) proposed *Matching Pursuit Decomposition (MPD)*. In the matching pursuit decomposition (MPD), a set of basis functions are generated by scaling, translating, and modulating a single window function as (Chakraborty, Okaya 1995):

$$\psi_{(s,\tau,\xi)}(t) = \frac{1}{\sqrt{s}} \psi\left[\frac{t-\tau}{s}\right] e^{i\xi t} \quad (10)$$

where s is the scale, τ is the translation and ξ is the frequency modulation. The basis functions are called "timefrequency" atoms. If is Gaussian then ξ are called Gabor atoms. It was established that Gabor atoms provide excellent time-frequency resolution (Fig. 7).

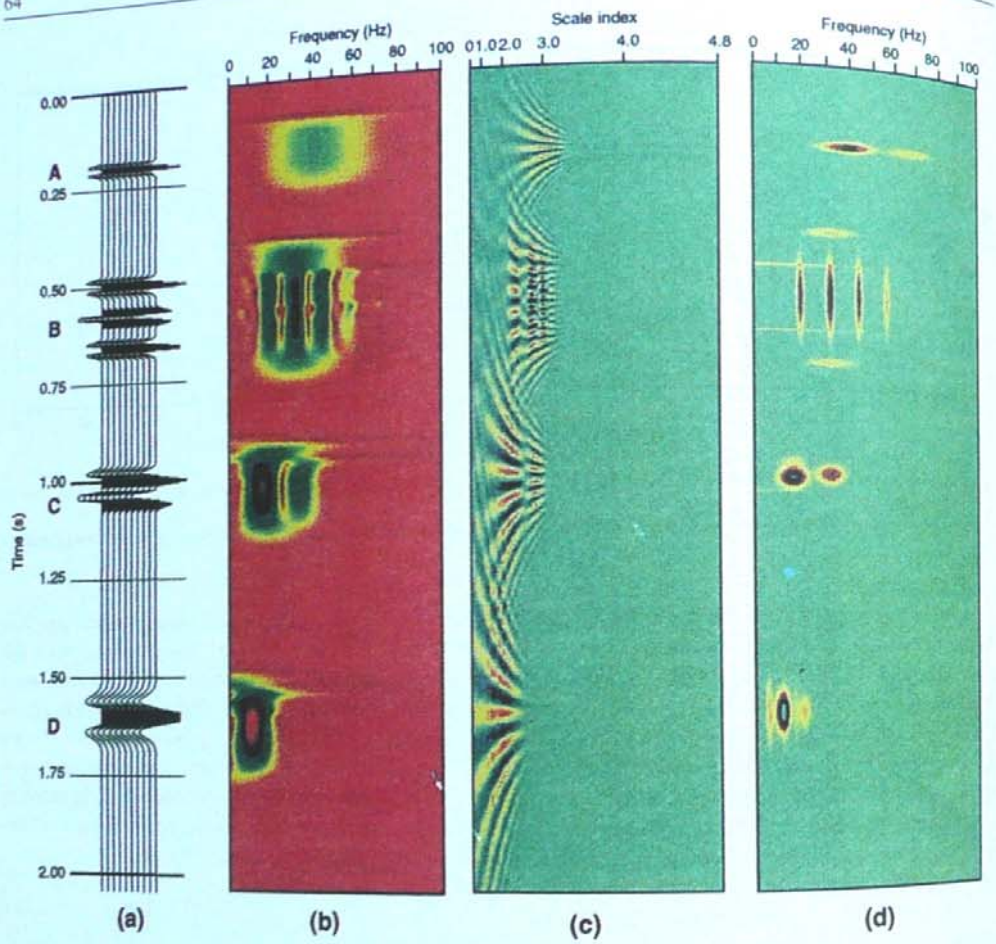


Fig. 7. A comparison of different kinds of synthetic seismogram decomposition: (a) synthetic seismogram, (b) seismogram decomposition using Short Time Fourier Transform (STFT), (c) seismogram decomposition using CWT presented as a scale index-translation plot, (d) seismogram decomposition using MPD presented as an frequency-time energy distribution (after Chakraborty & Okaya 1995)

Rys. 7. Porównanie różnych rodzajów dekompozycji sejsmogramu syntetycznego: (a) sejsmogram syntetyczny, (b) dekompozycja sejsmogramu z zastosowaniem *Short Time Fourier Transform* (STFT), (c) dekompozycja sejsmogramu z zastosowaniem CWT przedstawiona jako wykres indeks skali-translacja, (d) dekompozycja sejsmogramu z zastosowaniem MPD przedstawiona jako czasowo-częstotliwościowy rozkład energii (wg Chakraborty & Okaya 1995)

The important application of wavelet transform is Instantaneous Spectral Analysis (ISA). This is the technique of a continuous time-frequency analysis providing a frequency spectrum for each time sample of the signals which in the case of seismic method create a seismic trace. The method applied to seismic data involves the following steps (Castagna et al. 2003):

– decomposition of the seismogram into constituent wavelets using wavelet transform methods such as MPD,
– summing the Fourier spectra of the individual wavelets in the time-frequency domain to generate a so-called frequency gathers,
– sorting the frequency gathers to provide common (constant) frequency cubes, sections, time slices, and horizon slices.

An application of the discussed method considerably improves the spectral analysis resolution, identification of thin layers and defining direct hydrocarbon indicators (DHI). The possibilities of the technique shows Fig. 8, containing the synthetic seismic trace with constituent signals and results of time-frequency analysis of the trace. The resulted graph shows amplitude spectra of each time sample and is often called *frequency gather*. The constituent signals creating the complicated interference wave pattern in the time domain are easy to identify on the map of time-frequency analysis resulting from ISA.

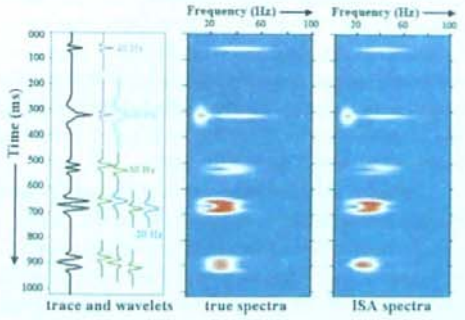


Fig. 8. A comparison of the true spectra with the ISA spectra for the case of synthetic trace: the trace and its constituent wavelets (on the left), true spectra (in the middle), ISA spectra (on the right) (after Castagna et al. 2003)

Rys. 8. Porównanie prawdziwych widm z widmami *ISA* dla trasy syntetycznej: trasa i jej składowe fale elementarne (z lewej), prawdziwe widma (w środku), widma z *ISA* (po prawej (wg Castagna et al. 2003)

A comparison of the ISA spectra with FFT spectra (the time analysis window 200 ms was used) are presented – for the discussed model data – in Fig. 9 and the comparison with Discrete Fourier Transform spectra are contained in Fig. 10. The main feature of FFT spectra is the lack of time resolution. The vertical notches indicate the interference wave pattern without possibility to identify the individual reflectors. The DFT spectra have at first look the resolution comparable with ISA spectra. However, a careful analysis confirms the presence of false indications at low frequencies, the loss of frequency resolution and lateral stretching of frequency bands.

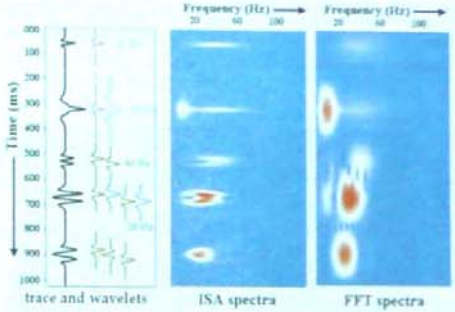


Fig. 9. A comparison of ISA spectra with FFT spectra obtained for 200 ms window (after Castagna et al. 2003)

Rys. 9. Porównanie widm ISA z widmami FFT określonymi w oknie 200 ms (wg Castagna et al. 2003)

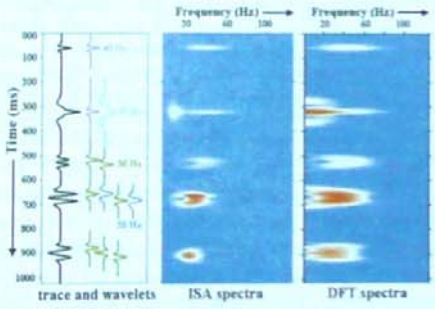


Fig. 10. A comparison of ISA spectra with DFT spectra (after Castagna et al. 2003)

Rys. 10. Porównanie widm ISA z widmami DFT (wg Castagna et al. 2003)

One of the important applications of ISA to seismic data processing is a creation of the so-called common frequency section. Fig. 11 explains how the common frequency section is created using synthetic data. Each trace is decomposed using ISA giving time-frequency gather. From the time-frequency gathers corresponding individual section traces with CDP index the established frequency is selected. The graph of the ISA spectrum for the established frequency treated as a function of the time creates the constituent trace of common frequency section. The sections of this type are used to identify the hydrocarbon indicators like low frequency shadows of the gas deposits (Castagna et al. 2003).

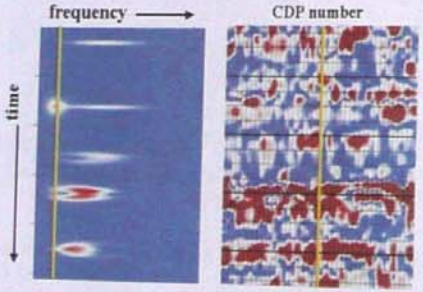


Fig. 11. An idea of the common frequency section creation: the result of ISA decomposition of a single section trace (on the left) with a yellow line denoting selected frequency and its position on the common frequency section (on the right) (after Castagna et al. 2003)

Fig. 11. Idea generowania sekcji wspólnej (stalej) częstotliwości: wynik dekompozycji ISA pojedynczej trasy sekcji (po lewej) z zaznaczoną żółtą linią wybranej częstotliwości i jej położenie na sekcji wspólnej częstotliwości (po prawej) (wg Castagna et al. 2003)

The next application of wavelet transform to seismic data processing is the attenuation of the coherent noise, especially the surface wave (the so-called ground roll) (Deighan, Watts 1997). The coefficients of the WT decomposition let us identify the time of surface wave occurrence (on the horizontal axis) as well as the frequencies of that wave (on vertical axis) corresponding to the significant values of the decomposition coefficients. The process of zeroing these coefficients is equivalent removing the coherent noise. The filtering with the help of wavelet transform does not change the frequency content of the resulting record contrary to the frequency filtering seriously narrowing the frequency band of the record from the side of low frequencies. Any other additional, linear noise is introduced like in the case of application of F-K filter.

One of the application of wavelet transform is a method for computing instantaneous polarization attributes of multicomponent signals using the continuous wavelet transform (Diallo 2006). The next application, most popular for several years is related to spectral decomposition (Peyton et al. 1998, Partyka et al. 1999, Zhang et al. 2006)

3. Singular Value Decomposition and Karhunen-Loeve Transform

Singular Value Decomposition (SVD) is used to decompose the matrix into its eigenvectors and eigenvalues. SVD was frequently applied to seismic data processing (Ursin & Zheng 1985; Freire & Ulrych 1988; Stork 1992a, b; Michelena 1993, Kasina 1984), mainly in solving inverse problems. One of the special application of SVD is Karhunen-Loeve transform used to attenuate the coherent and random noise (Jones & Levy 1987, Al-Yahya 1991, Bitri & Grandjean 2004, Kritski et al. 2007).

Singular Value Decomposition of the symetric $n \times n$ matrix $\hat{A}$ has the form:

$$\hat{A} = \hat{U}\,\hat{\Lambda}\,\hat{U}^T \tag{11}$$

where:

$\hat{U}$ – matrix with columns identical with eigenvectors of matrix $\hat{A}$

$\hat{U}^T$ – transposition of matrix $\hat{U}$

$\hat{\Lambda}$ – diagonal matrix with diagonal elements identical with eigenvalues of matrix $\hat{A}$

Eigenvalues of matrix $\hat{A}$ are the scalars λ for which the equation:

$$\hat{A}\vec{x} = \lambda\vec{x}$$

has nonzero solutions $\vec{x}$ called eigenvectors of matrix $\hat{A}$.

K-L transform of the seismic record has the form:

$$\psi_j(t) = \sum_{i=1}^{n} a_{ij} x_i(t), j = 1,...,n \tag{12}$$

where $x_i(t)$ is the input seismic traces, a_{ij} is the transform coefficients. The above relation may be written in the matrix form:

$$\hat{\Psi} = \hat{A}\,\hat{X} \tag{13}$$

The inverse K-L transform has the form:

$$x_i(t) = \sum_{i=1}^{n} b_{ij}\psi_j(t),\ i = 1,...,m;\ m < n \tag{14}$$

or in the matrix form:

$$\hat{X} = \hat{B}\,\hat{\Psi} \tag{15}$$

To find the coefficients of matrix $\hat{A}$ and $\hat{B}$ we must realize the spectral decomposition of covariance matrix $\hat{C}$:

$$\hat{C} = \hat{X}\,\hat{X}^T$$

In the decomposition we are looking for matrix $\hat{R}$ and $\hat{\Lambda}$ from the matrix equation:

$$\hat{C} = \hat{X}\,\hat{X}^T = \hat{R}\,\hat{\Lambda}\,\hat{R}^T$$

where $\hat{R}$ is the eigenvector matrix of covariance matrix, $\hat{\Lambda}$ is the eigenvalue matrix of covariance matrix. The rows of matrix $\hat{A}$ are the normalized eigenvectors of covariance matrix and matrix $\hat{B}$ is identical with matrix $\hat{A}$ (Jones & Levy 1987). The K-L filtering relies on a modification of matrix $\hat{A}$ by means of its eigenvalues selection. Matrix $\hat{B}$ obtained this way is used to the reconstruction described in the relation (15).

The K-L transform was mainly used to solve two problems of seismic data processing:

– the separation of signal from incoherent and dipping coherent noise in stacked seismic data,

– the suppression of multiples in common-depth-point (CDP) gathers.

To remove noise by means of K-L filter we can use the process named *Eigenvector Filter* in the seismic processing system ProMAX. The parameter window of the *Eigenvector Filter* process is illustrated in Fig. 12.

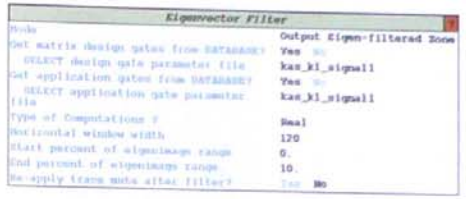


Fig. 12. The parameter window of the *Eigenvector Filter* process in the ProMAX® system

Rys. 12. Okno parametrów procesu *Eigenvector Filter* w systemie ProMAX®

Beside the time gates parameters the most essential parameter is the range of eigenvalues included from the first to the last eigenvalue *(start/end percent of eigenimage range)* expressed in percentage. The eigenvalues are ordered from the maximum to minimum eigenvalue. For example, the range 0 – 10% means the range of eigenvalues defined from the first eigenvalue to the value defined by 10% of all the eigenvalues. The range 85% – 100% means that we include in the filtering only the last 15% of the least eigenvalues. Before the process of filtering we introduce the kinematic corrections (NMO). The example of the records after NMO corrections and before or after K-L filtering for two ranges of the eigenvalues is illustrated in Figs. 13 and 14. After NMO corrections the residual moveouts of multiples are visible at times greater then 600 ms. The analysis of the record in Fig. 13 confirms that the application of K-L filtering in the eigenvalue range 0 – 10% essentially improved the coherent signal to random noise ratio. An application of the second range of eigenvalues 85 – 100% including the least eigenvalues retained only random noise after the reconstruction process (Fig. 14).

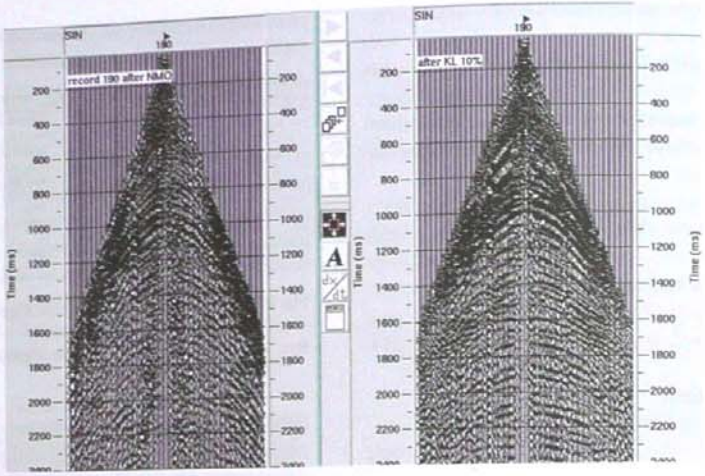


Fig. 13. The seismic record with kinematic corrections (NMO) applied before K-L filtering (on the left) and after K-L filtering with eigenvalues preserved in the range 0 – 10%

Rys. 13. Rekord sejsmiczny po wprowadzeniu poprawek kinematycznych (NMO) przed filtracją K-L (po lewej) i po filtracji K-L z zachowanymi wartościami własnymi z zakresu 0 – 10%

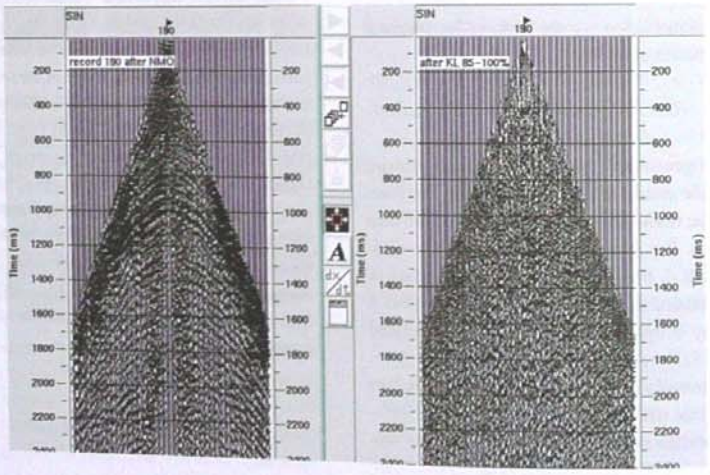


Fig. 14. The seismic record with kinematic corrections (NMO) applied before K-L filtering (on the left) and after K-L filtering with eigenvalues preserved in the range 85 – 100%

Rys. 14. Rekord sejsmiczny po wprowadzeniu poprawek kinematycznych (NMO) przed filtracją K-L (po lewej) i po filtracji K-L z zachowanymi wartościami własnymi z zakresu 85 – 100%

4. Curvelet Transform

The curvelet transform is a multiscale pyramid with many directions and positions at each length scale, and needle-shaped elements at fine scales called curvelets. Curvelets have useful geometric features which make them essentialy different from wavelets. Curvelets obey a parabolic scaling relation which says that at scale 2^{-j}, each element has an envelope which is aligned along a "ridge" of length $2^{-j/2}$ and width 2^{-j} (Cand`es et al. 2006).

In the frequency domain *curvelet* at scale *j* is a wedge whose frequency support is again inside a rectangle, but of size 2^{j} by $2^{j/2}$. Unlike wavelets, curvelets are localized not only in position (the spatial domain) and scale (the frequency domain), but also in orientation.

Curvelets may be used to solve many problems (Cand`es et al. 2006) providing the tools for optimally sparse representation of objects with edges, optimally sparse representation of wave propagators. Curvelets may also be a very significant tool for the analysis and the computation of partial differential equations and for solving reconstruction problems with missing data (for example to reconstruct an object from noisy and incomplete tomographic data).

To define the curvelets we start with a pair of windows W(r) and V (t), which are called the "radial window" and "angular window" respectively. These are both smooth, nonnegative and real-valued, with W taking positive real arguments and V taking real arguments. The arguments of W(r) an V(t) satisfy the conditions (Cand`es et al. 2006):

$$r \in (1/2, 2) \qquad t \in [-1, 1]$$

The window functions fulfil the conditions:

$$\sum_{j=-\infty}^{\infty} W^2(2^j r) = 1, \qquad r \in (3/4, 3/2)$$

$$\sum_{l=-\infty}^{\infty} V^2(t - l) = 1, \qquad t \in (-1/2, 1/2)$$

Now for each $j \geq 0$ we introduce a frequency window U_j defined in Fourier domain (Cand`es et al. 2006):

$$U_j(r, \theta) = 2^{-3j/4} W(2^{-j} r)\, V\left[\frac{2^{\lfloor j/2 \rfloor} \theta}{2\pi}\right] \qquad (16)$$

where $\lfloor j/2 \rfloor$ is the integer part of $j/2$. The support of U_j is a polar "wedge" defined by the support of W and V, the radial and angular windows, applied with scale-dependent window widths in each direction. To obtain real-valued curvelets, we work with the symmetrized version of (4.57):

$$U_j(r, \theta) + U_j(r, \theta + \pi)$$

Now we define the waveform $\varphi_j(x)$ by means of its Fourier transform:

$$\mathcal{F}[\varphi_j(x)] = \hat{\varphi}_j(\omega) = U_j(\omega)$$

We can treat the waveform $\varphi_j(x)$ as „mother"curvelet in the sense that all curvelets at scale 2^{-j} are obtained by rotations and translations of φ_j. If we introduce now two sequences:

– the sequence of rotation angles:

$$\theta_l = 2\pi\, 2^{-\lfloor j/2 \rfloor} l \quad \text{where} \quad l = 0,1,\ldots \quad \theta \leq \theta_l < 2\pi,$$

– the sequence of translation parameters:

$$k = (k_1, k_2) \in Z^2$$

We can define *curvelets* (as function of $x = (x_1, x_2)$) at sca-le 2^{-j}, orientation θ_l and position

$$x_k^{(j,l)} = R_{\theta_l}^{-1}(k_1 2^{-j}, k_2 2^{-j/2})$$

by the relation (Cand`es et al. 2006):

$$\varphi_{j,l,k}(x) = \varphi_j(R_{\theta_l}(x - x_k^{(j,l)})) \qquad (17)$$

where R_θ is the rotation by angle θ (in radians), and R_θ^{-1} is its inverse identical with its transpose:

$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \qquad R_\theta^{-1} = R_\theta^T = R_{-\theta}$$

A *curvelet coefficient* is defined by the integral (Cand`es et al. 2006):

$$c(j,k,l) = \int_{\mathbb{R}^2} f(x)\, \overline{\varphi}_{j,l,k}(x)\, dx \qquad (18)$$

Curvelet tiling of space and frequency in curvelet transform are presented in Fig. 15.

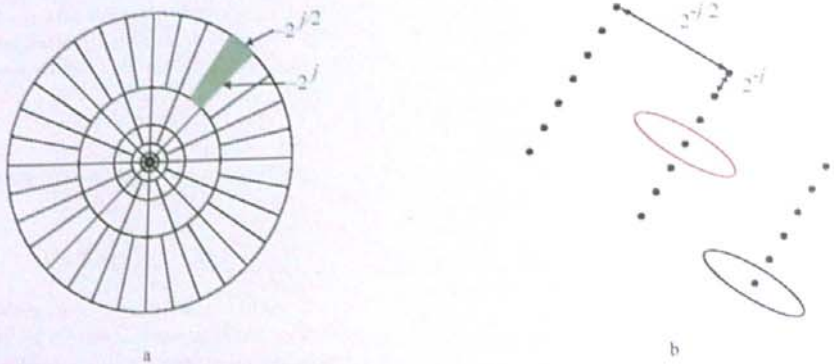


Fig. 15. Curvelet tiling of space and frequency in the curvelet transform called second dyadic decomposition (SDD): (a) the tiling of the frequency plane, (b) the spatial Cartesian grid associated with a given scale and orientation (after Cand`es et al. 2006)

Rys. 15. Podział przestrzeni i częstotliwości w transformacji krzywkowej zwany *second dyadic decomposition* (SDD): (a) podział płaszczyzny częstotliwości, b) kartezjańska siatka przestrzenna związana z daną skalą i orientacją (wg Candes et al. 2006)

The algorithms for 3D discrete curvelet transforms are similar to their 2D analogs. We first decompose the object into dyadic annuli based on concentric cubes. Each annulus is subdivided into trapezoidal regions satistying the usual frequency parabolic scaling (one long and two short directions) (Fig. 16).

Examples of curvelets in space and frequency domain are presented in Fig. 17. In the spatial domain, curvelets are smooth along and oscillatory across the ridge sharp (Fig. 17a), while in the frequency domain their localization is sharp (Fig. 17b).

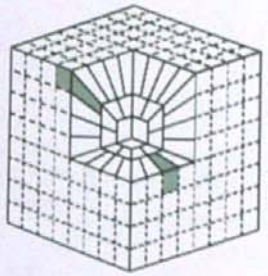


Fig. 16. The dyadic-parabolic frequency tiling in 3D Discrete Curvelet Transform (after Ying et al. 2004)

Rys. 16. Podział diadyczno-paraboliczny częstotliwości w Dyskretnej Transformacji Krzywkowej 3D (wg Ying et al. 2004)

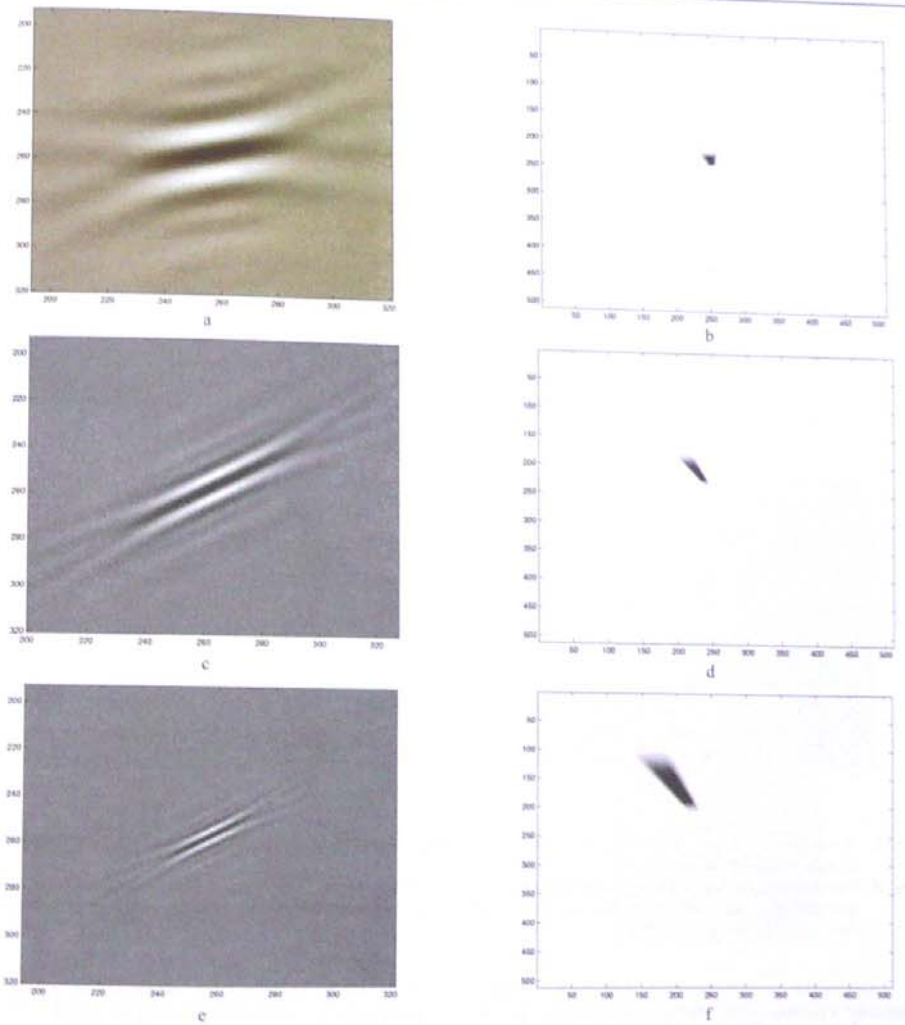


Fig. 17. Curvelets at increasingly fine scales: (a) panels representing curvelets (real part) in the spatial domain (as functions of the spatial variable x), (b) panels showing the modulus of the Fourier transform (as functions of the frequency variable) (after Cand`es et al. 2006)

Rys. 17. *Curvelets* dla wybranych rosnących skal: a) panele reprezentujące *curvelets* (część rzeczywistą) w domenie przestrzeni (jako funkcje zmiennej przestrzennej *x*), b) panele z modułem transformacji Fouriera (jako funkcje zmiennej częstotliwościowej) (wg Candes et al. 2006)

A comparison of wavelet and curvelet at the finest scale is presented in Fig. 18. Even undersampled in the space domain a curvelet (Fig. 18c) shows very good resolution in the frequency domain (fig. 104d) in spite of visible aliasing (low part of Fig. 18d).

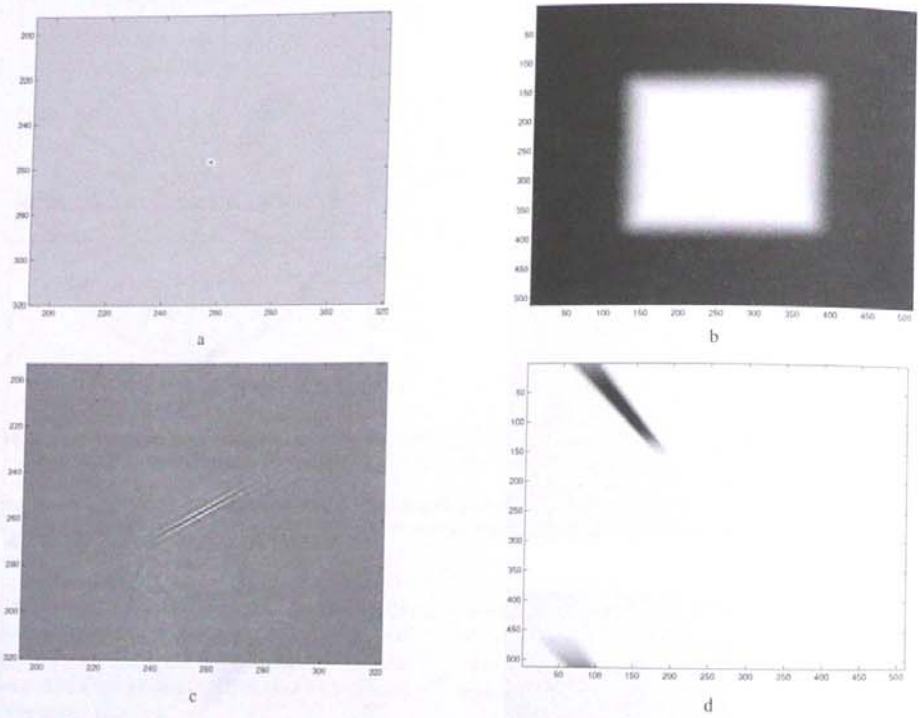


Fig. 18. Wavelets and curvelets at the finest scale. Meyer wavelet in space (a) and frequency (b). Undersampled curvelet in space (c) and frequency (d) (after Cand`es et al. 2006)

Rys. 18. Fale elementarne i *curvelets* dla najsubtelniejszej skali: fala elementarna Meyera w domenie przestrzeni (a) i częstotliwości (b), zbyt rzadko spróbkowana fala elementarna w domenie przestrzeni (c) i częstotliwości (d) (wg Cand`es et al. 2006)

One of the curvelet transform application is image denoising. The example of this application in the case of seismic record is illustrated in Fig. 19. The results are compared with the effects of the wavelet transform.

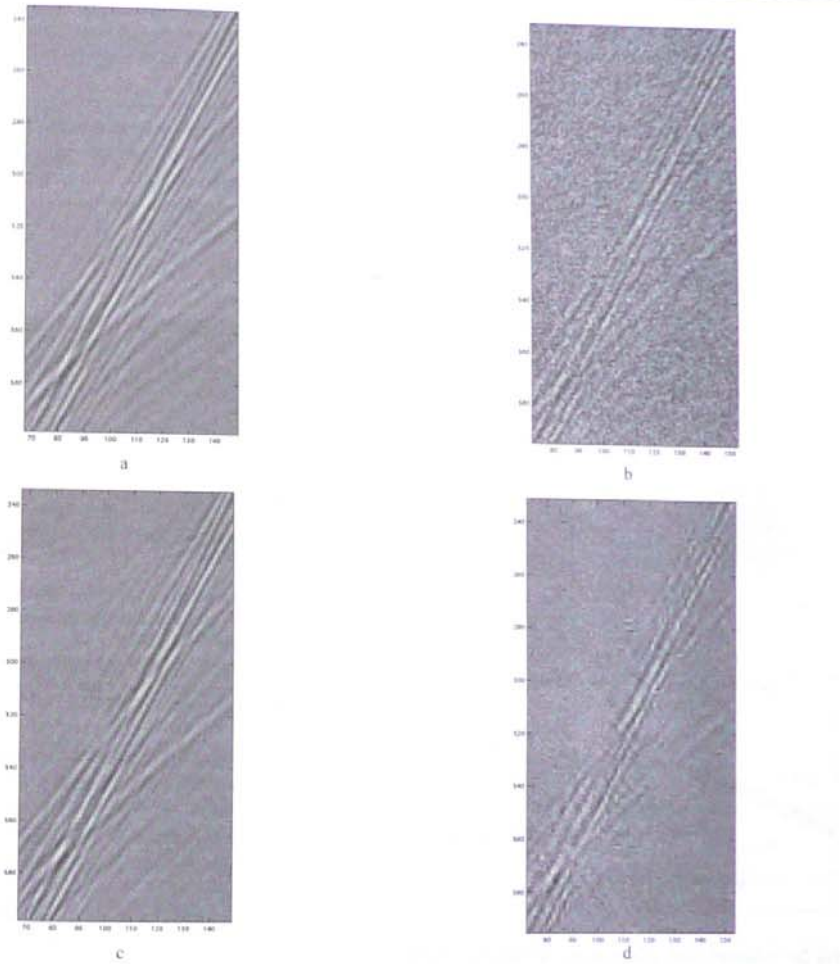


Fig. 19. Image denoising using curvelets: (a) Original image of the seismic record (zoom), (b) Noisy image (Gaussian white noise), (c) Denoised image using curvelets, (d) Denoised image using wavelets (after Cand`es et al. 2006)

Rys. 19. Usuwanie szumu z obrazu z zastosowaniem *curvelets*: (a) oryginalny obraz rekordu sejsmicznego (powiększenie), (b) obraz z nałożonym szumem (biały szum gaussowski), (c) obraz po usunięciu szumu z zastosowaniem *curvelets*, (d) obraz po usunięciu szumu z zastosowaniem *wavelets* (wg Cand'es et al. 2006)

The parabolic scaling used in the curvelet transform is very helpful in better resolving the edgelike components of the images (Cand`es, Guo 2001). We can achieve better accuracy in the vicinity of edges while using many fewer terms in an approximation comparing with other techniques (for instance with wavelet transform). The essence of imaging the objects with edges using the curvelet transform is illustrated in Fig. 20. In the figure three kinds of curvelets are showed:

– curvelets of type A, whose essential support does not overlap with the discontinuity,

– curvelets of type B, whose essential support overlap with the discontinuity but are not tangent to the singularity,

– curvelets of type C, which overlap with the singularity and are nearly tangent to the singularity.

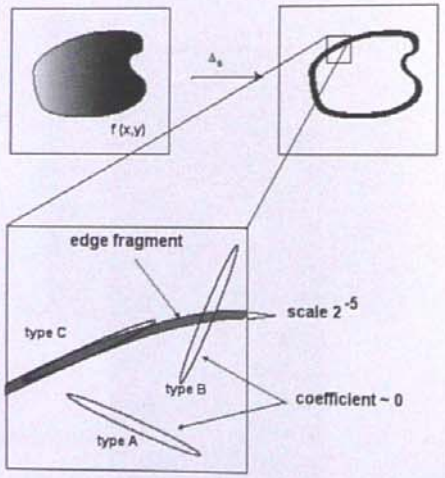


Fig. 20. Schematic decomposition of a subband. The top figures represents an object with an edge and that same object after applying a bandpass. The bottom picture represents a bandpassed edge fragment together with the three types of curvelets described in the text (after Cand`es 2002)

Rys. 20. Schematyczna dekompozycja obiektu z krawędzią. Górne figury reprezentują obiekt z krawędzią oraz ten sam obiekt po zastosowaniu filtru pasmowego. Dolny obraz reprezentuje fragment krawędzi po filtracji pasmowej z trzema typami *curvelets* opisanymi w tekście (wg Cand`es 2002)

The main feature of curvelets type A which are located near the edge curve is rapid decay of their coefficients as the distance between the edge curve and the center of the curvelets increases. Generally, curvelet coefficients decay as the angle between their orientation and that of the edge increases. The comparison of the image reconstructions using the curvelet and wavelet transforms is presented in Fig. 21.

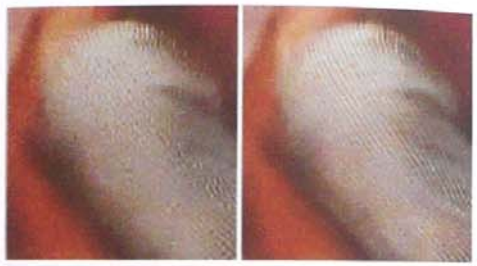


Fig. 21. A comparison of the image reconstruction using: (a) wavelet transform, (b) curvelet transform (after Cand`es & Guo 2001)

Rys. 21. Porównanie wyników rekonstrukcji obrazu z zastosowaniem: (a) transformacji falkowej, (b) transformacji krzywkowej (wg Cand`es i Guo 2001)

The curvelet transform has been recognized as a very effective tool of seismic data processing. The main applications are related to:

– the reconstruction of seismic wave fields using regularly sampled data with many traces missed (Hennenfent & Herrmann 2008),

– the separation of primaries and multiples (Herrmann et al.2007, Herrmann et al. 2008a, b),

– the restoring migration amplitude by amplitude scaling of wave fields with spatially varying and dip-dependent amplitude deterioration (Herrmann et al. 2008a, Chauris & Nguyen 2008),

– the improving the signal to noise ratio in the case of random noise and strong coherent linear dipping noise (Neelamani et al. 2008).

In the process of seismic data processing we can treat the curvelet transform as mathematical transform that represents an m-dimensional signal $D(x_k)$ using linear, weighted combination of curvelet functions $C_j(x_k)$ (Neelamani et al. 2008):

$$D(x_1, \ldots, x_m) = \sum_j \lambda_j \times C_j(x_1, \ldots, x_m) \qquad (19)$$

where weights λ_j are referred to as the curvelet coefficients of data D. The curvelet transform is linear. Therefore, each curvelet coefficient of noisy data is the sum of the corresponding curvelet coefficient of the signal and the noise:

$$\lambda_{\lambda,noisy\ data} = \lambda_{\lambda,signal} + \lambda_{\lambda,noise}$$

In 2D (m = 2) each curvelet function resembles a small piece of a band-limited seismic reflector. Each curvelet function is oscillatory in one direction, but varies more smoothly in the orthogonal direction. The oscillations in different curvelet functions occupy different frequency bands. Each curvelet function is spatially localized because its amplitude rapidly decays to zero outside a certain region. Summarizing we can state that each 2D curvelet function has a characteristic dip, frequency (thickness), and location. In 2D the region that essentially localizes a curvelet function resembles an oriented needle. In 3D a curvelet function resembles an oriented disc and it has characteristic dip, azimuth, frequency (thickness), and location. The idea of using the curvelet transform for denoising is based on the fact that the signal and noise map into different sets of coefficients after curvelet transformation. Removing the noise is achieved by selective attenuation (multiplication by small scalar) the curvelet coefficients of the noise before reconstruction (mm). The same technique may be is used to improve not only seismic images (Fig. 21).

Bibliography

Al-Yahya K. M., 1991, *Application of the partial Karhunen-Loeve transform to suppress random noise in seismic sections*. Geophysical Prospecting, vol. 39, 77–93.

Bastiaans M. J., 1980, *Gabor's expansion of a signal into Gaussian elementary signals*. Proc. IEEE 68, 538–539.

Bitri A., Grandjean G., 2004, *Suppression of guided waves using the Karhunen-Loève transform*. First Break, vol. 22, May, 45–47.

Castagna J. P., Sun S., Siegfried R. W., 2003, *Instantaneous spectral analysis: Detection of low-frequency shadows associated with hydrocarbons*. The Leading Edge, 120–127.

Chakraborty A., Okaya D., 1995, *Frequency-time decomposition of seismic data using wavelet-based methods*. Geophysics, vol. 60, 1906–1916.

Deighan A. J., Watts D. R., 1997, *Ground-roll suppression using the wavelet transform*. Geophysics, vol. 62, 1896–1903.

Diallo M. D., Kulesh M., Holschneider M., Scherbaum F., Adler F., 2006, *Characterization of polarization attributes of seismic waves using continuous wavelet transforms*. Geophysics, vol. 71, V67–V77.

Freire S.L., Ulrych T.J., 1988, *Application of singular value decomposition to vertical seismic profiling*. Geophysics, vol. 53, 778–785.

Gabor D., 1946, *Theory of communication*: J. IEEE (London), 93(III), 429–457.

Glinsky M. E., Clark G. A., Cheng P. K. Z., Devi K. R. S., Robinson J. H., Ford G. E., 2001, *Automatic event picking in prestack migrated gathers using a probabilistic neural network*. Geophysics, vol. 66, 1488–1496.

Jones I. F., Levy S., 1987, *Signal-to-noise ratio enhancement in multichannel seismic data via the Karhunen-*

-Loeve Transform. Geophysical Prospecting, vol. 35, 12–32.

Kasina Z., 1984, *Liniowa analiza poprawek statycznych w ujęciu teorii inwersji ogólnej*. Technika Poszukiwań Geologicznych, zesz. 3, 1–5.

Kasina Z., 1998a, *Przetwarzanie danych sejsmicznych*. Wyd. Centrum PPGSMiE PAN, Kraków.

Kasina Z., 1998b, *Metodyka badań sejsmicznych*. Wyd. Instytutu GSMiE PAN, Kraków.

Kritski A., Vincent A. P., Yuen D. A., Carlsen T., 2007, *Adaptive wavelets for analyzing dispersive seismic waves*. Geophysics, vol. 72, V1–V11.

Margrave G. F., Lamoureux M. P., 2002, *Gabor deconvolution*. CSEG Geophysics, p.1.

Michelena R. J., 1993, *Singular value decomposition for cross-well tomography*. Geophysics, vol. 58, 1655–1661.

Morlet J., Arens G., Fourgeau E., Giard D., 1982a, *Wave propagation and sampling theory-Part I: Complex signal and scattering in multilayered media*. Geophysics, vol. 47, 203–221.

Morlet J., Arens G., Fourgeau E., Giard D., 1982b, *Wave propagation and sampling theory-Part II: Sampling theory and complex waves*. Geophysics, vol. 47, 222–236.

Partyka G., Gridley J., Lopez J., 1999, *Interpretational applications of spectral decomposition in reservoir characterization*. The Leading Edge, vol. 18, 353–360.

Peyton L., Bottjer R., and Partyka G., 1998, *Interpretation of incised valleys using new 3-D seismic techniques: A case history using spectral decomposition and coherency*. The Leading Edge, vol. 17, 1294–1298.

Sinha S., Routh P. S., Anno P. D., Castagna J. P., 2005, *Spectral decomposition of seismic data with continuous-wavelet transform*. Geophysics, vol. 70, P19-P25.

Stork C., 1992a, *Singular value decomposition of the velocity-reflector depth tradeoff, Part 1 Introduction using a two-parameter model*.Geophysics, vol. 57, 927–932.

Stork C., 1992b, *Singular value decomposition of the velocity-reflector depth tradeoff, Part 2: High-resolution analysis of a generic model*.Geophysics, vol. 57, 933–943.

Ursin B, Zheng Y., 1985, *Identification of Seismic Reflections using Singular Value Decomposition*. Geophysical Prospecting, vol. 33, 773–799.

Yongwang M., Margrave G., 2006, *Prestack Depth Migration with the Gabor Transform*. CSPG-CSEG-CWLS Convention, p. 85.

Zhang J., Chapman M., Liu E., Lix X. Y., 2006, *Application of spectral decomposition to detection of dispersion anomalies associated with gas saturation*. The Leading Edge, 206–210.

Streszczenie

W pracy przedstawiono kierunki najnowszych zastosowań transformat matematycznych w przetwarzaniu i analizie danych sejsmicznych. W analizie uwzględniono transformację Gabora, transformację Karhunena-Loevego (*K-L Transform*), transformację falkową (*Wavelet Transform*) oraz transformację krzywkową (*Curvelet Transform*). Przedstawiono podstawy teoretyczne tych transformacji oraz ich główne zastosowania związane z usuwaniem zakłóceń koherentnych i przypadkowych w procesie przetwarzania danych sejsmicznych. Opisano także zastosowania związane z analizą sekcji sejsmicznych opartą na chwilowej analizie spektralnej (*ISA*) oraz zastosowania związane z analizą obrazów. Przedstawiono porównanie wyników różnych typów dekompozycji spektralnej na wybranych danych modelowych i polowych. Analizując zastosowania transformacji falkowej w przetwarzaniu danych sejsmicznych szczególną uwagę zwrócono na procedurę konstruowania sekcji wspólnej częstotliwości, wykorzystywanej do analizy cienkich warstw oraz definiowania bezpośrednich wskaźników węglowodorów. Opisano także rolę transformacji falkowej w tłumieniu fali poowierzchniowej. Analizując zastosowania transformacji K-L opisano jej rolę w poprawianiu stosunku sygnałów koherentnych do szumu przypadkowego oraz w tłumieniu refleksów wielokrotnych. W przypadku transformacji krzywkowej uwypuklono jej rolę w rekonstrukcji pól falowych nieregularnie spróbkowanych, poprawianiu stosunku sygnałów koherentnych do szumu przypadkowego, separacji fal jednokrotnych i wielokrotnych, poprawiania rozdzielczości bitmap.

ZBIGNIEW PERSKI[1]

EARTH'S SURFACE DEFORMATION MEASUREMENTS WITH SAR INTERFEROMETRY. METHODS AND NEWEST ARCHIVEMENTS

Key words:

SAR interferometry, SAR imagery, digital processing, terrain deformation, mining

1. Abstract

In this paper the SAR inteferometric techniques and their applications to Earth's surface deformations are presented. The practical use of satellite SAR (Synthetic Aperture Radar), is discussed based on examples from Poland, where these methods have been applied for the measurement of terrain deformations. The papier is focused on the studies of the natural earth surface displacements in Poland performed within GEO-IN-SAR project. The issues related to newest high resolution SAR systems are discussed with TerraSAR-X data. Due to short data archiving histoty those data are not sutable for natural surface movement detection. However, the performance of TerraSAR-X data against Envisat are discussed. The preliminary PSInSAR processing in presented studies did not allow to clearly detect any recent tectonic movement. However, the „blind" experiments over relatively small areas proved high reliability of PSInSAR results.

POMIARY DEFORMACJI POWIERZCHNI ZIEMI PRZY UŻYCIU SATELITARNEJ INTERFEROMETRII RADAROWEJ. METODY I NAJNOWSZE OSIĄGNIĘCIA

Słowa kluczowe:

Inreferometria SAR, zobrazowania SAR, przetwarzanie cyfrowe, deformacje terenu, górnictwo

Abstrakt

Artykuł prezentuje metodykę satelitarnej interferometrii radarowej (InSAR) i jej zastosowania do pomiarów deformacji powierzchni Ziemi. W oparciu o przykłady zastosowań z terenu Polski przedstawiono wykorzystanie satelitarnych zobrazowań SAR (radar obrazowy z aperturą syntetyczną). Artykuł przedstawia przede wszystkim wykorzystanie technik InSAR do badania deformacji powierzchni terenu o przyczynach naturalnych będących przedmiotem studiów w ramach projektu GEO-IN-SAR. Omówiono również dane wysokorozdzielcze z sensora TerraSAR-X. Z powodu krótkiego okresu obserwacji, dane te nie nadają się jeszcze do badania deformacji naturalnych. Przeprowadzono jednak stadium porównawcze danych TerrSAR-X z danymi z satelity ENVISAT. Przeprowadzone dotychczas badania nie pozwoliły na wykrycie deformacji terenu w sposób jednoznaczny. Przeprowadzone testy na niewielkich obszarach wykazały ich ogromny potencjał.

[1] Department of Fundamental Geology, Faculty of Earth Sciences University of Silesia, Sosnowiec, Poland.

2. Introduction

SAR interferometry (InSAR) is a remote sensing technique that makes use of the (SAR) phase information by subtracting the phase value of one image (usually called master) from that of the other (usually referred to as a slave image), for the same point on the ground. This technique, briefly described in this paper, has been used in many succesful applications related to earth-surface deformation. The paper is focused on the studies of natural earth surface displacements in Poland which were performed within GEO-IN-SAR project.

The measurements of natural terrain deformations are from technical viewpoint very difficult, time consuming and expensive. Moreover, for reliable results the measurements must be done repeatedly over many years. Due to that fact and due to a small practical importance such measurements are performed rarely. InSAR is one of the most advanced techniques for the measurements of subtle, very slow movements which are usually announcing the seismic risk.

Natural recent surface movements had not yet been studied with InSAR in Poland before. Similar works abroad were focused on relatively high rate displacements like e. g. rifting activity in Iceland (Jonsson 2002) or the development of the anticline in Taiwan (Fruneau et al 2001). In Europe, the most of InSAR deformation studies were focused on anthropogenic (technogenic) phenomena (e.g. Perski, Jura 1999). Few works were done on landslides or natural subsidence (e.g. Colesanti, Wasowski, 2006, (Strozzi et al., 2003).

3. SAR imagery

SAR, also known as coherent radar, overcomes the limitation of the antenna size typical of real-aperture radars by synthesizing an antenna which receives a series of reflected waves and electronically combines them with reference wavelengths. The resolution of an SAR effectively remains the same over all ranges (Gens and Logan 2003). SAR technology became very popular in the past decade; it provided, for example, terrain structural information for mineral exploration, oil spill boundaries on water, sea state and ice hazard maps, and reconnaissance and targeting information for military operations. There are also many other operational or potential applications. For general mapping purposes, SAR imagery is very interesting because it can produce high-quality images independent of cloud cover and solar illumination. A typical geological image interpretation in the case of radar is based on backscatter images, where variation of radar look angle and wavelengths can have a substantial effect (Henderson and Lewis 1998). The interpretation capabilities in the case of radar could be extended by the application of color composition images constructed in the following ways (Perski, 2005):

(1) Multi-frequency: constructed from images acquired at different frequencies/bands. The application of such data is usually limited to multi-frequency sensors such as SIR (shuttle imaging radar) or airborne missions, however it is possible to construct such images in 'multi-platform mode' using data from different satellites, e.g. SAR images acquired by ERS-1 and JERS-1 satellites;
(2) Multi-temporal: compositions developed from data from the same sensor but acquired at different times. Colors are used to represent changes in backscatter over time between acquisitions;
(3) Multi-polarization: composed images acquired simultaneously by the same SAR system but with different transmitter/receiver polarization configurations, e.g. H/H, V/V, H/V.

For the last 10 years SAR technology has been developing very fast. Now there are available multi-polarisation space borne systems of different wavelengts and spatial resolution: mid-resolution C-band ERS–2/Envisat, Radarsat–1, L-band ALOS-PALSAR, high-resolution X-band TerraSAR-X, constellation of COSMO-SkyMed satellites. In the following years first high-resolution C-band SENTINEL satellite will be lauched.

4. InSAR and PSInSAR methods

SAR Interferometry (InSAR) is a technique for extracting information related to the topography of the Earth's surface (Goldstein et al., 1988). It uses the phase difference between the radar echoes from repeated SAR (Synthetic Aperture Radar) observations of the same area. The result of this operation is known as an interferograms, presenting

relative phase differences 'wrapped' within 2Π radians (Fig. 1). By applying phase unwrapping techniques (Ghiglia and Pritt, 1998) it is possible to reconstruct a full unambiguous signal. First interferometric studies, focused on topography retrieval, demonstrated the applicability of InSAR to digital elevation model generation (Ferretti et al., 1997). Differential InSAR (D-InSAR) represents a branch of InSAR that exploits the temporal baseline between consequent SAR acquisitions to derive phase differences which correspond with terrain displacements. D-InSAR has already been successfully used in different applications: the monitoring of volcanic activity, earthquakes, glacier dynamics, landslides and urban subsidence. In many cases D-InSAR has demonstrated its capability in measuring surface movements of the order of centimeters. A good overview of DInSAR technique and its applications was given by (Bamler and Hartl, 1998; Massonnet and Feigl, 1998; Rosen et al., 2000 and Hanssen, 2001).

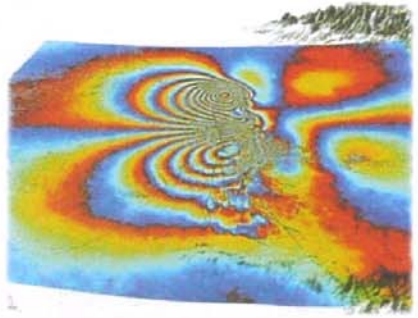


Fig. 1. SAR interferogram generated from Envisat ASAR images that depicts terrain deformation resulted from devastating earthquake in Bam city (Iran). The full colour cycle represents ~3 cm of LoS (Line of Sight) deformation. For the purposes of 3D presentation the interferogram was wrapped over SRTM–3 DEM data of the area. (Perski, Hanssen 2005)

Rys. 1. Interferogram SAR wygenerowany w oparciu o dane ASAR z satelity ENVISAT przedstawiający deformacje terenu powstałe w wyniku niszczycielskiego trzęsienia Ziemi w rejonie miasta Bam (Iran). Pełny cykl barwnych prążków interferomerycznych reprezentuje ok. 3 centymetrową zmianę wysokości w kieryunku do/od satelity. Dla potrzeb prezentacji interferogram przedstawiono w widoku trówymiarowym na modelu terenu SRTM-3. (Perski, Hanssen 2005)

The main problem associated with D-InSAR is the so-called temporal decorrelation, due to changes in the electromagnetic properties and/or relative positions of scatterers within a resolution cell. Moreover, D-InSAR is very sensitive to atmospheric signal delay. The variable water vapor distribution related to the turbulent character of the atmosphere creates an interferometric phase contribution (Hanssen et al., 2006). For singular interferograms, this atmospheric phase screen (APS) is impossible to remove and therefore the accuracy of measuring small deformations is significantly reduced. Due to those properties the operational use of D-InSAR is limited to a) short temporal baselines, b) phenomena with strong deformation gradient in respect to radar wavelength within the acquisitions, c) areas with limited vegetation, and d) advantageous weather conditions during master and slave acquisitions. Subject to these limitations, many successful studies were performed regarding subsidence due to underground mining (Perski and Jura, 1999; Wright and Stow, 1999). Fewer studies focused on landslides due to slow movement in the presence of vegetation (Colesanti et al., 2003; Hilley et al., 2004).

To bypass the limitations mentioned above point wise InSAR techniques have been developed since 2000, see e.g. (Ferretti et al., 2001). The initial method developed by POLIMI (Politechnico Milano Team) used radar point targets as 'natural' corner reflectors. The phase of such targets (labeled Persistent Scatterers, or PS) is not sensitive to small incidence angle variations and temporal decorrelation. The time series of interferometric phase is decomposed into (linear) deformation, topography (relative height) and APS. Such operation can only be performed for scatterers with coherent phase behavior over time. The exploration of all available SAR images (typically more than 20 acquisitions) by coregistering and resampling them to the same master is fundamental in these point wise methods. Then the interferograms between all slave images and the master are computed. An overview of PSI applications for landslide studies was given by (Colesanti and Wasowski, 2006). Other techniques, such as the Small Baseline Subset technique (SBAS) (Berardino et al., 2002) utilize several master images to construct an optimal set of interferograms with the smallest temporal and perpendicular baselines. The crucial element in the point wise analysis is the identification of potentially coherent points (Fig. 2). The first selection might be based on amplitude dispersion as in the original PSI algo-

rithm (Ferretti et al., 2001) or coherence as in the case of SBAS or StaMPS (Hooper et al., 2004). Redundant observations are consecutively used to estimate the APS and the (linear) deformation. This estimation is based on the properties of the specific signals: APS is strongly correlated in space but not in time, whereas deformation is usually strongly correlated in time (Colesanti et al., 2003).

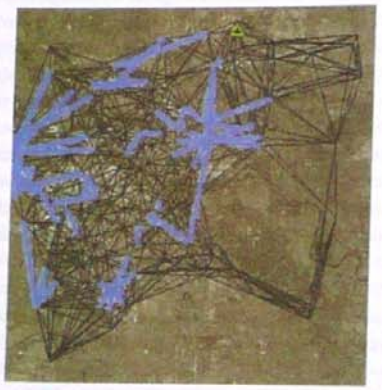


Fig. 2. PSInSAR method. PSC network construction and arc rejection (blue). The reference point is marked as a green triangle. Wrocław area (Perski, Mróz 2008).
Rys. 2. Metoda PSInSAR. Utworzona sieć PSC oraz wyeliminowane po wyrównaniu segmenty (niebieskie). Położenie punktu referencyjnego oznaczono zielonym trójkątem. Obszar Wrocławia (Perski, Mróz 2008).

It should be noted that all the InSAR measurements are completed along Line of Sight (LoS) direction. To properly decompose that measurement into horizontal and vertical deformation the measurements taken from other directions are required and/or the measurements taken by other techniques. Unfortunately within this study the data acquired from a single direction were available. In case of mining subsidence, where we assumed mainly vertical deformation, the vertical component from InSAR data was derived. For the landslides since no ground truth data were available the only LoS deformation is considered.

In the case of PSInSAR method the most serious doubty are associated with the proper phase unwrapping. The maim problems are associated here with low density networks and long arcs. For non-linear deformations the pahse decomposition may give wrong results where the deformation signal will be mixed with the atmospheric component. Therefore, the results must be very carefully analyzed.

5. PSInSAR studies with ERS–1/2 data

Thanks to huge data archive containing observatiuons since mid 1991 the ERS–1/2 of European Space Agency is the best suitable system for PS Interferometry. ERS–1/2 SAR data have been used in GEO-In-SAR project, whose main idea was to use the newest achievements in a field of SAR interferometry to measure recently occurring earth surface movements. Three study areas were selected where the evidence of vertical movements had been already archived (Fig. 3) The first site includes Sambia peninsula and Warmia area which were affected by the strong earthquake in September 2004. The second site is located in Lower Silesia where the system of recently active faults has been mapped. The third region is located in Podhale and the Tatra Mounatins, which is dominated by young alpine tectonics associated with geothermal springs and frequent earthquakes.

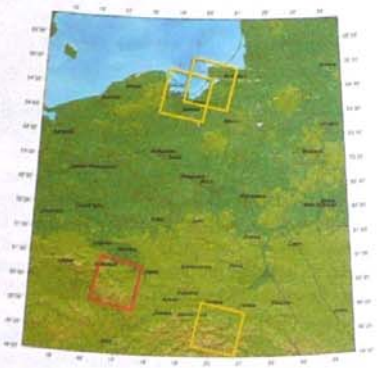


Fig. 3. Areas of interest of the GEO-In-SAR project. Yellow and red squares shows the extent of processed ERS–1/2 SAR frames.
Rys. 3. Obszary badawcze projektu GEO-In-SAR. Zółte I czerwone czworoboki pokazują zasiięgi użytycg scen satelitarnych ERS-1/2 SAR

5.1. Natural surface movements – Paczków area

The interferometric processing for the study area was performed based on 67 ERS–1/2 SAR images covering almost 8-year period from 04 April 1992 to 11 January 2001. The data acquired from the descending satellite track (Fig. 3) were used. For PSInSAR processing the TUDelft implementation (Leijen et al. 2005) of original Persistent Scatteres algorithm was applied. For D-InSAR part of the processing the Delft Object Oriented Interferometric Software (DORIS) was used (Kampes at al 2003). The final analysis of the PSInSAR results requires GIS environment, which allows putting all interferometric results and external data into one common reference system. For this purpose open-source GRASS (Geographic Resources Analysis Support System) was applied (Grass Development Team 2006).

Due to very sparse urbanisation and the presence of agriculture and large areas occupied by water the problem of coregistering of the SAR images occurred. The problem was solved by the application of the coregistration with optimization of window distribution based on amplitude and optimization of the computation of coregistration polynomial. Also 64 interferograms were used as the input for PSInSAR processing. The results show the appearance of the clusters of PS points associated with the following towns: Nysa, Paczkow and surrounding villages. The connection between those two main clusters is made through sparse PS points located around water reservoirs and on dams. As in the case of Wrocław area the linear model was applied for PSInSAR processing. The obtained linear velocities are ranged between –5 and +5 mm/year for the majority of highly coherent points (Fig. 4). The most interesting feature is the cluster of PS points located within the town Nysa. Nysa, with 47 thousands of inhabitants occupies the area of 27 km2 which allows to obtain high PS density (> 100 PS/km2) with the/a reference point located in the north-western part of the city. The densely urbanised centre of the town shows relative upward deformation of 5 mm/year. Accordingly, the most of the surroundings of Nysa Lake and the town of Paczków reveal subsidence of some of 2 mm/year. Those areas might be considered as "stable". The relative uplift of Nysa is very consistent and showing zones of gradual decrease of deformation towards the limits of the urbanised area. The high consistency of the first results convinced us to process the slightly larger crop of the same dataset. The obtained results of that second PSInSAR processing shows the same pattern. As it was described in section two the upward movement of Nysa had already been reported by (Cacoń et al., 2005) and (Kowalczyk, 2006) based on levelling and gravity data. However the origin of that deformation is not clear. The most probable explanation is the groundwater level rise due to changes in industrial technology and economy of the region. The role of active faults mapped in that area is still not confirmed.

Fig. 4. The PSInSAR results of Paczków area. The relative velocities are varying between –5 to 5 mm/yr. The crop size is approx 58 × 17.5 km
Rys. 4. Wyniki przetwarzania danych metodą PSInSAR dla obszaru Paczkowa. Obliczone względne prędkości deformacji wachają się pomiędzy -5 do 5 mm/rok. Rozmiary obszaru ok. 58 × 17.5 km

5.2. Landslides and mining subsidence in Wieliczka

More than 2000 underground chambers have been excavated in the Wieliczka salt mine within the last 700 years. The slow convergence of theses cavities has induced several meters of subsidence. Apart from slow convergence, the mine suffers from catastrophic events such as collapse and water inflow, causing sinkholes and rapidly developing localized subsidence (e.g. up to 1.5 m during 6 months in 1992). Additional geohazards in the Wieliczka region are landslides, as it is located at the edge of the

Carpathian Mountains. Up to now, no clear relation between mining subsidence and the triggering of landslide movement has been found.

(Wójcik and Mrozek, 2002) found five large landslides to the south of Wieliczka in the region of Lednica, Grabowki and Choragwica hamlets (Fig. 5). Here, large fragments of slopes are occupied by landslides covering the total area of 2 km^2. The landslide zone begins with 23 m high niches close to Choragwica hill. The activity of the landslides manifests in various parts of the colluvia causing damages to buildings and infrastructure. In the vicinity of Siercza and Grabowki landslides have damaged the roads in 3 places. According to Varnes (1978) the described landslides could be classified as the compound landslides type. The upper parts of these landslides are permanently active and because of that a section of the road surface has been subsiding by 5 cm. Landslide velocities are varying from a few mm/y to 20 cm/year. The lower parts of landslides are most active. Slide surfaces (shear surfaces) occur commonly at depths of 3–4.5 m, which was confirmed by observations in tectonic windows.

In the PSI processing 51 interferograms were used. After obtaining the first results the processing was repeated with a modified setup to ensure that the reference point to which all the observations are referred to was not located within an area likely to be affected by deformation. For the extraction of subsidence signals from SAR observations various models may be used (Leijen and Hanssen, 2007). However the application of a specific model should be tuned to a priori knowledge about the subsidence development in space and time. As terrestrial geodetic measurements since the 1980s proved that the deformation is steady-state (Kortas, 2007), a constant velocity model was applied. To improve the separation of that part of the interferometric signal related to topography the interferograms were regenerated and subtraction of topography with an external DEM was performed. As external DEM (Digital Elevation Model) the one-arcsecond DTED (Digital Terrain Elevation Data) of level–2 (Defence-Mapping-Agency, 1986) was used. The requirement of topography removal comes from the fact that Wieliczka is located in the Carpathians, in a hilly terrain of 200 m topography (Fig. 5).

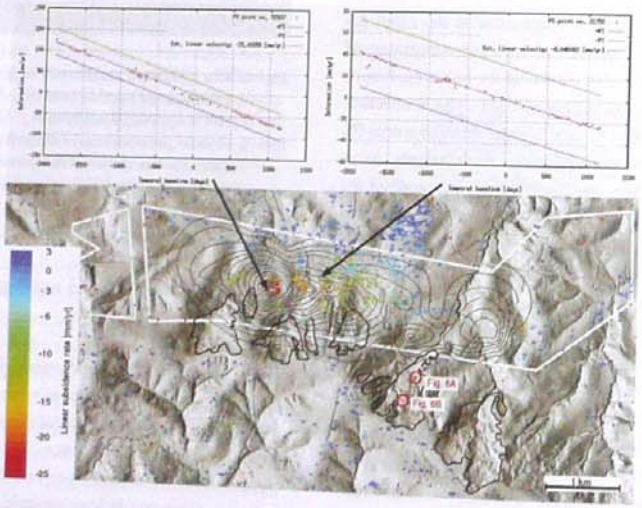


Fig. 5. Map of PS points in the Wieliczka region. The PS used as the refernce point is located outside the presented crop. (Perski et al 2009). Plots present the history of deformation for selected PS

Rys. 5. Mapa punktów PS dla obszaru Wieliczki. Punkt PS wykorzystany jako referencyjny znajduje się poza zasięgiem prezentowanego fragmentu. (Perski et al 2009). Wykresy przedstawiają historie deformacji dla wybranych punktów obszaru.

A large number (30.000) of PS points is detected within the study area of 18 × 18 km (average density of 92 PS/km^2). Within the mine concession area (164 PS/km^2) and surroundings (Fig. 5) 1895 PS points are detected. Most of the scatterers are related to buildings and are located within the densely urbanized center of Wieliczka, where the maximum PS density was 480 PS/km^2. Wieliczka is surrounded by villages and hamlets characterized by sparse urbanization which allow to get a good spatial distribution of scatterers (~30 points/km^2) for the rest of the area of interest. Most of the scatterers are related there to buildings but also to other metallic and concrete objects like fences, roads, and poles. General knowledge of the scatterer characteristics is obtained by the comparison of PS locations with detailed topographic maps followed by field inspection.

Some PS with different behavior, i.e., variable velocities have been found on the areas of known active landslides. The PS velocities are highly dispersed over the area: up to 25mm/yr between neighbor points. Due to small number of man-made structires the PS density is very low, which results in the weaker network and then in higher errors. In our study only points of overall very high coherence, whose locations correspond with the extent of the active landslides (Fig. 5) were considered. To verify the reliability of these points the field examination was performed. The coordinates of the points was plotted on the topographic map and loaded into field GPS receiver. During the field investigation some PS located on buildings with visible damage were identified, see Fig. 6. Unfortunately, a large number of PS was not identified and might relate to scattering from poles, power lines, fences, and roads, without clear evidence of damage. The field verification was focused mainly on the points with coherence higher than 0.8 and PS on landslides were verified in the field against the evidence of visible damage. During that work, some buildings where construction damage and failures are visible were identified (Fig. 6). For those located on the lowest part of the slopes the origin of the damage is not clear since these parts are also within the range of the influence of the mining-induced subsidence. However, it should be noticed that many buildings located on the upper parts of the slopes have been damaged. The origin of the damage is not always clear since all failures are cracks, which are quickly repaired. Some of them may be caused by construction mistakes or bad quality building material.

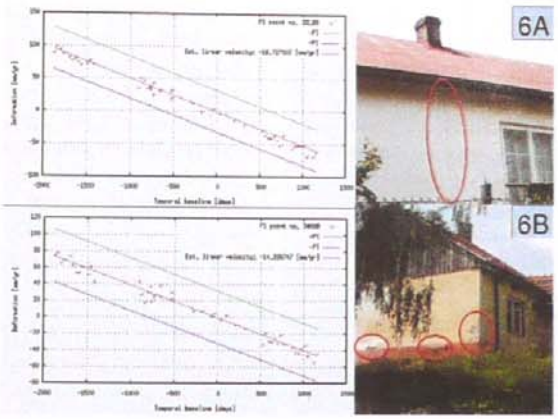


Fig. 6. Identification of PS points on landslides and visible evidence of construction damage caused by the landslide. Presented time series of corresponding deformation show linear behavior. It could be noted that the roofs remains not damaged, which probably allows preserving phase of backscattered SAR signal over the time. (Perski et al 2009)

Rys. 6. Identyfikacja punktów PS na obszarach osuwisk powiązana z widocznymi uszkodzeniami konstrukcji budynków. Prezentowane historie deformscji punktów ujawniają liniowy character deformacji. Uwagę zwracają nieuszkodzone dachy budynków, co prawdopodobnie pozwoliło na zachowanie fazy interfeormetrycznej dla calej serii obserwacji. (Perski et al 2009)

As the northward direction of the landslide movement (horizontal component) is not suitable for interferometry with polar satellite orbits (Colesanti and Wasowski, 2006), the observed movement is not spatially uniform. Each point behaves individually according to its local conditions and local movement direction. It should also be emphasized that PS velocity was calculated against a linear model of deformation. Thus, the analyzed cases present the cases close to linear movemets of relatively low magnitude (up to 25mm/year) that allow to preserve interferometric coherence and high backscatter amplitude over the time. Unfortunately, our observations cannot be verified by any quantitative measurement, because the landslide movement is not measured there systematically.

6. Interferometry with TerraSAR-X

High resolution passive satellite earth observing systems have been known for the last 8 years. Images from IKONOS and Quickbird have become a standard for many photogrammetric applications. The first available civilian high resolution SAR (Synthetic Aperture Radar) system is TerraSAR-X. It acquires data in different modes with different resolutions: SpotLight (1.3 m), StripMap (3.3 m) i Scan-SAR (14.8 m) and also polarimmetric mode (Fritz, Eineder 2008). In this study the interferometric capabilities of TerraSAR-X StripMap data were evaluated. The test was performed for the area affected by terrain subsidence caused by underground copper ores exploitation. For the same area Envisat ASAR C-band data were acquired covering approximately similar period of time.

Most of the InSAR processing algorithms were developed and optimized for widely available ERS–1/2, Envisat/ASAR and Radarsat–1 data. However, for the high resolution interferometry, in order to achieve optimal results, the coregistration approach that utilizes the imaging geometry information and a-priori information of the topography of AOI has to be used – the so-called DEM-assisted method, see e.g. (Huanyin 2004). Moreover, for processing TerraSAR-X SpotLight data the adaptive filtering needs to be performed in order to account for time-varying Doppler centroid frequency, specifically, (Kampes et al 2003) optimization of the image resampling and spectral filtering algorithms is needed.

All the optimizations and new algorithmic development for processing SpotLight and StripMap TerraSAR-X data, including reader of COSAR format, are fully implemented and will be released in a new version of DORIS software (Marinković 2008).

In the case of Legnica-Głogów area two TerraSAR-X StripMap images were used (Tab. 1). For the coregistration of TerraSAR-X dataset both "conventional" and "dem-assisted" method were tested. In this case study with rather moderate topography, there was no significant difference between both methods observed. The final processing was completed with "conventional" method based on orbital geometry and estimation of the resempling polynomial through the correlation optimization procedure.

Legnica-Głogów Copper Mining Area has been already studied with SAR interferometry using ERS–1 data (Krawczyk, Perski 2000) and Envisat (Popiołek 2006). For the evaluation of TerraSAR interferometry a pair of Envisat ASAR data was acquired at the closest dates to TerraSAR-X acquisitions (Tab. 1).

Table 1. SAR scenes used for processing
Tabela 1. Sceny SAR użyte do przetwarzania

	TerraSAR-X scenes used for processing	
	1	2
Orbit No.	02408 ASCENDING	02575 ASCENDING
Acquisition date and time	2007–11–20 16:34:52	2007–12-01 16:34:51
Incidence angle	31°	31°
Polarization	HH	HH
Temporal baseline	11 days	
Geometrical baseline (perpendicular)	–165.2 m	

Altitude of ambiguity	28.7 m	
	Envisat ASAR scenes used for processing	
	1	2
Orbit No.	29570 ASCENDING	30071 ASCENDING
Acquisition date and time	26-OCT–2007 23:38:44	17-DEC–2007 23:37:44
Incidence angle	23°	23°
Polarisation	VV	VV
Temporal baseline	35 days	
Baseline (perpendicular)	423.6 m	
Altitude of ambiguity	19.3 m	

Envisat data were also processed with DORIS software, and the topographic phase contribution was removed using the same DTED dataset as in the processing of TerraSAR-X data.

Even though high coherence of TerraSAR-X interferogram of the area was not expected, because of the sensitivity of short wavelength (X band) and relatively rural area, the computed TerraSAR-X interferogram presents very good coherence. TerraSAR-X data were acquired in November and December, i.e. during the season of lowest vegetation, which resulted in overall good coherence of the interferogram. However, the regional atmospheric effect should be noted (Fig. 2C).

The visual inspection of both, TerraSAR-X and Envisat interferograms, shows that the same areas affected by deformation (Fig. 7) could be identified. The location of subsidence phenomena is similar in both cases. Due to different temporal baseline and wavelength the measured subsidence ratios are different, but the question whether they are comparable remains.

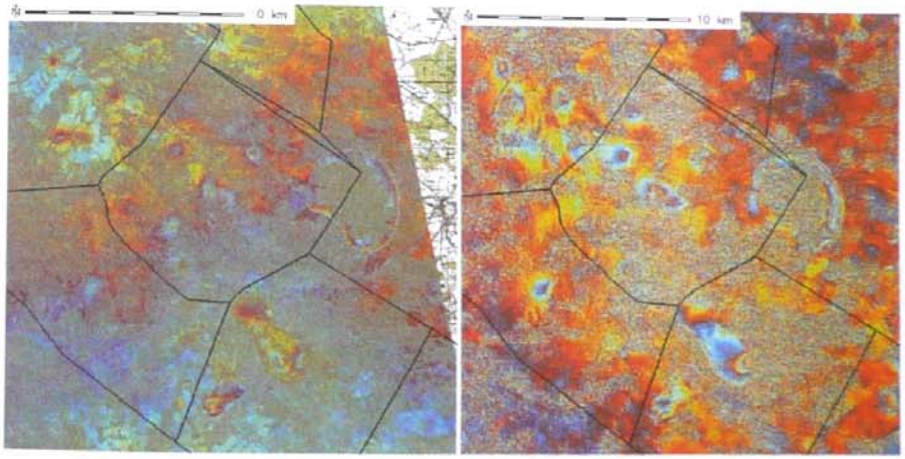


Fig. 7. TerraSAR-X (left) and Envisat (Right) interferograms of the same area In LGOM. Data geocoded to the state coordinate system "2000". In the eastern part of interferograms Żelazny Most waste pond is clearly visible (Perski et al. 2008)
Rys. 7. Interferogramy TerraSAR-X (lewy) I Envisat (prawy) dla tego samego obszaru z rejonu LGOM. Dane zostały przetransformowane do Państwowego Układu Współrzędnych 2000. We wschodniej części interferogramów widoczny jest zbiornik odpadów poflotacyjnych „Żelazny Most" (Perski et al. 2008).

The comparison between ASAR and TerraSAR-X was performed based on the deformation gradient obtained from different interferometric products. In order to compute it, for the selected subsidence bowls, the phase unwrapping was performed by the semi-manual approach taking into account coherence information, and considering a-priori subsidence information/prognosis. In this manner the deformation gradient was calculated. In order to do so, the interferometric phase difference between the highest subsidence in the center and the lowest subsidence at the edge of the bowl was taken. The distance between those points was about 500 m, thus atmospheric signal of long spatial wavelength was not affecting the measurements. To increase the (relative) accuracy and reduce the effect of noise, the phase values were collected over 3x3 window and were averaged. The window size was estimated empiricaly taking into account the small spatial size analysed subsidence bowl. The data was unwrapped over (pre)selected profiles using a-priori information on the expected subsidence. In order to make the phase measurements comparable, the gradients were recalculated into a vertical component of the subsidence velocity, by means of formulas as described in (Hanssen 2001). Results are presented in mm/day and discussed in Tab. 2 (Perski et al 2008). To calculate the subsidence velocity its linearity over the temporal baselines of interferograms was assumed. The experience with the analysis of subsidence within the area of study shows that the subsidence velocity variation becomes important within the frame of 3–4 monts.

Table 2. Results of subsidence velocity calculations for the selected subsidence bowls. Both SAR systems show very similar capability for subsidence velocity measurement. (Perski et al 2008)
Tabela 2. Wartości prędkości osiadania obliczone dla wybranych niecek. Oba systemy SAR prezentują podobne możliwości dla pomiarów prędkości osiadania (Perski et al 2008)

SAR system	*Interferometric phase gradient [radians]*	*Subsidence gradient along LoS [mm]*	*Subsidence velocity vertical component [mm/day]*
Gilów mine waste dump area			
TerraSAR-X	7.0542	17.402	**1.238**
Envisat	10.5674	43.092	**1.356**
Vicinity of Moskorzyn village			
TSX	2.1668	5.345	**0.416**
Envisat	3.8108	16.982	**0.446**

LoS – *Line of Sight*

The obtained values show that both radar systems could measure the same phenomena with the similar accuracy. However, it has to be stressed again, the analysis was performed using only a single interferogram with no removal of the atmospheric phase screen. Thus, these results should be considered as preliminary, however, they still could be used to indicate the expected precision and to evaluate the performance of TerraSAR-X for the deformation monitoring.

7. Conclusions

The preliminary PSInSAR processing on SAR data and its interpretations old not allow to clearly detect any recent tectonic movement. However, the „blind" experiments over relatively small areas proved high reliability of PSInSAR results. The obtained results show that the deformation signal detected within urban areas is very sensitive to non-tectonic phenomena like the ground water level changes.

The further works are focusing on the selection of larger and more representative crops, the alternative processing with different master scene for cross-validation of the results and the removal of periodic signal related to seasonal groundwater level change. The tests with other than linear model are planned in the near future. The cross-validation with other point wise InSAR techniques (e.g. StaMPS, SBAS) is also planned.

As for as the new high resolution sensors like TerraSAR-X are concerned, it is too early to draw any conlusion about their applications to natural surface movement detection. It was already proven that these data could be used for mining subsidence monitoring allowing the same subsidence phenomena detection as Envisat C-band SAR system, however, with more details than the previous system.

8. Acknowledgement

ERS–1 and ERS–2 SAR data used in this work are the courtesy of ESA project C1P.3915. The research was funded by Polish Ministry of Scientific Research (grant no. T12E04329). The research on TerraSAR-X data was performed thanks to funding provided by Polish Ministry of Science no. 4T12E 001 26.

9. Bibliography

Bamler R. and Hartl P., 1998, *Synthetic aperture radar interferometry*. Inverse Problems, 14: R1-R54.

Berardino P., Fornaro G., Lanari R. and Sansosti E., 2002, *A New Algorithm for Surface Deformation Monitoring Based on Small Baseline Differential SAR Interferograms*. IEEE Transactions on Geoscience and Remote Sensing, 40(11): 2375–2383.

Cacoń S., Dyjor S. and Kontny B., 2005, *The recent movements of earth crust surface in the Sudety and their foreland*, [In:] Z. W., P. B. and B. J. (Editors), IV Conference on Neotectonics of Poland: Active Faults of Central Europe. Polsh Academy of Science, Srebrna Góra 26–28.09, pp. 60–63.

Cattabeni M., Monti-Guarnieri A. and Rocca F., 1994, *Estimation and Improvement of Coherence in SAR Interferograms*. International Geoscience and Remote Sensing Symposium. Pasadena, CA, USA, 8–12 August 1994, pp. 720–722.

Colesanti C. and Wasowski J., 2006, *Investigating landslides with space-borne Synthetic Aperture Radar (SAR) interferometry*. Engineering Geology, 88: 173–199.

Colesanti C., Ferretti A., Novali F., Prati C. and Rocca F., 2003, *SAR Monitoring of Progressive and Seasonal Ground Deformation using the Permanent Scatterers Technique*. IEEE Transactions on Geoscience and Remote Sensing, 41(7): 1685–1701.

Defence-Mapping-Agency, 1996, *DMA Product Specifications for Digital Terrain Elevation Data Level 1 and 2*, MIL-PRF–89020A

Ferretti A., Prati C. and Rocca F., 2000, *Nonlinear Subsidence Rate Estimation using Permanent Scatterers in Differential SAR Interferometry*. IEEE Transactions on Geoscience and Remote Sensing, 38(5): 2202–2212.

Ferretti A., Prati C. and Rocca F., 2001, *Permanent Scatterers in SAR Interferometry*. IEEE Transactions on Geoscience and Remote Sensing, 39(1): 8–20.

Fritz T., Eineder M. (ed.) (2008), TerraSAR-X Ground Segment Basic Product Specification Document, TX-GS-DD–3302, http://www.dlr.de/tsx/documentation/SAR_Basic_Products.pdf

Fruneau B., Pathier E., Raymond D., Deffontaines B., Lee C. T., Wang H. T., Angelier J., Rudant J. P., and Chang C. P., 2001, *Uplift of Tainan Tableland (SW Taiwan) revealed by SAR interferometry Geophysical Research Letters*, v. 28, p. 3071–3074.

Gens R. and Logan T., 2003, *Alaska Satellite Facility software tools: Manual*. Geophysical Institute, UAF, 180 pages.

Ghiglia D. C. and Pritt M. D., 1998, *Two-dimensional phase unwrapping: theory, algorithms, and software*. John Wiley & Sons, Inc, New York.

Goldstein R. M., Zebker H. A. and Werner C. L., 1988, *Satellite radar interferometry: Two-dimensional phase unwrapping*. Radio Science, 23(4): 713–720.

GRASS-Development-Team, 2006, *Geographic Resources Analysis Support System (GRASS) Software*. ITC-irst, Trento, Italy. http://grass.itc.it

Hanssen R. F., 2001, *Radar Interferometry: Data Interpretation and Error Analysis*. Kluwer Academic Publishers, Dordrecht.

Hanssen R. F. et al., 2006, *Atmospheric Phase Screen (APS) estimation and modeling for radar interferometry*, Fourth International Workshop on ERS/Envisat SAR Interferometry, 'FRINGE05', Frascati, Italy, 28 Nov-2 Dec 2005, pp. 6 pp.

Henderson F. M., Lewis A. J., 1998, *Principles and applications of imaging radar. Manual of remote sensing*, 3rd ed, vol. 2. John Wiley, New York

Hilley G. E., Burgmann R., Ferretti A., Novali F. and Rocca F., 2004, *Dynamics of Slow-Moving Landslides from Permanent Scatterer Analysis*. Science, 304: 1952–1955.

Hooper A., Zebker H., Segall P. and Kampes B., 2004, *A new method for measuring deformation on volcanoes and other non-urban areas using InSAR persistent scatterers*. Geophysical Research Letters, 31: L23611, doi:10.1029/2004GL021737.

Huanyin Y. et al., 2004, *Sensitivity of topography of InSAR data coregistration*, ENVISAT & ERS Symposium, Salzburg, Austria, 6–10 September, 2004, 6 pp.

Jonsson S., 2002, *Modeling of Volcano and Earthquake Deformation from Satellite Radar Interferometric Observations*, Stanford University, 164 pp. pp.

Kampes B. M., Hanssen R. F. and Perski Z., 2003, *Radar Interferometry with Public Domain Tools*, Third International Workshop on ERS SAR Interferometry, 'FRINGE03', Frascati, Italy, 1–5 Dec 2003, pp. 6 pp.

Kortas G., 2007, *Surface displacements above the historical mine in Wieliczka*. Przegląd Górniczy, 3: 4–12.

Krawczyk A. and Perski Z., 2000, *Application of satellite radar interferometry on the areas of underground exploitation of copper ore in LGOM – Poland*, 11th International Congress of the International Society for Mine Surveying. AGH, Kraków, pp. 209–218.

Kowalczyk K., 2006, *New Model of the Vertical Crustal Movements in the Area of Poland*. Geodesy and Cartography, 32(4): 83–87.

Leijen F. J. v., Ketelaar V. B. H., Marinkovic, P.S. and Hanssen R. F., 2005, *Persistent Scatterer Interferometry: Precision, Reliability and Integration*, ISPRS Workshop, High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany, 17–20 May 2005, pp. 7.

Leijen F. J. v. and Hanssen R. F., 2007, *Persistent Scatterer interferometry using adaptive deformation models*, ESA ENVISAT Symposium, Montreux, Switzerland, 23–27 April 2007, 6 pp. CD-ROM ESA SP–636

Marinković P. (2008), *New developments in Earth Observation: using TerraSAR-X satellite data for monitoring geophysical parameters*. NAC–9, Nederlands Aardwetenschappelijk Congres, 18–19 March 2008, Veldhoven. – poster

Massonnet D. and Feigl K. L., 1998, *Radar interferometry and its application to changes in the earth's surface*. Reviews of Geophysics, 36(4): 441–500.

Perski Z., 2003, *InSAR and POLinSAR for land subsidence monitoring – a user perspective, Applications of SAR Polarimetry and Polarimetric Interferometry*. ESA, Frascati.

Perski Z., 2005, *Application of SAR imagery and SAR interferometry in digital geological cartography*, [In:] S. Ostaficzuk (Editor), The Current Role of Geological mapping in Geosciences. NATO Science Series IV. Earth and Environmental Sciences pp. 255–244.

Perski Z. and Jura D., 1999, *ERS SAR Interferometry for the Land Subsidence Detection in Coal Mining Areas, Earth Observation*. Earth Observation Quartery 63: 25–29.

Perski Z., Leijen F. v. and Hanssen R., 2007, *Applicability of PSInSAR for building hazard identification*. Study of the 29 January 2006 Katowice exhibition hall collapse and the 24 February 2006 Moscow Basmanny market collapse, ESA ENVISAT Symposium, Montreux, Switzerland, 23–27 April 2007, ESA CDROM SP–636.

Perski Z. and Mróz M., 2008, *Zastosowanie metod interferometrii radarowej InSAR do badania naturalnych ruchów powierzchni terenu w Polsce*. Projekt GEO-IN-SAR. Archiwum Fotogrametrii i Teledetekcji, 17.

Perski Z., Hanssen R., Wójcik A., Wojciechowski T., 2009, *InSAR analyses of terrain deformation near the Wieliczka Salt Mine, Poland*. Engineering Geology, Vol. 106, No. 1–2, pp. 58–67.

Popiołek E. (ed.) (2006), *Analysis of vertical displacements of mining terrains of KGHM „Polska Miedź S.A. with application of SAR interferometric techniques*. Stowarzyszenie Naukowe im. S.Staszica w Krakowie. [In polish, unpublished report].

Rosen P. et al., 2000, *Synthetic Aperture Radar Interferometry*. Proceedings of the IEEE, 88(3): 333–382.

Strozzi T., Wegmuller U., Werner C.L., Wiesmann A. and Spreckels V., 2003, *JERS SAR interferometry for land subsidence monitoring*. IEEE Transactions on Geoscience and Remote Sensing, 41(7): 1702–1708.

Wójcik A. and Mrozek T., 2002, *Landslides in the Carpathian Flysch*, [In:] J. Ciesielczuk and S.Ostaficzuk (Editors), Proceedings of the Tenth International Conference and Fieldtrip on Landslides (ICFL); Polish Lowlands – Carpathians – Baltic Coast, Poland, 6–16 September 2002, pp. 151–167.

Wright P. and Stow R., 1999, *Detecting mining subsidence from space*. Int. J. Remote Sensing, 20(6): 1183–1188.

Varnes D. J., 1978, *Slope Movement Types and Processes*, [In:] Schuster R.L. & Krizek R. J. (Editors), Landslides: Analysis and Control. Special Rep. 176. Transportation Research Board, Nat. Acad. of Science, Washington.

Streszczenie

Niniejszy artykuł stanowi przegląd obecnie stosowanych metod satelitarnej interferometrii radarowej (InSAR) oraz jej najnowszych osiągnięć w dziedzinie pomiarów deformacji powierzchni terenu. Omówiono klasyczną metodykę InSAR, jej ograniczenia i potencjał oraz najnowszą jej modyfikację jaką jest metoda interferometrii rozpraszaczy stabilnych (PSInSAR). W artykule przedstawiono również pierwsze wnioski z prób wykorzystania wysokorozdzielczych danych radarowych z satelity TerraSAR-X. Omówiono wyniki badań porównawczych interferogramów z danych TerraSAR-X oraz Envisat (dane średniorozdzielcze) dla tego samego obszaru i podobnego okresu obserwacyjnego. Metodykę InSAR i PSInSAR omówiono w oparciu o przykłady opracowań wykonanych przez autora z terenu Polski. W sensie poznawczym artykuł skupia się przede wszystkim na możliwościach pomiarów deformacji naturalnych wywołanych geodynamiką, hydrogeologią lub ruchami masowymi. W odniesieniu do danych wysokorozdzielczych przedstawiono deformacje wywołane eksploatacją podziemną rud miedzi. Dane wysokorozdzielcze są rejestrowane dopiero od kilku lat, co uniemożliwia ich wykorzystanie do badań geodynamicznych.

Współcześnie stosowana metodyka InSAR, a przede wszystkim PSInSAR umożliwia już wykrywanie i mierzenie deformacji terenu o przyczynach naturalnych, choć otrzymywane wyniki nie są jednoznaczne. Stwierdzane interferometrycznie prędkości deformacji nie zawsze odpowiadają mierzonym metodami geodezyjnymi. Dokładne poznanie przyczyn takiego stanu rzeczy wymaga wnikliwych badań porównawczych oraz metodycznych.

GEOINFORMATICA POLONICA
9: 2009

MARIUSZ SZUBERT[1]

PROBLEM OF THE PALAEO-WARTA RIVER VALLEY ON THE WOŹNIKI-WIELUŃ UPLAND IN THE LIGHT OF GEOSTATISTICAL SPATIAL ANALYSIS OF THE SUB-PLEISTOCENE SURFACE HYPSOMETRY

Key words:

geostatistics, kriging, palaeogeomorphology, elevation model, sub-Pleistocene hypsometry, palaeo-Warta River valley, Woźniki-Wieluń Upland

Abstract

Digital elevation model of the sub-Pleistocene surface on the Woźniki-Wieluń Upland was constructed using ordinary point kriging with the aim of verifying the existence of the buried valley of the palaeo-Warta River. The problem of its existence, often dealt with in literature, is of key importance in the restoration of the palaeogeographic evolution of the northern part of the Silesia-Kraków Upland. The application of a geostatistical model appropriate for the spatial structure of the data allowed to elaborate a digital elevation model of the buried surface, which seems to be the most accurate representation of palaeorelief. The models reveals the depressions described in literature. However, their pattern does not support the thesis that they are fragments of a large valley. Their shapes are characteristic of subglacial kettles, troughs and basins.

PROBLEM DOLINY PRA-WARTY NA WYŻYNIE WOŹNICKO-WIELUŃSKIEJ W ŚWIETLE GEOSTATYSTYCZNEJ ANALIZY PRZESTRZENNEJ HIPSOMETRII POWIERZCHNI PODPLEJSTOCEŃSKIEJ

Słowa kluczowe:

geostatystyka, kriging, paleogeomorfologia, model wysokościowy, hipsometria podplejstoceńska, dolina pra-Warty, Wyżyna Woźnicko-Wieluńska

Abstract

Metodą geostatystyczną – krigingiem zwyczajnym punktowym opracowano cyfrowy model wysokościowy powierzchni podplejstoceńskiej na Wyżynie Woźnicko-Wieluńskiej, w celu rozsądzenia, czy istnieje kopalna dolina pra-Warty. Problem ten, wciąż dyskutowany w literaturze, jest zarazem kluczowym w opracowaniu koncepcji rozwoju paleogeograficznego północnej części Wyżyny Śląsko-Krakowskiej. Zastosowanie modelu geostatystycznego właściwego dla struktury przestrzennej danych pozwoliło opracować cyfrowy model wysokościowy kopalnej powierzchni, który

[1] Institute of Geography, Pedagogical University of Cracow.

z dużym prawdopodobieństwem obrazuje paleohipsometri. Widoczne są na nim obniżenia opisywane w literaturze. Jednak ich układ nie jest potwierdzeniem tezy, że są to fragmenty dużej doliny. Kształtem odpowiadają one kotłom, rynnom i misom subglacjalnym.

Introduction

A buried valley of the palaeo-Warta River on the Woźniki-Wieluń Upland has been repeatedly described in geological and geomorphological publications as one of the major forms of pre-Pleistocene relief. However, no convincing evidence of its existence has been hitherto presented, as its course was reconstructed basing on various depressions found at various, often distant, locations. The lack of evidence for their intercommunication raises doubts about the existence of a buried pre-glacial valley of the Palaeo-Warta River in the sub-Pleistocene relief of the Woźniki-Wieluń Upland.

Geostatistical methods allow for the development of a credible digital elevation model of the sub-Pleistocene surface for a large area defined by the extent of boreholes and exposures of sub-Quaternary rocks. Such a model would allow for a unanimous interpretation of the buried depressions.

The aim of this paper is to present the results of an analysis of a digital elevation model of the sub-Pleistocene surface in the central part of the Woźniki-Wieluń Upland, constructed using geostatistical methods.

The digital elevation model has been constructed using all available data from borehole archival data and geological maps and it covers the whole area defined by the available data. It is also made as accurate as possible in order to make the palaeogeomorphological interpretation reliable.

The Woźniki-Wieluń Upland is the north-western part of the Silesia-Kraków Upland. It borders the Silesian Upland on the west and the Kraków-Częstochowa Upland on the south (Fig. 1). The study was done in the Górna Warta Depression and the Krzepice Depression, between Częstochowa and Krzepice. This fragment of the Upland was densely drilled for iron ore prospection by Przedsiębiorstwo Geologiczne in Częstochowa.

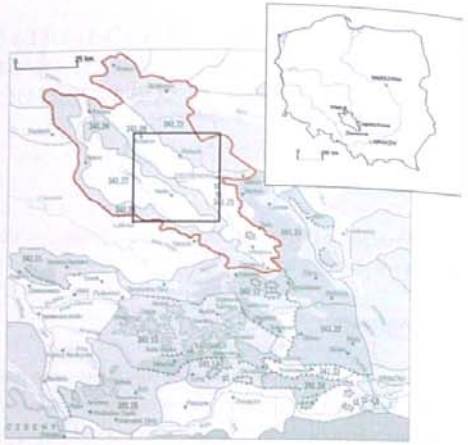


Fig. 1. Location of the study area with respect to the mezoregions of the Woźniki-Wieluń Upland according to J. Kondracki (2002). Mezoregions within the Woźniki-Wieluń Upland: 341.21 –Wieluń Upland, 341.22 – Liswarta-Prosna Depression, 341.23 – Woźniki Cuesta, 341.24 –Herby Cuesta, 341.25 Górna Warta Depression, 341.26 –Krzepice Depression.

Ryc. 1. Położenie terenu badań na tle mezoregionów Wyżyny Woźnicko-Wieluńskiej wg J. Kondrackiego (2002). Mezoregiony wydzielone w obrębie Wyżyny Woźnicko-Wieluńskiej: 341.21 – Wyżyna Wieluńska, 341.22 – Obniżenie Liswarty-Prosny, 341.23 – Próg Woźnicki, 341.24 – Próg Herbski, 341.25 Obniżenie Górnej Warty, 341.26 – Obniżenie Krzepickie.

Methods

The digital elevation model was constructed using ordinary point kriging. It is based on values of altitudes of Pleistocene base at sample points. The values are regionalized. Each of the values is a realization of a random function. In such a context, the geostatistical method is optimal for data manipulation.

In geostatistical (stochastic) approach, altitude, as a regionalized variable, has two aspects: random and structural. The random aspect accounts for local irregularities, while the structural one accounts for the large-scale tendencies (Matheron, 1989). Randomness of the analyzed variable is represented by fluctuations around the fixed surface – drift. The fluctuations are vanishing properties of the studied phenomenon, featuring their own structure (Wackernagel, 2003).

The geostatistical analysis of elevation of the sub-Pleistocene surface was performed in two stages. The first was geostatistical analysis of data, the second was spatial analysis (interpolation) using ordinary point kriging.

Geostatistical analysis

The aim of the geostatistical (structural) analysis was to present the spatial structure of the data set variability using an empirical semivariogram, and then to construct a geostatistical model of this structure. The analysis was preceded by the characterization of the source data and their statistical distribution.

Source data and their evaluation

The source data – altitudes of the sub-Pleistocene surface were obtained for 6499 sample points. Of these, 4862 values were calculated using borehole data and 1637 were read from the Detailed Geological Map of Poland, within the outcrops of sub-Pleistocene strata (Bardziński et al., 1982, Bednarek et al., 1987, Haisig et al., 1981a, 1981b, 1985, Kaziuk et al., 1986).

Statistical distribution of the sample data set is not favourable for geostatistical analysis because of the right-handed asymmetry, high concentration of values and high variability of data (Table 1).

Table 1. Parameters of statistical distribution of the data set
Tabela 1. Parametry rozkładu statystycznego zbioru danych

Frequency	Mean	Min.	Median	Max.	Standard deviation	Variance	Skewness coefficient	Curtosis coefficient
6499	260.35	166.58	260.99	365.53	22.98	528.12	0.0505	1.77

The sample points are unevenly distributed over an area of 1540 km². A distinct concentration is present in an NW-SE oriented belt, several kilometres wide, between Poraj, Częstochowa, Kłobuck and Krzepice (Fig. 2). It corresponds to the subcrop of iron ore-bearing Middle Jurassic clays to the sub-Pleistocene surface. The density of the sample points is variable. It is greatest within ancient mining fields, 78 points per 1 km² at Wręczyca. One hundred seventy points of the kilometric grid contain only one sample point and 770 fields contain none. On average, there are 4.2 points per 1 km². Distances between points vary from ca. 1 km to ca. 50 m.

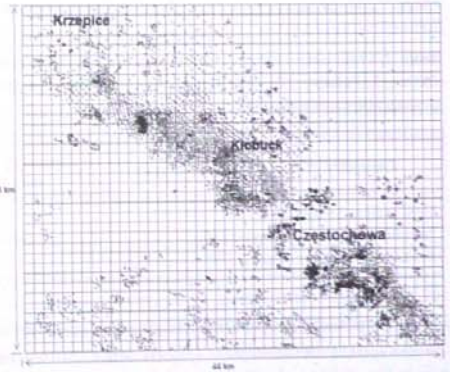


Fig. 2. Areal distribution of the sample points
Ryc. 2. Rozmieszczenie punktów próbkowych

The accumulated material was evaluated at the stage of data accumulation. The evaluation involved first the correctness of borehole locations, then the correctness of identification of the base Pleistocene sediments in the borehole sections.

The degrees of detail in lithological descriptions of the Pleistocene and Jurassic rocks are comparable. The depths down to the soles of penetrated strata in the whole sections were determined to an accuracy of 1 cm. The base of the Pleistocene sediments is easy to identify in almost all boreholes because they overlie limestones (Upper Jurassic), sandstones (Middle Jurassic) or firm dark micaceous clays, locally even black, with fauna.

These clays, known as „ore-bearing clays" (Middle Jurassic) contain deposits of iron ores. The upper ore horizon lies a few to several metres below the top of the clays.

De identyfication of the Pleistocene clays directly overlying Middle Jurassic clays is not difficult, because the Pleistocene clays are lighter-coloured than the Middle Jurassic ones, gray or light-gray, often varved, and include an admixture of quartz sand or gravel and pebbles of chert or limestone, often also northern (Scandinavian) rocks.

The correctness of identification of the Pleistocene base raised doubts in only few tens (ca 30) boreholes. The Middle Jurassic clays are overlain there by mixed Middle Jurassic and Pleistocene clays. The position of the Pleistocene base in such cases was identified by comparing sections in neighbouring boreholes and drawing a geological cross-section using them. When the interpretation was considered erroneous, a correction was introduced.

Borehole data are valuable evidence, suitable as a base for constructing a reliable digital elevation model of the buried sub-Pleistocene surface.

Calculation of isotropic empirical semivariogram

The spatial structure of altitude of the sub-Pleistocene surface is shown on an empirical isotropic semivariogram.

The measure of the variability in the spatial structure is provided by semivariance calculated for the semivariogram classes as a half of the mean of squares of the deviations of the studied parametre, in points distant by vector h

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2$$

where:

$N(h)$ – sample frequency,

x_i, y_i – values of the studied parameter (at the start and the end, respectively).

Parameters of the empirical isotropic semivariogram of the altitude values for the sub-Pleistocene surface are shown in Table 2.

Table 2. Main characteristics of the empirical isotropic semivariogram of the altitude values for the sub-Pleistocene surface
Tabela 2. Podstawowe charakterystyki izotropowego semiwariogramu empirycznego wysokości bezwzględnej powierzchni podczwartorzędowej

Geometry of the semivariogram grid	Frequency of the data set	Number of variogram classes	Class range [m]	Number of point pairs
Grid radius: 23 000 m Number of direction classes: 180 Number of distance classes: 100	6499	25	1000	14 841 515

Empirical isotropic semivariogram of the sub-Pleistocene surface elevation (Fig. 3) starts with a distinct discontinuity. At $h=0$ $\gamma(h)$ it attains the value of 36 m². A rapid rise in values of the semivariance function in successive distance classes follows up to the distance of 4100 m. The log is nearly rectilinear and steep over this distance. Farther, up to 7000 m, the value of the function becomes stable at 340 m². Over the distance of 7000 – 11,000 m the function value rises again, initially, to 9000 m, first slowly, then more intensely.

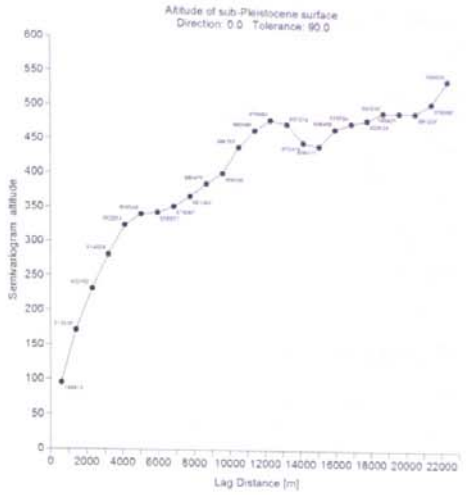


Fig. 3. Empirical isotropic semivariogram of the sub-Pleistocene surface elevation
Ryc. 3. Izotropowy semiwariogram empiryczny wartości wysokości bezwzględnej powierzchni podplejstoceńskiej

Within the distance range of 14,000 – 16,000 m the values of the semivariance function decrease to 440 m². They slightly rise again in the final section of the semivariogram, to 21,000 m, with two notable levels of stabilization — at 460 m² and 480 m². It rises to 550 m² in the two final distance sections of the semivariogram.

The empirical isotropic semivariogram shows the complex structure of the sub-Pleistocene surface elevation values. The structure involves:

– nugget variance (nugget effect) that equals 36 m² at h=0, which may partly result from measurement errors caused by errors in determination of the thickness of Pleistocene sediments, which resulted in errors of Pleistocene base elevations;

– distinct spatial correlation of data at distances between samples 1500m – 4200 m, 7000 m – 11,600 m and greater than 20,400 m;

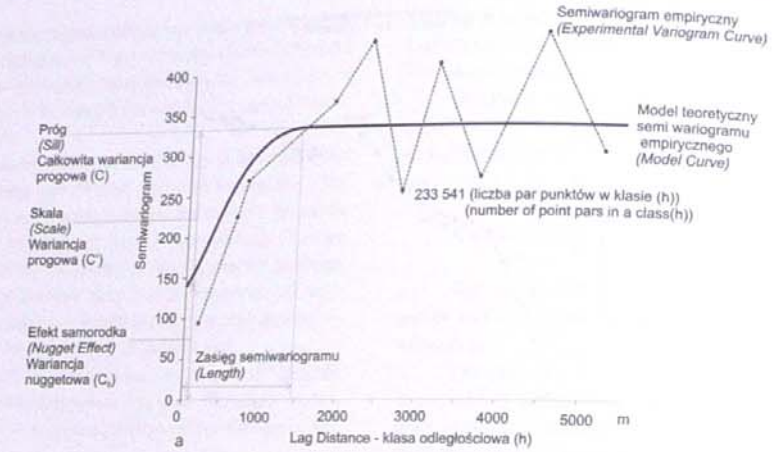
– stabilization of the semivariogram function values, expressed in flattening of the diagram at various, successively greater, values of the semivariance function, if the distance between samples equals 5000–6000 m, 12,000–13,100 m and 18,600–24,000 m;

– reversal of the growing trend of the semivariance function at the distances between samples equal to 13,100–15,000m.

The semivariogram is not smooth; it displays fluctuations. It also has no distinct flattening over a large distance. In general, the values of the semivariance function increase with increasing distance, with one exception when this trend is reversed. This is an unlimited semivariogram (Namysłowska-Wilczyńska, 2006). The persistent increase in semivariogram values with distance is interpreted as indicative of the presence of a trend or drift in the analyzed data.

Geostatistical model of variability

The next step in the geostatistical analysis of the empirical semivariogram involved approximation to an analytical function, which is a geostatistical model of the variance of studied parameter (Mucha, 1994), called also the theoretical model of the semivariogram or the variance structure model (Fig. 4).



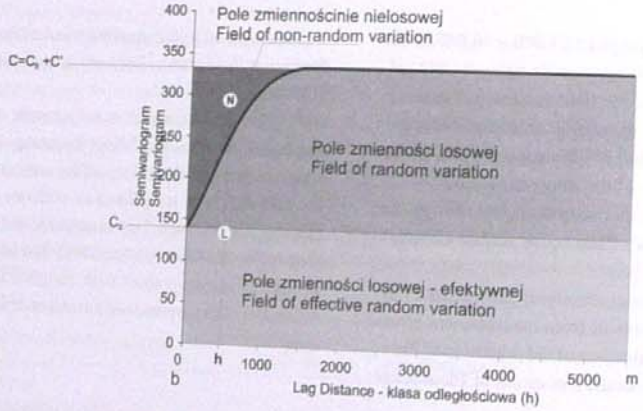


Fig. 4. Experimental semivariogram fitted using a theoretical model (a) (*Surfer 8*); variance structure of the studied parameter described by exemplary theoretical semivariogram (geostatistical model) after J. Mucha, (2002): C – total sill variance (semivariogram amplitude), C_0 –nugget variance (local variability of the parameter), C' – sill variance, L, N – random and non-random components of the studied parameter for distance h between sample points.

Ryc. 4. Semiwariogram empiryczny dopasowany za pomocą modelu teoretycznego (a) (*Surfer* 8); struktura zróżnicowania badanego parametru opisana przykładowym semiwariogramem teoretycznym (modelem geostatystycznym) za J. Mucha, (2002): C – całkowita wariancja progowa (amplituda semiwariogramu), C_0 – wariancja nuggetowa (zmienność lokalna parametru), C' – wariancja progowa, L, N – losowy i nielosowy składnik badanego parametru dla odległości, h między punktami próbkowymi.

After unsuccessful attempts at approximating the empirical semivariogram with single basic models, complex embedded models were used. These combined nugget effect models with one, two or three basic models (Table 3).

Table 3. Chart of parametric values of the geostatistical elevation models of the sub-Pleistocene surface
Tabela 3. Zestawienie wartości parametrów modeli geostatystycznych wysokości bezwzględnej powierzchni podplejstoceńskiej

Semivariogram model	Nugget effect (C_0) (Nugget variance) (Nugget Effect) [m a.s.l.]^2	Root of nugget variance $\sqrt{C_0}$ [m a.s.l.]	Scale (C') (Sill variance) (Scale) [m a.s.l.]^2	Root of sill variance $\sqrt{C_0}$ m a.s.l.	Sill $C=C_0+C'$ (Total sill variance (Sill) [m a.s.l.]^2	Tangent of slope of linear function (Slope)	Exponent of power function (Power)
Model (a): Nugget + linear	252	15.87				0.0129	
Model (b): Nugget + power	44.9	6.70	0.768	0.876	45.67		0.635
Model (c): Nugget + exponential + power	23.5	4.85	37.2 0.73	6.099 0.854	61.43		0.635
Model (d): Nugget + spherical + power	42.8	6.54	27.9 0.694	5.282 0.833	71.39		0.635
Model (e): Nugget + pentaspherical + power	65.36	8.08	9.2 0.847	3.033 0.92	75.41		0.625

Geostatistical modelling began with an approximation of the experimental model by a model consisting of a linear function (default model in GRID module of SURFER software) and a nugget model (Fig. 5a). Satisfactory results were obtained in successive attempts by combination of the nugget model with a power function (Fig. 5b) or with several basic models. The best fit of the theoretical and experimental models was obtained when the power function was one of the elements of the composite model (Fig. 5c–e).

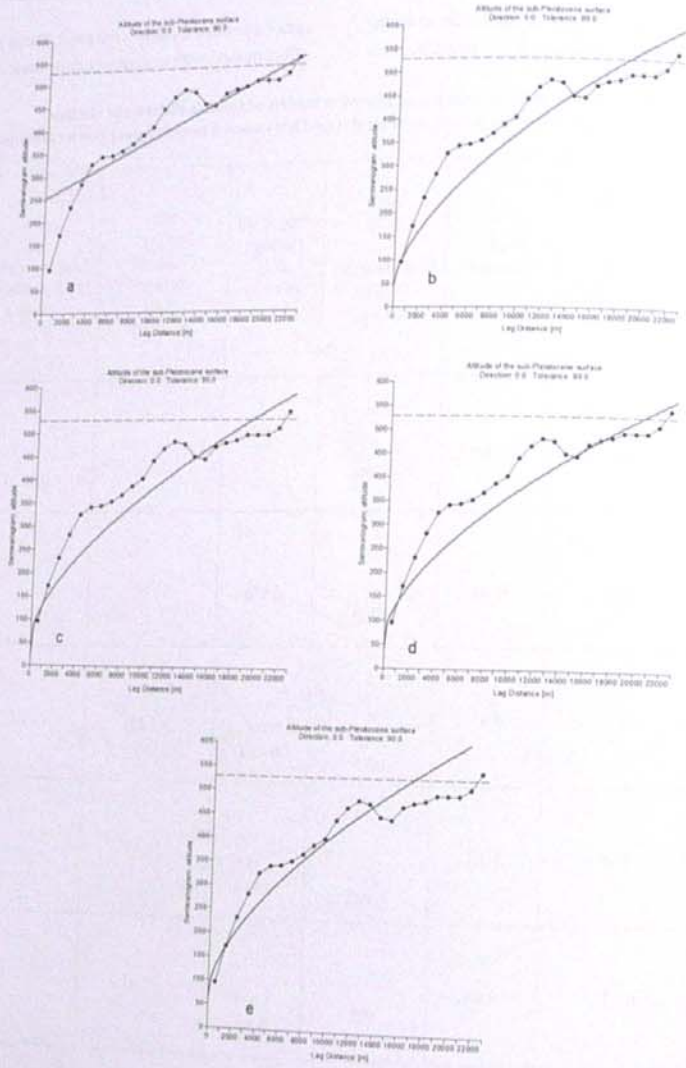


Fig. 5. Geostatistical elevation models of the sub-Pleistocene surface
a – nugget + linear model, b – nugget + power model, c – nugget + exponential + power model, d – nugget + spherical + power model, e – nugget + pentaspherical + power model.
Ryc. 5. Modele geostatystyczne wysokości bezwzględnej powierzchni podplejstoceńskiej
a – nugget effect (efektu bryłki) i liniowy, b – nugget effect i power (potęgowy), c – nugget effect, exponential (wykładniczy) i power, d – nugget effect, spherical (sferyczny) i power, e – nugget effect, pentaspherical i power.

Noteworthy in the tested models (except for model c) is the high value of the nugget variance – C_0 (random component of the variability) at h=0 (Table 3). It is greatest in model a, thus precluding its use in spatial analysis. The exponent of the power function in models b, c and d is the same; it is much lower in model e.

Cross-validation of the estimation and validation of the interpolation

The validation of the quality of estimation and interpolation was performed with the aim of selecting the best model of data variability for optimal interpolation of the sub-Quaternary surface elevation. The quality of fit of the theoretical model to the experimental one has a direct influence on the accuracy of the estimation and interpolation. This in turn influences the degree of conformance of the digital model with reality.

The validation was done in two steps. First the theoretical models were validated using cross-validation, then the quality of interpolation performed was assessed using the successive models.

The process of cross-validation consisted in the elimination of successive sample points $Z(x_\alpha)$ from the data set and in the estimation of $Z^*(x_{[\alpha]})$ values at the location of the eliminated sample point using the remaining data (Issaks and Srivastava, 1989). Then what followed was the calculation of the cross-validation error (true estimation error), that is the difference between the original value and the mean estimated value, showing to what degree the value of data fits the neighbourhood of surrounding data values (Namysłowska-Wilczyńska, 2006):

$$Z(x_\alpha) - Z^*(x_{[\alpha]})$$

where:

$Z(x_\alpha)$ – original value

$Z^*(_{[\alpha]})$ – mean value estimated at location x_α calculated after the elimination of the original valule.

The validation of the theoretical models of the semivariographs took into account: mean of cross-validation errors (ME), root of mean of square cross-validation errors (RMSE), standard deviation of kriging ($\sigma_{[\alpha]}$), kriging variance (σ_k^2) and coefficient of linear correlation (r) between data in sample points $Z(x_\alpha)$, and mean estimated values in these points $Z(x_{[\alpha]})$ (Table 4).

Table 4. Results of cross-validation of the estimation of elevation of the sub-Pleistocene surface
Tabela 4. Wyniki kross-walidacji estymacji wysokości bezwzględnej powierzchni podplejstoceńskiej

Estimation with semivariance model	Mean estimation error ME [m a.s.l.]	Mean square of estimation error MS [m a.s.l.]]^2	Root of mean square estimation error RMSE [m a.s.l.]	Standard deviation of kriging $\sigma_{[\alpha]}$ [m a.s.l.]	Variance of estimation error [m a.s.l.]]^2	Mean of standardized cross--validation error	Variance of standardized cross--validation error	Coefficient of linear correlation r between Z i Z^*
a	-0.025499	104.580894	10.226480	10.580244	104.580244	-0.002494	1.000006	0.885
b	0.063270	72.518652	8.515788	72.514649	72.514649	0.007430	1.000006	0.920
c	0.055100	74.016182	8.603266	74.013146	74.013146	0.006405	1.000041	0.919
d	0.049184	75.191711	8.671315	75.189292	75.189292	0.005672	1.000032	0.917
e	0.050173	74.746623	8.645613	74.744105	74.744105	0.005803	1.000034	0.918

The choice of the theoretical semivariogram model used in the kriging estimation was based on the principle that the chosen model should lead to the smallest difference between the real values Z (original data) and the mean estimated values Z^* (Namysłowska-Wilczyńska, 2006). Consequently, the most important were two indices: mean of cross-validation errors (ME) and root of mean of square cross-validation errors (RMSE). This semivariogram model was chosen, which resulted in their smallest values.

During the validation of the theoretical semivariogram models the cross-validation error was referred to the standard deviation of kriging ($\sigma_{[a]}$), which represents the assumed error, in order to compare the real and assumed error (Issaks and Srivastava, 1989, Namysłowska-Wilczyńska, 2006):

$$\frac{Z(x_a) - Z^*(x_{[a]})}{\sigma_{[a]}}$$

where:

$Z(x_a)$ – original value

$Z^*(x_{[a]})$ – mean estimated value at location x_a calculated after the elimination of the original value,

$\sigma_{[a]}$ – standard deviation of kriging.

Also the mean of square standard errors of cross-validation (variance of standardised error) was calculated (Namysłowska-Wilczyńska, 2006):

$$\frac{1}{n}\sum_{a=1}^{n}\frac{(Z(x_a) - Z^*(x_{[a]}))^2}{\sigma 2_{[a]}}$$

This index gives an insight into the adequacy of the theoretical model that approximates the course of the experimental semivariogram (Wackernagel, 2003). If the true error of estimation were on average equal to the assumed error, than the value of the index would be 1.

The quality of interpolation was estimated basing on statistical parameters of the set of estimated values in all nodes of the elementary grid (Table 5) and on the estimate of cartographic correctness of the digital elevation models (Fig. 7). Mean error of estimation was used as a criterion of the estimation:

$$Z - Z_{grid}$$

where:

Z – value measured in sample point (Z)

Z_{grid} – value interpolated in point Z

as well as the estimate of variance (s^2) and standard deviation (s) of the interpolation error.

Table 5. Statistical parametres of the set of estimated values in all nodes of grid.
Tabela 5. Parametry statystyczne zbioru wartości estymowanych we wszystkich węzłach siatki.

Interpolation with semivariance model	Mean [m]	Min. [m]	Median [m]	Max. [m]	Standard deviation [m]	Variance $[m]^2$	Variability coefficient	Coefficient of skewness
a	256.7020	197.68	258.3865	328.62	23.2601	541.0311	0.091	0.035
b	254.9486	182.77	257.3212	344.00	26.3190	692.6891	0.103	-0.121
c	255.1060	184.05	257.2364	342.19	26.0596	679.1013	0.102	-0.108
d	255.2074	185.17	257.2394	341.42	25.8778	669.6602	0.101	-0.098
e	255.1882	184.96	257.2425	341.67	25.9248	672.1947	0.102	-0.101

Taking into account the magnitude of the nugget variance and the results of cross-validation, the choice of a theoretical variability model was limited to models c and b. Model c (nugget + exponential + power) features the lowest nugget variance. Model b (nugget + power) has the lowest values of: root of mean square error of estimation (RMSE), standard deviation of kriging ($\sigma_{[a]}$), and variance of estimation error. This allows to accept the estimation using model b as the most accurate. Consequently, this model was chosen for the estimation of the values of elevation of the sub-Quaternary surface.

The choice of model b is also supported by the highest value of the linear correlation coefficient between the sampled and the estimated values. The adequacy of the choice of model b is also confirmed by the variance of the standardized cross-validation error, closest to 1 among the validated models. The mean of cross-validation errors close to zero indicates that there is no systematic overestimation or underestimation of the estimated values.

The mean of the values estimated using model b (Table 6) is lower by 4.48 m than the mean sample. The estimated minimum value exceeds the minimum sample value by 30.76 m. The maximum interpolated value is lower by 15.4 m than the lowest sample value. The distribution was thus „flattened", especially with respect to the minimum value. The „flattening" of distributions of the estimated values is greater when the other models are used.

Table 6. Selected statistical parameters of the set of estimated values Z^* of the sub-Pleistocene surface elevation
Tabela 6. Wybrane parametry statystyczne zbioru wartości estymowanych Z^* wysokości bezwzględnej powierzchni podplejstoceńskiej

Estimation with semivariance model	Estimated value Z^* [m a.s.l.]				Standard deviation Z^* [m]	Variance Z^* $[m]^2$
	Min.	Median	Max.	Mean		
a	210.09	258.81	328.57	259.41	17.71	313.84
b	197.34	259.37	341.43	259.49	19.41	376.68
c	198.70	259.32	340.76	259.49	19.25	370.64
d	199.53	259.25	340.28	259.48	19.15	366.70
e	199.44	259.26	340.35	259.48	19.18	367.97

The digital models (Fig. 6) were constructed using as a framework an elementary grid consisting of 6319 nodes distributed every 500 m. The mean error of interpolation equalled 0.231 m in all cases.

The variation of parameters of the sets of estimated values in all grid nodes is rather small, especially with respect to the estimation using models b, c, d and e (Table 5). Also the digital elevation models interpolated using them are similar in general outlines and in details.

A conclusion drawn from the evaluation of the quality of interpolation estimation is that only hypsometry interpolated using model a (Fig. 6a) differs essentially from those obtained using the other models. This is related to the model's markedly higher nugget variance.

Anisotropy of the elevation values of the sub-Quaternary surface

The first part of the geostatistical analysis was done using isotropic variograms. At the angular tolerance level of 90° they take into account the maximum number of sample point pairs in each class and their course is the most „smoothed". According to B. Namysłowska-Wilczyńska (2006) an isotropic averaged semivariogram provides the best estimate of the nugget effect (C_0) and sill variance (C'), it also facilitates the choice of the theoretical model that should „fit" the experimental curve.

The second stage of the analysis consisted in construction of a directional semivariogram. The initial study of anisotropy was made at experimentally established angular tolerance of the semivariogram of 40°, in four directions: 0°, 45°, 90° and 135° (Fig. 7a–d).

The elevation of the sub-Quaternary surface shows the best spatial correlation at azimuth 147°, semivariogram tolerance 40° and anisotropy coefficient 2 (Fig. 7e); the nugget variance equals 44.9 m^2, sill variance of the theoretical model – 0.768 m^2, the exponent of the power function 0.635, range 1.

Interpolation of the elevation of the sub-Quaternary surface

The choice of ordinary point kriging was dictated by the author's earlier experience in the application of computer-assisted methods of spatial data analysis (Szubert, 2004) and the other assesments of the ordinary kriging (Issaks and Srivastava, 1989, Wackernagel, 2003, Namysłowska-Wilczyńska, 2006). This method:

– belongs to the most frequently used kriging estimators,

– is a „strong" kriging system, accounting for the strong skewness of data distribution,

– when compared with other methods of interpolation it works well with unevenly distributed data; point

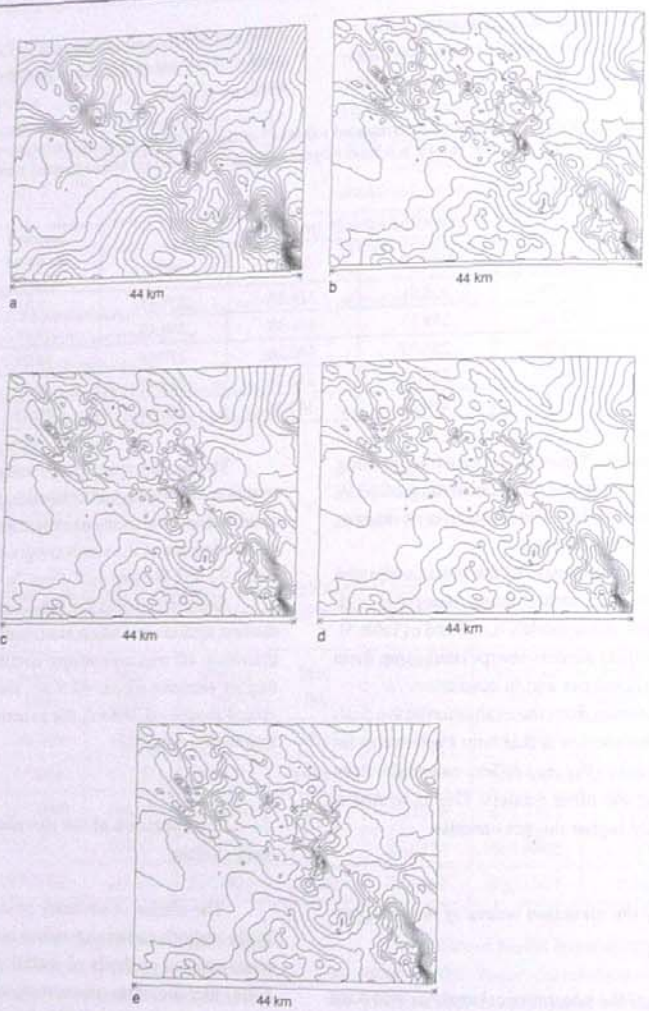


Fig. 6. Digital elevation models of the sub-Quaternary surface interpreted using various geostatistical models of data variability a – interpolation with model a (nugget + linear), b – interpolation with model b (nugget + power), c – interpolation with model c (nugget + exponential + power), d – interpolation with model d (nugget + exponential + power), e – interpolation with model e (nugget + pentaspherical + power).

Ryc. 6. Cyfrowe modele wysokościowe powierzchni podczwartorzędowej interpolowane z zastosowaniem różnych modeli geostatystycznych zmienności danych a – interpolacja z modelem a (nugget effect, linear), b – interpolacja z modelem b (nugget effect power), c – interpolacja z modelem c (nugget effect, exponential, power), d – interpolacja a modelem d (nugget effect, exponential, power), e – interpolacja z modelem e (nugget effect, pentaspherical, power).

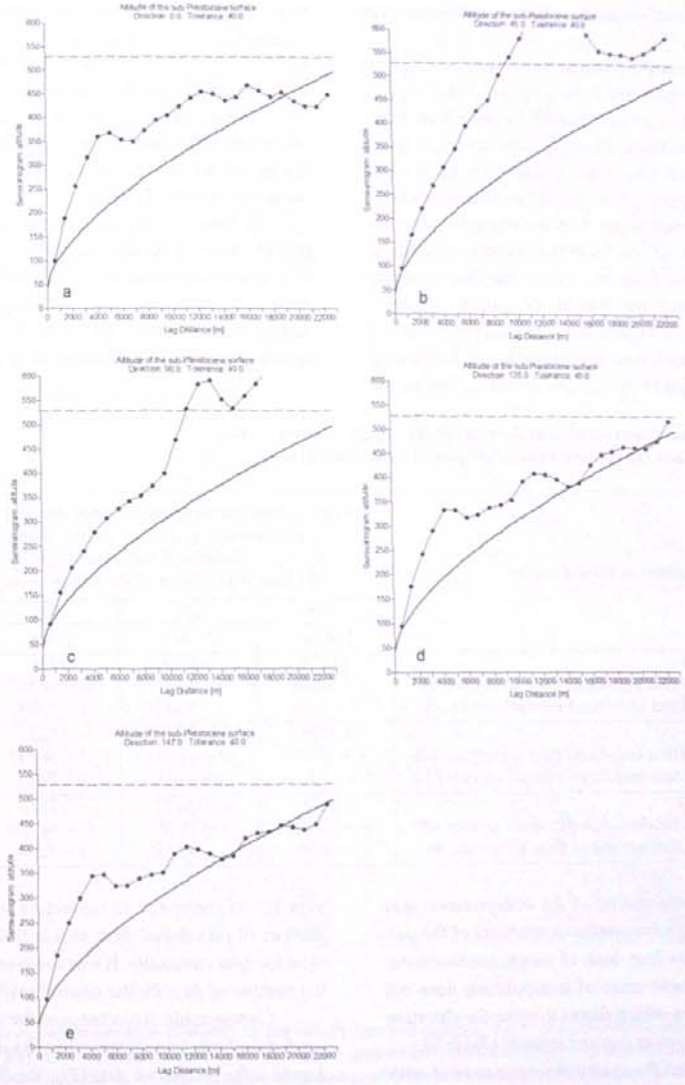


Fig. 7. Directional semivariograms of the elevation of the sub-Pleistocene surface. Angular tolerance of semivariograms 40°. Azimuths: a – 0°, b – 45°, c – 90°, d – 135°, e – 147°.

Ryc. 7. Semiwariogramy kierunkowe wysokości bezwzględnej powierzchni podplejstoceńskiej. Tolerancja kątowa semiwariogramów 40°. Kierunki: a – 0°, b – 45°, c – 90°, d – 135°, e – 147°.

concentration less influences the results of estimation than in other methods,

– it is a precise estimator, as is shown by the concordance between the estimated value *(Z*)* and data *(Z)*, when an estimated point (x_a) coincides with a sample point (x_a).

Moreover, estimation using ordinary kriging is based on a statistical description of data continuity by using a theoretical semivariogram model. The influence of the distances of the measurement points from the estimation point on the estimated mean *(Z*)* is lower than in other methods of spatial data analysis. Calculation errors are close to 0 and are distributed more evenly than in other kriging systems.

Geometry of the elementary grid and parameters of the area of point search were determined before performing the final interpolation of the sub-Quaternary surface elevation. Anisotropy of the data was reflected in the elliptical shape of the data area with semiaxes lengths R_1 (0X) and R_2 (0Y) equal to 35 000 m and 17 500 m, respectively, and semiaxis R_1 sloping toward 147°.

Another test performed was that of the accuracy of interpolation at various densities of the elementary grid and at various numbers of subsectors in the area searched for sample values (Table 7).

It was found that the value of error, hence the accuracy of interpolation, depends neither on the number of sectors in the area searched for sample values nor on the maximum and minimum numbers of sample points taken into account in the estimation of Z^* value, but rather on the distance between the nodes of the elementary grid (grid density).

Table 7. Mean errors of interpolation of the relief of the sub-Quaternary surface
Tabela 7. Średnie błędy interpolacji hipsometrii powierzchni podczwartorzędowej

Parameters of the search area	Distances between the nodes of elementary grid Number of nodes in elementary grid Mean error of interpolation [m] Variance of interpolation error (s^2) Standard deviation of the interpolation error (s)			
	100	250	500	1000
	154.791	24.957	6.319	1620
Number of sectors: **4** Maximum number of data introduced from all sectors: **64** Minimum number of data introduced from all sectors: **16**	0.0296 $s^2 = 33.506$ $s = 5.79$	0.0687 $s^2 = 43.717$ $s = 6.612$	0.231 $s^2 = 56.547$ $s = 7.520$	0.996 $s^2 = 85.028$ $s = 9.221$
Number of sectors: **8** Maximum number of data introduced from all sectors: **256** Minimum number of data introduced from all sectors: **32**	0,0229 $s^2 = 33.506$ $s = 5.79$	0,0625 $s^2 = 43.708$ $s = 6.611$	0,223 $s^2 = 56.517$ $s = 7.518$	0,983 $s^2 = 84.736$ $s = 9.205$
Number of sectors: **12** Maximum number of data introduced from all sectors: **600** Minimum number of data introduced from all sectors: **50**	0.0214 $s^2 = 33.506$ $s = 5.79$	0.0607 $s^2 = 43.711$ $s = 6.611$	0.221 $s^2 = 56.520$ $s = 7.518$	0.980 $s^2 = 84.741$ $s = 9.205$

Additionally, the quality of the interpretation was checked based on the cartographic correctness of the palaeoelevation contours (e.g. lack of loops, intersections, sharp bends). The mean error of interpolation does not exceed 1 m in all cases, which allows to refine the elevation model by using a narrower contour interval (Table 7).

The accuracy of the interpolation increased with increasing density of the elementary grid. The smallest interpolation error was obtained for a grid with 100 m distance between the nodes. The accuracy increased also with the increasing number of subsectors in the data search area. It was greatest at 12 subsectors and at the maximum number of introduced data, that is 600. Similar relations exist for data variability. It was lowest and independent of the number of data for the densest grid.

Cartographic correctness of the contour pattern depended on both grid density (Fig. 8a-d) and – to a lesser degree – the number of data (Fig. 8a, 8e-f). The model finally chosen was one interpolated on the basis of an elementary grid with 100 m node distance, estimated using the subdivision of the data searching area into 8 subsectors (Fig. 9 a).

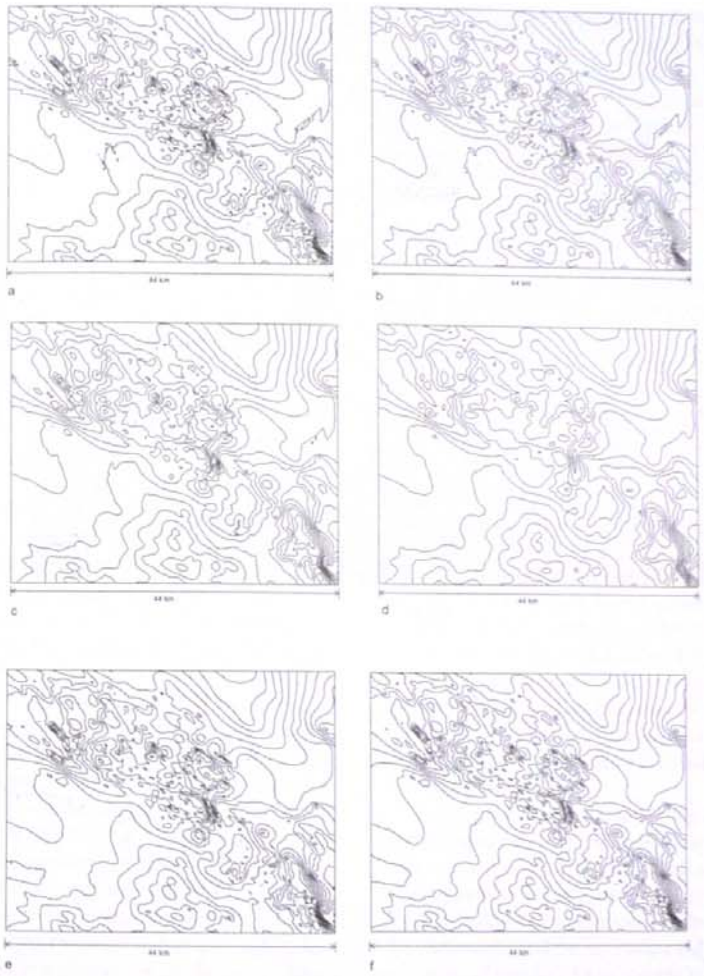


Fig. 8. Interpolation of palaeoelevation contours of the sub-Pleistocene surface: a-d – at 4 subsectors depending on the density of elementary grid, a – internodal distance of 100 m, b – internodal distance of 250, c – internodal distance of 500 m, d – internodal distance of 1000 m; e-f depending on the number of subsectors at internodal distances of 100 m, e – 8 subsectors, f – 12 subsectors.

Ryc. 8. Interpolacja paleoizohips powierzchni podplejstoceńskiej: a-d przy 4 podsektorach w zależności od gęstości siatki elementarnej a – odległości pomiędzy węzłami 100 m, b – odległości pomiędzy węzłami 250, c – odległości pomiędzy węzłami 500 m, d – odległości pomiędzy węzłami 1000 m; e-f w zależności od liczby podsektorów przy odległości pomiędzy węzłami siatki elementarnej 100 m e – 8 podsektorów, f – 12 podsektorów.

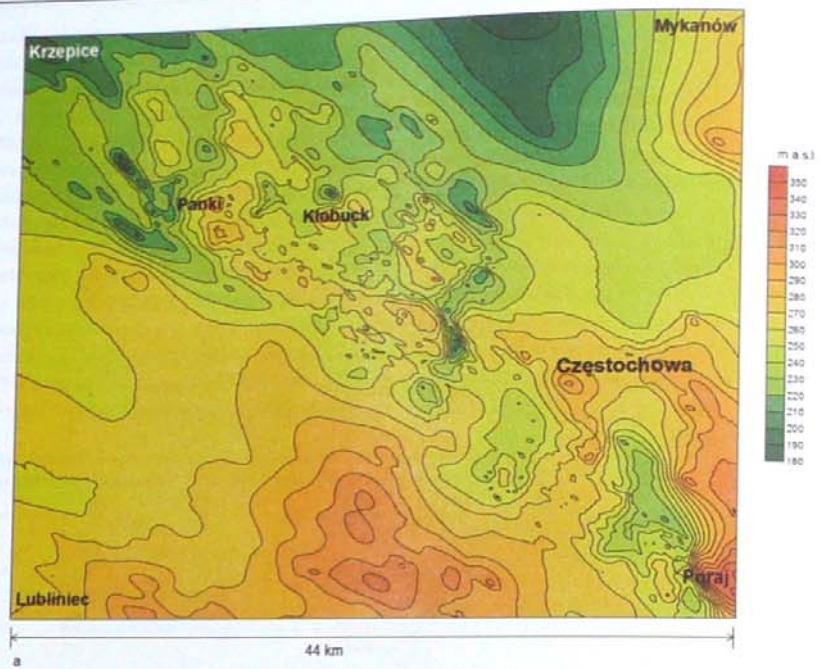


Fig. 9. Digital elevation model of the sub-Quaternary surface
a – digital elevation model, b – multidirectional geostatistical model, c –directional geostatistical model.
Ryc. 9. Cyfrowy model wysokościowy powierzchni podczwartorzędowej
a – cyfrowy model wysokościowy, b – wielokierunkowy model geostatystyczny, c – kierunkowy model geostatystyczny.

Discussion of results

The results of this study are visualized as a juxtaposition of the digital elevation model with an isotropic and a directional semivariograms. The spatial structure of the data displays a fluctuation of variable amplitude and drift. The spatial correlation of the data is disturbed. This may be interpreted as a manifestation of a local activity of one or several relief-forming factors that distorted the large-scale (prevailing over the whole area) features of the elevation pattern, shaped by a long-lasting geomorphological process. The spatial order (better correlation of data) is present in the NNE-SSW direction.

The last stage of shaping the sub-Pleistocene surface on the Woźniki-Wieluń Upland took place during the Odra Glaciation. It was then when the factor was active and distorted the relief characteristics shaped earlier, during the Mazovian Interglacial.

The Mazovian Interglacial was a period of a long-lasting denudation (Mojski, 2005), when the valley network was formed. The valleys had smooth long profiles. The smoothing of relief was interrupted by the invasion of the Odra ice sheet. Subglacial waters shaped troughs and basins, sharply marked in palaeorelief. The fluctuation in the spatial structure of data may record subglacial erosion. Subglacial troughs were an element of the valley network draining the subglacial waters during the pre-Warta interstadial.

Another stage of intense denudation followed during the Warta Stadial, when the northern part of the Upland was covered with an ice sheet. The studied area lay at its foreland. Periglacial climate intensified denudation of the depositional forms formed during the decay of the Odra ice sheet. Erosional subglacial forms were filled with periglacial sediments and so they were preserved.

This discussion is an attempt at explaining the fluctuation and drift basing on the previous knowledge on the palaeogeographical evolution of the northern part of the Kraków-Silesia Upland.

The digital elevation model (Fig. 9a) revealed a series of depressions in the sub-Pleistocene surface between Poraj and Krzepice. The performed validation of estimation and interpolation of the elevation of the sub-Pleistocene surface demonstrates that the pattern of these depressions, their shapes and sizes closely reflect the true relief.

The depressions south of Częstochowa were described by Z. Mossoczy (1955), who postulated that they are fragments of the pre-glacial valley of the upper Warta River (Figs 10, 11); a hypothesis accepted also in later works (Klimek, 1961, 1966). Similar depressions, also interpreted as fragments of buried valleys, are present in the northern part of the Woźniki-Wieluń Upland (Krzemiński, 1974) and west of Częstochowa (Lewandowski, 1993) (Figs 12, 13).

The digital elevation model (Fig. 9a) shows three separate groups of depressions: near Częstochowa, Kłobuck and Krzepice. They are incised into a plateau sloping to the northwest and north, toward an extensive depression.

The depressions near Częstochowa are clearly separated (Fig. 14c) by a ridge culminating in the Jasna Góra hill. They are shaped like wide bowls. The shallower one (30–35 m deep) is ovate, the deeper one (50–60 m) is elongated meridionally. They are probably a part of a greater system of depressions, as is shown by a trough branching to the south. Ovate and elongate secondary depressions are present in their bottoms.

One of the depressions, situated west of Częstochowa, resembles in its shape a classical subglacial trough (Fig. 14b). It consists of three parts. On the south it terminates in a bowl-shaped hollow with a radius of 2–2.5 km, ca. 35 m deep. The middle, meridionally elongated part of the trough is deepest (up to 70 m) and narrowest (1–2 km in the upper part). Secondary depressions are incised in the bottom. The third, northern part is up to 50 m deep and elongated in the NW-SE direction.

The depressions near Krzepice (Fig. 10a) are interconnected troughs with sharply outlined ovate or elongated hollows up to 50 m deep. The troughs are sloping to the northwest and north. Broad valleys with inner depressions in shape of subglacial bowls are present near Kłobuck (Fig. 9a).

The analysis of the sub-Pleistocene surface hypsometry does not support a thesis that the buried depressions between Poraj and Krzepice on the Wieluń Upland are fragments of the palaeo-Warta valley. The shapes and sizes of these forms support the thesis formulated by J. Lewandowski (1993) who, based on the lithological differentiation of the sediments filling some of the deep buried depressions, suggested that these are subglacial troughs or exaration depressions.

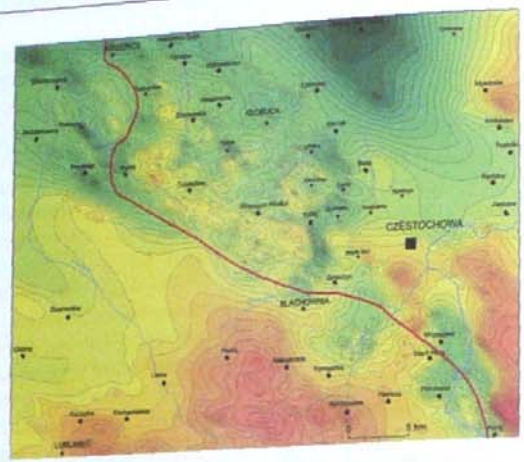


Fig. 10. Pre-glacial valley of the palaeo-Warta River after Z. Mossoczy (1955) against the digital elevation model of the sub--Pleistocene surface.

Ryc. 10. Preglacjalna dolina pra-Warty wg Z. Mossoczego (1955) na tle cyfrowego modelu wysokościowego powierzchni podplejstoceńskiej.

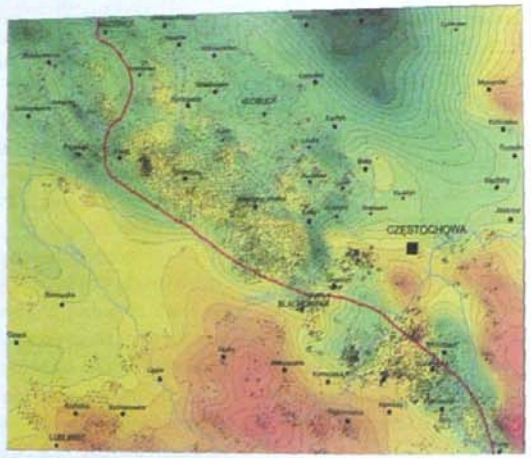


Fig. 11. Pre-glacial valley of the palaeo Warta River after Z. Mossoczy (1955) against the sample points used for construction of the digital elevation model of the sub-Pleistocene surface.

Ryc. 11. Preglacjalna dolina pra-Warty wg Z. Mossoczego (1955) na tle punktów próbkowych użytych do opracowania cyfrowego modelu wysokościowego powierzchni podplejstoceńskiej.

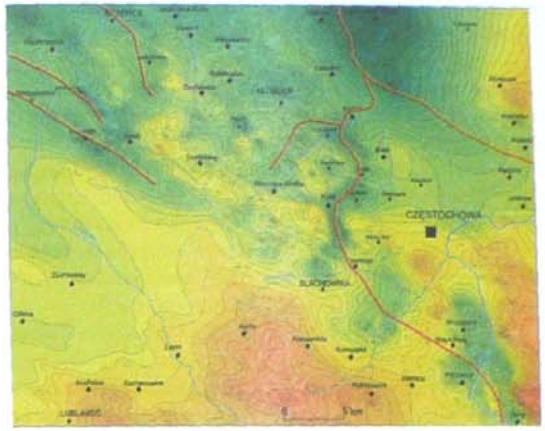


Fig. 12. Pre-glacial valley network after J. Lewandowski (1993) against the digital elevation model of the sub-Pleistocene surface.

Ryc. 12. Preglacjalna sieć dolinna wg J. Lewandowskiego (1993) na tle cyfrowego modelu wysokościowego powierzchni podplejstoceńskiej.

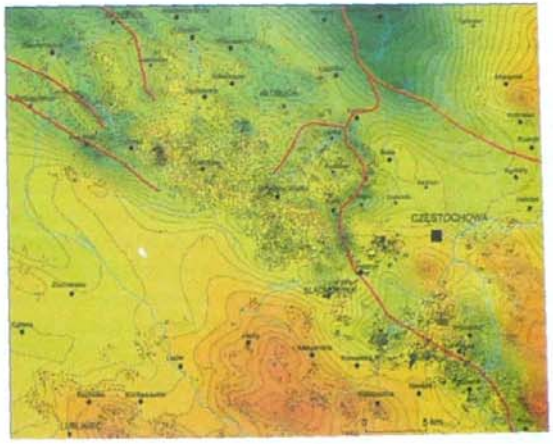


Fig. 13. Pre-glacial valley network after J. Lewandowski (1993) against the sample points used for the construction of the digital elevation model of the sub-Pleistocene surface.

Ryc. 13. Preglacjalna sieć dolinna wg. J. Lewandowskiego (1993) na tle punktów próbkowych użytych do opracowania cyfrowego modelu wysokościowego powierzchni podplejstoceńskiej.

The paper quoted above includes a palaeogeomorphological sketch of the sub-Pleistocene surface in the Silesia-Kraków region. A comparison of the Eopleistocene valley network in this scheme with the interpolated elevation model shows discrepancies (Fig. 12), which suggest that the depressions on the Woźniki-Wieluń Upland are not fragments of the Eopleistocene fluvial network. This conclusion is also supported by the pilot studies of the lithological differentiation of the sedimentary fill (Szubert, 2008a,b).

The distinguished depressions (Fig. 14a-c) should be considered as subglacial troughs formed during the Odra Glaciation – the last in this area. During the pre-Warta interstadial they were a part of the valley network.

The methodical aspect of the present paper provides an opportunity for reflection on the benefits of using for geomorphological purposes the geostatistical methods that seem to be difficult, complex and time-consuming when compared with other methods. A partial response was given in the paragraph explaining the choice of ordinary kriging as the method of interpolation.

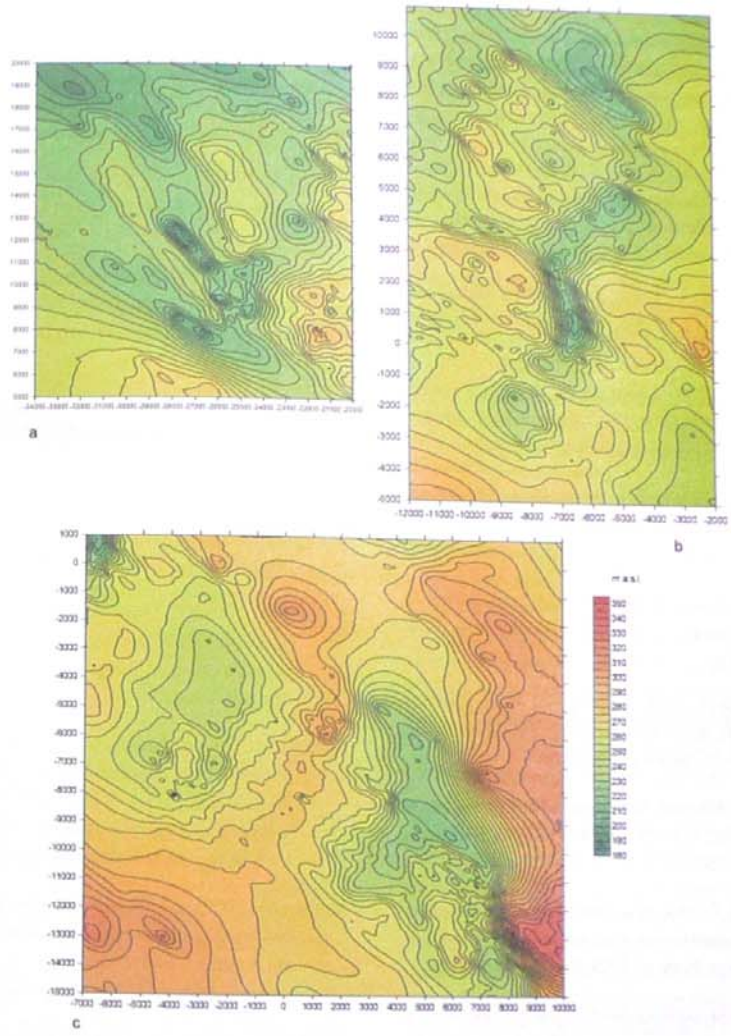
The advantages of geostatistical methods include:

– interpretation of the randomness of the phenomenon as fluctuation around drift, and fluctuations as vanishing properties of the phenomenon rather then errors, and acceptation of the drift as a reference surface for the fluctuations (Wackernagel, 2003). Multifactoriality of geomorphological phenomena and processes makes randomness their fundamental characteristic;

– structural analysis which allows to show in a semivariogram the spatial structure of regionalized data (studied phenomenon);

– data used in geomorphology are regionalized data. A analysis of their spatial structure may allow estimating the relative role of both, the structural and random components, showing to what degree they are correlated and to what degree chaotic. In geomorphology, it is the tendency to terrain planation, shaping of valleys with uniform slope. If a factor (one or more) disturbing denudation is (are) present, then structural analysis will reveal fluctuations and drift. It may be thus inferred to what degree is the analyzed hypsometry (relief) the result of local (random) processes. The structural analysis, by revealing the characteristics of the phenomenon which are not apparent in the hypsometrical pattern, is an important complement of a digital elevation model (Fig.9);

– possibility of using large data bases, conducting spatial analysis over a large area (interpolation and extrapolation of contour lines) and the valuation of the correctness of estimation and interpolation – judging from the accuracy of a digital model as a measure of its credibility. It is especially important when the object of study is not available for direct observation. Another conclusion stems from the comparison of Figures 10, 12 and 14. They show how the views on the pattern of buried valley network have evolved during the last 50 years with the growing data base, that is the knowledge on the relief of the buried surface and the development of interpolation methods from manual ones to those computer-assisted. A application of the latter, strengthened with geostatistical spatial analysis, allowed constructing a large-size credible image of palaeohypsometry, which led to the reinterpretation of the interpolated. In the case of the Woźniki-Wieluń Upland the geostatistical analysis indicated also a new dimension in the studies on buried relief – the evolution of this relief during the pre-Warta interstadial and subglacial erosion during the Odra Glaciation as important factors in shaping the valley network in the northern part of the Silesia-Kraków Upland.



Ryc. 14. Buried depressions on the Woźniki-Wieluń Upland
a – between Panki and Krzepice, b – between Konopiska and Czarny Las (west from Częstochowa), c – between Poraj and Częstochowa.

Ryc. 14. Kopalne obniżenia na Wyżynie Woźnicko-Wieluńskiej
a – pomiędzy Pankami i Krzepicami, b – pomiędzy Konopiskami i Czarnym Lasem (na zachód od Częstochowy), c – pomiędzy Porajem i Częstochową.

Reference

Bardziński W., Lewandowski J., Więckowski R., Zieliński T., 1982, *Szczegółowa Mapa Geologiczna Polski 1:50 000, ark. Częstochowa (845)*, Wyd. Geol., Warszawa.

Bednarek J., Haisig J., Lewandowski J., Wilanowski S., 1987, *Szczegółowa Mapa Geologiczna Polski 1:50 000, ark. Kłobuck (808)*, Wyd. Geol., Warszawa.

Haisig J., Wilanowski S., 1981a, *Szczegółowa Mapa Geologiczna Polski 1:50 000, ark. Boronów (844)*, Wyd. Geol., Warszawa.

Haisig J., Wilanowska H., Wilanowski S., Żurek W., 1981b, *Szczegółowa Mapa Geologiczna Polski 1:50 000, ark. Lubliniec (843)*, Wyd. Geol., Warszawa.

Haisig J., Wilanowski S., 1985, *Szczegółowa Mapa Geologiczna Polski 1:50 000, ark. Krzepice (807)*, Wyd. Geol., Warszawa.

Issaks E. H., Srivastava R. M., 1989, *An Introduction to Applied Geostatistics*. Oxford University Press, New York: 561 pp.

Kaziuk H., Nowak B., 1996, *Szczegółowa Mapa Geologiczna Polski, ark. Ostrowy (809)*, PIG, Warszawa, (maszynopis CAG PIG Warszawa).

Klimek K., 1961, *Morfologia przełomu Warty przez próg górnojurajski koło Częstochowy*. Przegl. Geogr., 33 (3): 421–442.

Klimek K., 1966, *Deglacjacja północnej części Wyżyny Śląsko-Krakowskiej w okresie zlodowacenia środkowopolskiego*, Prace IG PAN, Kraków, 53: 136 pp.

Krzemiński T., 1974, *Geneza młodoplejstoceńskiej rzeźby glacjalnej w dorzeczu środkowej Warty*, Acta Geogr. Lodz., 33: 171 pp.

Kondracki J., 2002, *Geografia regionalna Polski*, Wyd. Nauk. PWN, Warszawa: 440 pp.

Lewandowski J., 1993, *Rzeźba podczwartorzędowa region śląsko-krakowskiego i jej ewolucja morfogenetyczna*. Folia Quartern. 64: 101–121.

Matheron G., 1989, *Estimating and Choosing – An Essay on Probability in Practice*, Springer, Berlin: 141 pp.

Mucha J., 1994, *Metody geostatystyczne w dokumentowaniu złóż*, skrypty Wydz. Geologii, Geofizyki i Ochrony Środowiska AGH, Kraków: 155 pp.

Mucha J., 2002, *Struktura zmienności zawartości [Zn] i [Pb] w śląsko-krakowskich złożach rud Zn-Pb*, Studia Rozprawy i Monografie IGSM i E PAN Kraków, 108: 150 pp.

Namysłowska-Wilczyńska B., 2006, *Geostatystyka teoria i zastosowania*, Oficyna Wyd. Polit. Wrocł., Wrocław: 356 pp.

Surfer 8, Contouring and 3D Surface Mapping for Scientists and Engineers, Golden Software, Inc., Golden, 2002: 640 pp.

Szubert M., 2004a, *Cyfrowy model wysokości powierzchni podczwartorzędowej na Wyżynie Woźnicko-Wieluńskiej*. Archiwum Fotogrametrii, Kartografii i Teledetekcji, 13a: 233–242.

Szubert M., 2004b, *Utwory plejstoceńskie w kopalnych obniżeniach w południowej części Wyżyny Woźnicko-Wieluńskiej*, [w:] A. Kostrzewski (ed.), *Geneza, litologia i stratygrafia utworów czwartorzędowych*, t. IV, Wyd. UAM Poznań: 419–438.

Szubert M., 2008a, *Litologia osadów plejstoceńskich w obniżeniach w okolicach Częstochowy*, [w:] Kostrzewski A. (ed.), *Geneza, litologia i stratygrafia utworów czwartorzędowych*, UAM Poznań: 118–121.

Szubert M., 2008b, *Geomatyka w badaniach geomorfologicznych*, Dokumentacja Geograficzna, 37: 7–13.

Wackernagel H., 2003, *Multivariate Geostatistics An Introduction with Applications*, Springer, Berlin: 387 pp.

Streszczenie

W publikacji przedstawiono wyniki badania hipsometrii podłoża plejstocenu metodą geostatystyczną na Wyżynie Woźnicko-Wieluńskiej. Głównym problemem badawczym było rozstrzygnięcie, czy głębokie, kopalne obniżenia są fragmentami preglacjalnej doliny pra-Warty *(Mossoczy, 1955, Klimek, 1961, 1966)*.

Paleohipsometrię zrekonstruowano na podstawie 6499 punktów (4862 odwierty, 1637 punkty na Szczegółowej Mapie Geologicznej Polski) w południowej części Wyżyny Woźnicko-Wieluńskiej, na obszarze 1540 km^2 (ryc. 1, ryc. 2). W pierwszej części opracowania przeprowadzono analizę geostatystyczną danych, a w drugiej – metodą krigingu zwyczajnego punktowego wyinterpolowano wysokość bezwzględną podłoża plejstocenu i zobrazowano ją za pomocą cyfrowego modelu wysokościowego.

Analiza geostatystyczna ujawniła strukturę przestrzenną danych, której cechy odzwierciedla krzywa izotropowego semiwariogramu empirycznego (ryc. 3). Semiwariogram empiryczny jest nieograniczony oraz nie ma progu, co wskazuje na istnienie trendu lub dryftu w zbiorze danych. Krzywa empiryczna nie jest wygładzona, wykazuje fluktuacje. Cechami struktury przestrzennej wysokości bezwzględnej powierzchni podplejstoceńskiej są: wariancja nuggetowa, wyraźna korelacja przestrzenna danych na którą nakłada się stabilizacja wartości funkcji semiwariogramu oraz odwrócenie tendencji wzrostowej semiwariogramu.

Krzywa empiryczna wskazuje na znaczący udział czynnika losowego w kształtowaniu hipsometrii podłoża plejstocenu. Wyrazem tego są fluktuacje wartości semiwariancji oraz dryft w zbiorze danych. Stąd wniosek, że hipsometria prezentowana na cyfrowym modelu wysokościowym (ryc. 9) jest wynikiem działania procesu lokalnego (przypadkowego). Z uwagi na wyrazistość form erozyjnych takim procesem była erozja subglacjalna.

Krzywa empiryczna została przybliżona złożonymi modelami teoretycznymi (ryc. 5, tab. 3), z których na podstawie wyników kross-walidacji (tab. 4) oraz oceny interpolacji (ryc. 6, tab. 4) wybrano najlepiej „pasujący" do modelu empirycznego, zbudowanym z trzech modeli podstawowych: nugget effect, pentaspherical i power (ryc. 5e). Po zbadaniu anizotropii (ryc. 7) opracowano geostatystyczny model anizotropowy (ryc. 7e).

W powierzchni podplejstoceńskiej istnieją rynny oraz owalne zagłębienia o głębokości dochodzącej do 60 m. Grupują się one w okolicach Częstochowy (ryc. 14b, c), pomiędzy Pankami i Krzepicami (ryc. 14a) oraz w rejonie Kłobucka (ryc. 9a). Z uwagi na ich kształt oraz głębokość należałoby je uznać za rynny oraz misy subglacjalne, co potwierdza hipotezę postawioną przez *J. Lewandowskiego (1993)*. Ich geneza wiązałaby się ze zlodowaceniem odry– ostatnim na badanym terenie. Formy te wypełnione są zróżnicowanymi litologicznie osadami *(Szubert, 2004, 2008a, b)*, co wskazuje, że w schyłkowej fazie zlodowacenia odry oraz interstadiale przedwarciańskim były one elementem sieci dolinnej.

W czasie arealnej deglacjacji lądolodu odry poprzez obniżenia subglacjalne odpływały wody proglacjalne. Niektóre były jeziorami. W interstadiale przedwarciańskim rynny w okolicach Panek i Krzepic oraz Kłobucka zostały włączone w sieć dolinną odwadniającą południową część Wyżyny na północ i północny zachód. W czasie stadiału warty doliny wypełnione zostały osadami peryglacjalnymi. Badany teren znajdował się wówczas na przedpolu lądolodu, co sprzyjało intensywnej denudacji peryglacjalnej.

Ryc. 10, 12 i 14 ukazują rozwój poglądów w ciągu 50 lat na układ kopalnej sieci dolinnej, w sytuacji powiększania bazy danych, a w związku z tym wiedzy o ukształtowaniu kopalnej powierzchni oraz rozwoju metod interpolacji od manualnych po komputerowe. Zastosowanie tych ostatnich, wzbogaconych o geostatystyczną analizę przestrzenną pozwoliło opracować wielkoprzestrzenny, wiarygodny obraz paleohipsometrii skłaniający do reinterpretacji poznanych wcześniej *(Mossoczy, 1955)* obniżeń. Przeprowadzone badania nie potwierdziły tezy, że kopalne zagłębienia są fragmentami preglacjalnej doliny pra-Warty (ryc. 10, 12).

JADWIGA MACIASZEK[1]

THE ASSESSMENT OF THE USEFULNESS OF ARCHIVE MINING MAPS IN THE SPATIAL MANAGEMENT, FOCUS ON PIEKARY ŚLĄSKIE

Key words:

archive mining maps, system of co-ordinates *"Sucha Góra"*, information system MICARIS, filled shafts

Abstract

The paper presents the possibility of using archive mining maps in the supplementation of present mining and topographic maps with important situation details, e.g. filled mining shafts. The localization of 98 mining shafts was compared in the region of Piekary Śląskie to the made based on the archive maps from the inventory made by the mine. The system of co-ordinates *Sucha Góra-Soldnera (USG-S)* was characterized. In this system many mining maps (present and archival for Silesia) were made. The approximate parameters of the system were made. The analysis of the directions of the types of transformation from *USG-S* into national system was made. The necessity of making a special map of the filled mining shafts was indicated. This will be helpful in the management plans for of the areas of old metal mining in Poland. The information on archive mining maps can be easily found in the Internet in the MICARIS (Mine Cartography Information System – www.micaris.gis.edu.pl).

OCENA PRZYDATNOŚCI ARCHIWALNYCH MAP GÓRNICZYCH W ZAGOSPODAROWANIU PRZESTRZENNYM NA PRZYKŁADZIE PIEKAR ŚLĄSKICH

Słowa kluczowe:

archiwalne mapy górnicze, układ współrzędnych *„Sucha Góra"*, system informatyczny MICARIS, zasypane szyby

Abstrakt

Opracowywanie nowych map górniczych, topograficznych i mapy zasadniczej z zastosowaniem nowoczesnych technologii, a także zgodnie z wymogami wynikającymi z dyrektywy INSPIRE i właściwymi standardami ISO/CEN jest doskonałą okazją, aby odzwierciedlały one rzeczywisty stan zjawisk i procesów zachodzących – także w przeszłości – na danym terenie. Inwestorzy w Polsce coraz częściej natrafiają na zasypane szyby dawnego kopalnictwa, których nie ma na aktualnych mapach górniczych. Autorka zbadała pruskie mapy z atlasu *„Karte des Oberschlesischen Erzbergbaues"* z rejonu Górnego Śląska, uzyskując wysoką kartometryczność i zgodność faktograficzną tych map. Porównano lokalizację

[1] AGH, Wydz. Geodezji Górniczej i Inżynierii Środowiska, Katedra Ochrony Terenów Górniczych, Geoinformatyki i Geodezji Górniczej.

98 szybów znajdujących się na archiwalnych mapach z inwentaryzacją tych szybów, przeprowadzoną przez kopalnię, uzyskując średnią różnicę położenia około 10 m w rzeczywistości (1mm w skali mapy). Na mapach są także szyby obecnie nieznane. Opracowanie nakładki tematycznej z lokalizacją wszystkich szybów na podstawie badanych map pozwoli uniknąć w przyszłości wielu problemów inwestycyjnych. O archiwalnych mapach można dowiedzieć się więcej korzystając z internetowego systemu MICARIS (www.micaris.gis.edu.pl).

1. Introduction

Every investor carrying out their activities in post-mining areas should realize that new objects are under greater threat, especially in places of old shafts, liquidated workings of shallow exploitation, places where there were landslides in the past, where riverbeds were changed and hydrogeological conditions were modified etc. In many cases moving a new structure just several metres can improve technical conditions for the object and safety of its situation. The presented paper emphasizes the fact that archive documents including historical mining maps are very important in making decisions in the questions of environmental protection in the post-mining areas. It is easier to get them due to the internet system MICARIS (Mine Cartografy Information System www.micaris,gis.edu.pl). Proper use of the information included in old mining maps paves the way to progress and safety of many activities in mining, both in the rock mass and on the surface of post-mining areas. The MICARIS system – an on-line catalogue of archive mining maps – was made at the Faculty of Mining Surveying and Environmental Engineering, AGH-UST within a research project[2] the author was the manager of. It enables potential users to browse and search the resources and learn the conditions of using the service (Maciaszek J., 2009). At present there are about 2 500 archive descriptions in the MICARIS system in the Internet. Also there and in the paper by J. Maciaszek (2009) there are basic data and instructions how to seek necessary maps and information. Proper use of the information contained in old mining maps increases the safety of many mining activities both in the rock mass and on the surface of post-mining areas.

[2] Research project no. 4T12E05728: *System informacji o zbiorach archiwalnych dotyczących eksploatacji górniczych na terenie Polski* [Information system on archive collections referring to the mining exploitation in the area of Poland].

The presented paper emphasises the fact that in the questions of environmental protection and the management of post-mining areas, archival documents, including historical mining maps have a great significance in decision making. Gaining access to them, them is easier due to the Internet system MICARIS (Mine Cartography Information System).

In the areas of Bytom, Tarnowskie Góry and Piekary Śląskie, metal mining has been developed since prehistoric times. It went through different developmental periods. The fall of demand for metals caused the fall of mining, and the revival of demand increased mining. From the 18th century until the Second World War more mining fields, mines, shafts and metallurgy plants appeared in the described regions, which proves great demand for the resources.

The exploitation included the ores of iron, zinc and lead, and in small amount silver and copper. After the historical exploitation of ores thousands of shafts remained, yet, at the end of the 19th century and the beginning of the 20th century they were liquidated. The way of their liquidation was unknown due to the lack of reliable sources. The experience and inventory made by the mine [*Zbiorcze zestawienie...*] (tab.1) show that many of the liquidated shafts can pose threat to the ground surface and the objects located there. The examples were: flooding from the Szarlejka river in 1972 and a sudden disappearance of water by getting into the old mine workings; numerous cases of the activation of small shafts; the collapse of the floor in the building of the Technical High School in Bytom (Kubajak A. et al., 2007). Finding the location of every unknown working of shallow mining exploitation is helpful in the management of the post-mining areas.

2. Studies of the Usefulness of Archive Mining Maps from the Region of Piekary Śląskie

The problem of the usefulness of archive maps was presented taking as an example the region of Piekary Śląskie,

where old exploitation was very intensive and rich historic cartographic documentation exists. The subject of detailed analyses was establishing the location of old mining shafts.

To study the reliability, carthometric accuracy and the possibility of present application of archive mining maps, the comparison of the co-ordinates 98 shafts was carried out for the region of Piekary Śląskie. The co-ordinates were taken from the mine large-scale maps in the *Sucha Góra* system, with the situation of these shafts in the maps of the atlas *Karte des Oberschlesischen Erzbergbaues* (The Map of the Upper Silesia Ore Mining) later on referred to as *Erzkarte*.

The *Erzkarte* atlas consists of 26 maps with a scale of 1:10 000 and 1 review map with a scale of 1:100 000 (fig. 1) was issued by the Royal Higher Mining Office in Breslau in 1911-1912. The maps were made by Brück – a surveyor of the Royal Higher Mining Office. The following deposits were marked in the map: red and white calamine, sphalerite (Zn), galena (lead ore), iron ore and limonite (fig.1). The occurrence of mixed deposits and secondary deposits is marked on exploitation surfaces with points and lines of different colours, explained in the map legend. The other lines were marking definite and estimated borders of the occurrence of: metal-bearing dolomites, muschelkalk and other borders. The cover page and legend were written in German. There are points of the observation line in some sections of the maps, which shows that these maps were based on triangulation and altitude measurements, which is discussed in the further part of the article. Taking into account the size of a locality the following categories were distinguished: cities and quarters within cities, towns, villages, hamlets. Selected objects were also marked (among others churches, shrines, schools, windmills, watermills, railway stations and post offices). Hydrography was very poorly illustrated, only some rivers, wetlands and other water bodies were marked. Communication links are roads, railways and canals; in cities and towns streets, the names of which as well as other descriptions in maps are in German. The borders and names of mining fields were included, which enables the historians to know the property situation before 1912[3].

Fig. 2. marks the localities giving the names for individual maps presenting ore exploitation. The sheets of the maps with mining shafts used for this paper were marked with hatching. The reader can learn more about each map section after logging into the system using the following address: www.micaris.gis.edu.pl. Section maps have a very rich content, mainly geological, at each of them, on their bottom (sometimes also on the sides of the map) there are geological profiles with explanations. These maps should be interesting not only for communes having mining industry, but also for scientific centres included in the European project *GeoEurope 3D*[4] or making geoportal *IKAR*[5]. They give the possibility of systematic complementation of missing geological and mining data in the databases formed there according to the geostandards and prepared to make many analyses regarding changes in time.

[3] The papers containing the description of historic mining fields for hard coal based on archival mining maps, are for: the Cieszyn Silesia [4], Wałbrzych region [6], Dąbrowa Górnicza region [5].

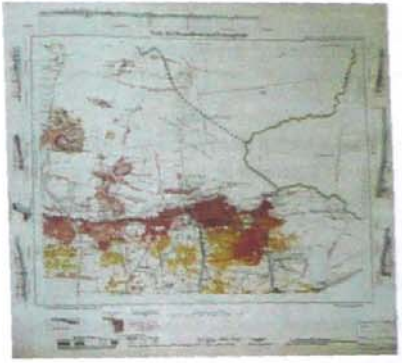


Fig. 1. The map from *Erzkarte* atlas No 16 presenting the area of Piekary Śląskie – Szarlej. Exploited ores: yellow – sphalerite; red – calamine; blue – galena; brown – limonite. The discontinuous line presents probable borders of the occurrence of the deposits, the continuous line – borders of mining fields. The maps from the atlas are in the archives of the Higher Mining Office and State Archive in Katowice

Rys. 1. Mapa z atlasu *Erzkarte* Nr 16 przedstawiająca obszar Piekar Śląskich – Szarleja. Oznaczenia rodzajów eksploatowanych rud: żółty – blenda cynkowa; czerwony – galman; niebieski – galena; brunatny – żelaziak brunatny (ruda darniowa). Linia przerywana oznacza przypuszczalne granice występowania tych złóż, zaś linia ciągła – granice pól górniczych. Mapy z atlasu znajdują się w archiwach: WUG oraz Archiwum Państwowym w Katowicach

[4] http://3d2f.com/tags/gis/geo/europe/excel/

[5] http://ikar.pgi.gov.pl/Portal/

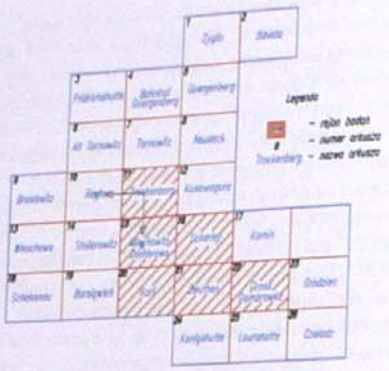


Fig. 2. Mutual situation of the sheets of maps in *Karte des Oberschlesischen Erzbergbaues*

Rys. 2. Wzajemne położenie arkuszy map atlasu *Karte des Oberschlesischen Erzbergbaues*

For the studies connected with the aim of this paper 6 maps were selected from the atlas. Their numbers were the following (compare fig. 2):

- Nr 11 – Trockenberg (Sucha Góra);
- Nr 15 – Miechowitz Dombrowa (Miechowice Dąbrowa);
- Nr 16 – Scharley (Szarlej – a part of Piekary Śląskie);
- Nr 20 – Karf (Karb);
- Nr 21 – Beuthen (Bytom);
- Nr 22 – Gross Dombrawka (Dąbrówka Wielka).

The Coal Company *Kompania Węglowa S.A.* (*Zakład Górniczy* "Piekary") carried out within the borders of mining areas "Rozbark III", "Piekary Śląskie" and "Brzeziny Śląskie" the inventory of most shafts, defining for them such parameters as: the name of the shaft, its function, depth, coordinates (some parameters were read from archive maps). after field investigation it was also established if the shaft poses threat for the surface (*Zbiorcze zestawienie…*). For these shafts a numeric database was made employing Microsoft Excel program. Due to much more accurate localization of shafts carried out by the mine, this set was treated as the base for the comparison with shafts that were put on the *Erzkarte* maps made with a scale of 1:10000. After the calibration and vectorization of the analyzed section of the atlas, the co-ordinates of shafts were read in the co-ordinates system "Sucha Góra", as well as their names and depths. The values were compared to the data obtained from the mine (a fragment of the database – without geographic co-ordinates shafts – is illustrated in Tab. 1.)

Table. 1. The fragment of the database from the inventory of shafts in the region of Piekary (source: autor's own work based on the maps of *Erzkarte* and materials from *Kompania Węglowa S.A.*)

Tab.1. Fragment bazy danych z inwentaryzacją szybów w rejonie Piekar (źródło: opracowanie własne na podstawie map *Erzkarte* oraz materiałów z Kompani Węglowej S.A.)

No.	Shaft No.	The Name of the Shaft		Shaft Depth	Threat for Surface According to the Assessment from the Mine	The Difference in the Situation of the Shaft [m]
		From the Mine	From the Maps			
1	Shaft 8	Jankiel	Jaeckel	6.0	No threat	18.8
2	Shaft 13	Adolf	Adolph	13.5	No threat	13.3
3	Shaft 24	Paweł	Paul	62.8	Possible threat	4.4
4	Shaft 27	Adam	Adam	41.6	Possible threat	7.9
5	Shaft 28	Paulina	Pauline	36.6	Possible threat	6.2
6	Shaft 29	Dorota	Dorotea	25.5	No threat	3.6
7	Shaft 30	Aleksander	Alexander	20.4	No threat	5.6
8	Shaft 32	Walery	Valerius	37.7	Possible threat	3.4
9	Shaft 34	Znajda	Fund	15.9	No threat	6.8
10	Shaft 35	Hugon	Hugo	37.2	Possible threat	6.7
11	Shaft 36	Konstantyn	Constantin	34.9	Possible threat	5.2

12	Shaft 41	Dreschler	Drechsler	36.1	Possible threat	15.7
13	Shaft 43	Agat	Edler	32.2	Possible threat	10.0
14	Shaft 45	Zachodni	West	30.0	No threat	2.4
15	Shaft 52	Wigury	Schmidt	85.0	No threat. The hole was filled.	8.7
16	Shaft 56	Jerzy	Georg	58.4	Possible threat	7.4
17	Shaft59	Książę	Furst Hugo	62.0	Possible threat	4.3
18	Shaft 60	Baildort	Baildort	52.6	Possible threat	1.3
19	Shaft 70	Żmija	Schlangen	32.8	No threat	4.0
20	Shaft 75	Karol	Scherbening	78.4	Possible threat	4.4
21	Shaft 76	Hajdy	Brauer	71.5	Possible threat	1.8
22	Shaft 80	Maria	Marie	9.6	No threat	3.4
23	Shaft 81	Adolf	Adolf	30.7	No threat	4.2
24	Shaft 82	Łazarz	Lazarus	43.2	Possible threat	14.9
25	Shaft 83	Walenty	August	72.2	No threat; filled before 1960	6.5
26	Shaft 99	Eliza	Elise	32.4	No threat	5.9
27	Shaft 101	Wojciech	Albert	59.9	Possible threat	0.3
28	Shaft 113	Graniczny	Grenz	97.3	No threat	5.6
29	Shaft 114	Powietrzny	Wetter	89.2	No threat	11.8
30	Shaft 118	Narutowicz	Holh	88.2	No threat	8.3
31	Shaft 120	Jacek I	Clothildes	108.7	No threat	5.6
32	Shaft 127	Walter	Walter	68.4	No threat	11.3
33	Shaft 128	Paweł	Paul	66.7	No threat	7.1
34	Shaft 129	Bethlen	Bethlen	62.6	Possible threat	2.9
35	Shaft 133	Jan	Johann	71.2	Possible threat	8.0
36	Shaft 134	Maszynowy	Maschinen	75.4	Possible threat	3.8
37	Shaft 135	Poszukiwawczy	Fund	70.8	Possible threat	4.2
38	Shaft 136	Helena	Helene	–	No threat	4.5
39	Shaft 137	Polowanie	Jagd	61.8	Possible threat	8.7
40	Shaft 139	Powietrzny	Wetter	–	No threat	5.0
41	Shaft 142	Witold	Abel	43.4	No threat	3.5
42	Shaft 149	Dołki	Maximilian	49.8	Possible threat	4.2
43	Shaft 152	-	-	32.5	No threat	9.0
44	Shaft 153	Marcin	Martin	24.0	No threat	4.6
45	Shaft 154	Poszukiwawczy IV	-	11.8	No threat	3.9
46	Shaft 156	Poszukiwawczy VII	Fund	27.2	No threat	2.5
47	Shaft 158	Skrajny	Roland	11.5	No threat	6.5
48	Shaft 162	Średni	Mittel	34.4	No threat	5.5
49	Shaft 163	Powietrzny	Luft	34.4	Possible threat	8.8
50	Shaft 164	Helena	Helene	21.8	No threat	4.9
51	Shaft 165	Stefan	Stephan	–	Possible threat	7.5
52	Shaft 166	Anna	Anna	37.4	Possible threat	2.6
53	Shaft 168	Wodny	Wasser	35.5	No threat	4.0
54	Shaft 169	Bieda	North	–	No threat	5.2
55	Shaft 170	Maria	Maria	16.3	No threat	1.8
56	Shaft 175	Olga	–	52.1	No threat	19.5

57	Shaft 176	Maurycy	Moritz	56.8	Possible threat	7.6
58	Shaft 177	Poszukiwawczy	Fund	18.0	No threat	10.8
59	Shaft 178	Poszukiwawczy IV	Vers	23.0	No threat	8.5
60	Shaft 179	Erazm	Erasmus	60.6	Possible threat	9.8
61	Shaft 180	Pomocniczy	Hilf	66.3	Possible threat	11.5
62	Shaft 181	Herman	Hermann	56.1	Possible threat	6.8
63	Shaft 182	Pomyślności	Gutglück	44.0	Possible threat	11.1
64	Shaft 183	Poszukiwawczy I	Vers	18.8	No threat	10.1
65	Shaft 184	Ludwik	Louis	26.5	Possible threat	7.0
66	Shaft 186	Gabriel	Gabriel	37.8	Possible threat	6.2
67	Shaft 187	Albert	Albert	33.8	Possible threat	8.5
68	Shaft 188	Paulina	Pauline	37.0	Possible threat	6.3
69	Shaft 189	Beniamin	Benjamin	38.0	Possible threat	8.0
70	Shaft 190	Jerzy	Georg	21.4	Possible threat	9.7
71	Shaft 191	Abel	Abel	45.8	Possible threat	10.3
72	Shaft 193	Józef	Friedrich	27.3	No threat	7.5
73	Shaft 194	Boża Pomoc	Gotthelf	26.2	No threat	3.2
74	Shaft 195	Ernestyna	Ernesthine	31.3	No threat	5.0
76	Shaft 196	Barbara	Barbara	34.0	Possible threat	7.6
75	Shaft 197	Poszukiwawczy X	–	36.6	No threat	9.6
77	Shaft 232	Jakub	Lieab	30.0	No threat	18.0
78	Shaft 234	Łąka	Wiese	17.2	No threat	12.8
79	Shaft 235	Mieszko	Otto Masch	47.0	Possible threat	17.0
80	Shaft 236	Poszukiwawczy	Wers	34.0	No threat	7.3
81	Shaft 240	Poszukiwawczy	Wers	30.0	No threat	7.3
82	Shaft 244	Udo	Udo	74.2	Possible threat	6.9
83	Shaft 249	Otylia	Ottilie	31.4	No threat	4.9
84	Shaft 250	Wschodni	Osten	33.4	No threat	22.4
85	Shaft 276	Shaft4	–	38.7	Possible threat	16.5
86	Shaft 277	Maszynowy	Masch	83.0	Possible threat	4.4
87	Shaft 278	Poszukiwawczy	–	31.5	No threat	4.9
88	Shaft 279	Maszynowy	Masch	38.4	No threat	8.2
89	Shaft 280	Olcha	Escher	27.3	No threat	7.8
90	Shaft 283	Karol	Carl	21.2	No threat	8.4
91	Shaft 284	Hugo	Hugo	21.7	No threat	1.3
92	Shaft 285	Gwidon	Guido	25.4	No threat	4.7
93	Shaft 286	S. I	S. I	25.0	No threat	4.7
94	Shaft 287	S.II	S.II	30.0	No threat	4.2
95	Shaft 298	Maciej	Mathiuss	87.8	Possible threat	2.4
96	Shaft 299	Zosia	Scotti	108.8	Possible threat	5.6
97	Shaft 306	Ryszard	Bertha	-	No threat	0.1
98	Shaft 308	Zbyszko	Wolfgang	79.2	No threat	7.4
			Mean Depth	**59.05**	**Mean Difference in Situation**	7.1

Table. 2. The comparative analysis of differences in the situation of shafts (differences in metres)
Tab. 2. Analiza porównawcza różnic położenia szybów (różnica położenia w metrach)

0 – 2 m	2-6 m	6 – 10 m	10 – 20 m	>20 m
6%	39%	37%	17%	1%

Table. 3. The comparison of depths of the analysed shafts
Tab. 3. Zestawienie głębokości analizowanych szybów

0 – 20 m	20 – 40 m	40 – 60 m	60 – 80 m	>80 m
10.5%	50%	13%	17%	8.5%

As it results from the carried out analysis (Tab. 1 and 2), a good compatibility was obtained referring to the location of shafts on both maps ranging from 0.1 m to 22.4 m – on average 7.1 m (10 m actually = 1mm in the map scale), which proves that archival maps fulfil their documentation role providing a lot of accurate and valuable information. Many shafts, as much as 50% of them, were between 20 and 40 m deep. This shows shallow exploitation of ores (Tab. 3). The zinc and lead ores are distributed very irregularly; the deposits can be in the shape of lenses, veins, pseudo-seams, thin bands and sometimes lumps of native silver, thus the exploited deposit often suddenly "disappeared". In the technique of metal mining a huge number of shallow shafts (often very small ones) was made to seek the deposit, which was cheaper than making searching excavations. The deepest shafts are 108.7m deep, because the ores are in the Triassic geological formation up to such a depth – to the limestone called basic limestone.

Here it is worth mentioning that in the studied maps shafts have depth descriptions (from 8.4 m to 76.0 m – on average 39.6 m, 27 other missing shafts do not have depth descriptions. These regions were marked in fig. 3. with a symbol with no number. On the other hand, the mine had many shafts which were not present on the analyzed maps. That means they came from the later period than the *Erzkarte* map were made, i.e. after 1912.

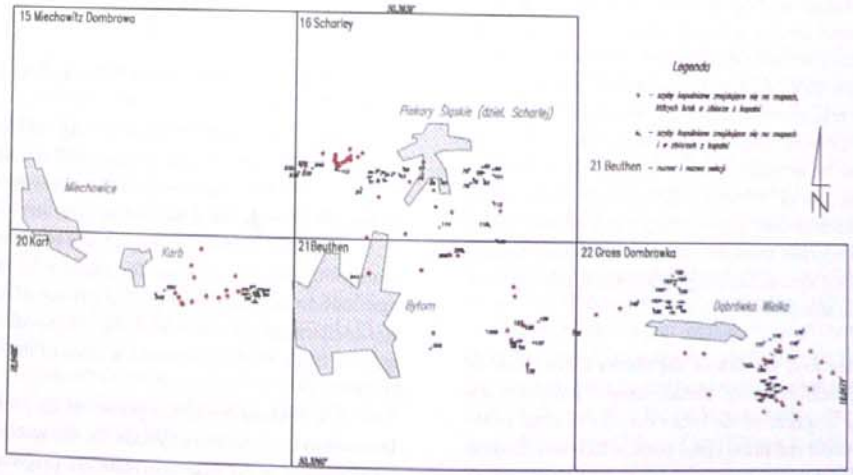


Fig. 3. The situation of the analysed shafts in the maps of the *Erzkarte* atlas (source: the author's work)
Rys. 3. Położenie analizowanych szybów na tle map atlasu *Erzkarte* (źródło: opracowanie własne)

3. The Analysis of the Impact of Co-ordinates Systems on the Location of the Filled Shafts.

A significant problem is to find the place where a given shaft was. Nowadays we can use GPS, provided we have geographic co-ordinates of the shaft (or co-ordinates according to a national reference grid).

3.1. A Local Co-ordinates System "Sucha Góra"

Erzkarte Maps were made in a system at local co-ordinates called "Sucha Góra" (later on referred to as USG). The studies were presented in papers [1,9,10]. So far there have not been exactly defined coefficients of the transformation from USG into national systems, thus it is not possible to automatically change into a geographic system. The problem of the transformation of co-ordinates from USG into system 1992 and 2000 is at present very important due to the Directive by the Polish Council of Ministers on the National System of Spatial Reference (Rozporządzenie...), according to which all basic maps and mining basic maps should, by the end 2009 be made in system 2000, while maps of smaller scales in system 1992. Nowadays mining maps are made on cartographic cards by hand and transformed into numeric maps. The rule was accepted that they would be made in the co-ordinates system 2000, uniform for the whole Poland.

At present many centres are working on a transformation coefficient, thus any new experience in this field should also be helpful for other researchers. It is worth mentioning that no Polish geodetic program containing transformation module between the systems of co-ordinates has transformation from the Sucha Góra system into any other national system, although many basic mining maps are made in this system.

In the first version of the Sucha Góra (USG-S) system, a modified Cassini-Soldner parallel projection was applied for a sphere of the mean curvature radius of the Bessel ellipsoid and in the initial point, which was the point of Sucha Góra (SG). In this projection, a German surveyor Soldner, at the end of the 18th century made a flat system of rectangular co-ordinates (fig. 4b) based on the ellipsoid system on the original – ellipsoid with spheroid co-ordinates (ξ, η) (fig.4a):

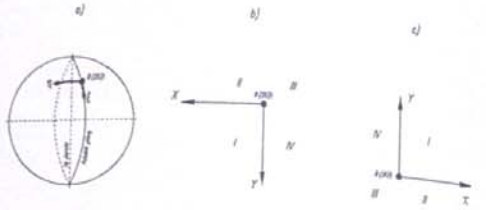


Fig. 4. Axes of co-ordinates in the Sucha Góra system: a) ellipsoid system (ξ, η), b) USG-S Soldner rectangular system, c) USG-G rectangular system in Gauss--Krüger projection.

Rys. 4. Osie współrzędnych w układzie Sucha Góra: a) układ elipsoidalny (ξ, η), b) USG-S prostokątny Soldnera, c) USG-G prostokątny w odwzorowaniu Gaussa--Krügera.

A flat system with ellipsoid co-ordinates (X_s, Y_s) on the projection field – later on referred to as Soldner co-ordinates – was designed in such a way that for any point $P(\xi, \eta)$ in the original and for each picture $P'(Y_s, Y_s)$ the following relation is true:

$$\xi = X_s, \quad \eta = Y_s$$

Historic roots of the flat USG-S system are shown by its axes:

- axis X overlapped with the image of the meridian going through point SG and was directed sounthwards.
- Axis Y, being the image of the surveying line going through SG was directed westwards (fig. 4b).

The points can have positive and negative co-ordinates. In the presented paper the quarters for this system were introduced to carry out the comparison of transformation coefficients for different signs of co-ordinates in respective quarters and to define the scale of the coefficient in directions.

The Soldner Projection preserves the proportions in distances in the direction of Y axis, i.e. the sections of bows perpendicular to the main meridian are projected without distorting the lengths. In the same way the main meridian is projected and the sections parallel to the main meridian

relatively quickly get elongated with the growing distance from this meridian. Also the growth of the deformations of angles and areas is quick. This projection was applied to mine surveying.

In the second version of (USG-G) system tangent Gauss-Krüger projection was applied to Bessel ellipsoid with the central meridian going threoug point SG. Axis X overlapped with the picture of the meridian and was directed northwards. Axis Y overlapped with the perpendicular line going through point SG and directed eastwards (fig. 4c). The co-ordinates in this version are later on referred to as Gauss co-ordinates X_G, Y_G in the "Sucha Góra" system The reference area in this projection version was an elliptic cilinder tangent to the surface of the Bessel ellipsoid in the initial point of SG. This is a multi-angle projection, which was applied to cadastre surveying work.

The transition between co-ordinates of both projections (from USG-G into USG-S) is made with the following formulae (Rajnich R. et al., 1986):

$$X_G = X_S$$

$$Y_G = Y_S + \frac{Y_S}{6N^2}$$

gdzie:

X_s, Y_s – Soldner co-ordinates,

X_G, Y_G – co-ordinates in Gauss-Krüger projection,

N – the radius of the curvature of the ellipsoid's meridian in the first vertical.

Geographic co-ordinates of SG point are according to two sources the same and equal to (Michałowski, Sikorski, 1932; Niemczyk, 1951):

B = 50°24'42.8922"
L = 18°52'39.9732".

The values of SG co-ordinates defined based on the contemporary measurements in EUREF – 89 system (Szafarczyk, Szymczyk, 2003) equal:

B = 50°24'38.90851";
L = 18°52'30.89102".

Differences in co-ordinates are mainly caused by different ellipsoids and their point of applications. In the first data the Bessel ellipsoid was applied with the point of application SG, in the second data – ellipsoid WGS84 with the centre of the Earth accepted for the calculations in such a way that it overlaps with the centre of the mass of the Earth and the polar axis with the "mean" rotation axis of our planet.

As has been mentioned earlier, the maps of the *Erzkarte* atlas were made in USG-S system. Situation measurements were based on triangulation of 1885 and 1901. The first triangulation is the Prussian Triangulation Network, which as the Higher Mining Office in Breslau ordered, was counted into USG-S. In the region of Rybnik this network was supplemented for the purpose of mining by Biler, a mine surveyor [Warchał B.]. The network supplementing the triangulation of 1885 was the triangulation of 1901, made for the needs of mining by the Trigonometric Division of the Royal Prussian of the Country Surveying Office on the request by the Higher Mining Office in Breslau. All the existing in this area points of the 1st and 2nd order of state triangulation were included into this network. The supplementary measurement allowed the definition of the co-ordinates of many points of the 3rd and 4th order. The co-ordinates of this triangulation were also determined in system USG-S.

It should be assumed that all the points of the control line in the maps of *Erzkarte* come from those triangulations, for which the catalogues with the co-ordinates can be found in Poland (also in existing mines, possessing mining maps made in system USG-S). To find the coefficients of transformation from this system and transfer into national systems it would be enough to find several permanent points from old triangulations, (well distributed to maximize the accuracy of the adjustment) and make the GPS measurement there. Unfortunately, after more than 100 years, these points are difficult to find, because most of them were destroyed. Even if some of them are found, one should realize that they could be displaced even by several metres due to the deformation of the surface in mining areas. The Region of the Upper Silesia Coal Trough is large, the range of the application of the USG-S system exceeds the area of 35 km × 40 km. We should realize that projection deformations at the margins of this system are larger than in its

central part. In transforming this system to state systems one should take into account the turning between these systems, as well as the co-ordinates of the length scale, the value of which can change, depending on the place and direction. As it is known – also from other systems of co-ordinates – projection corrections for the directions to the places are different. One can expect that in the future, in USG-S, the transformation would be made for smaller areas and for each from these areas, with their transformation coefficients, or from the tables with corrections, depending on geographic situation.

3.2. Co-Ordinates Transformations and the Analysis of the Influence of Co-Ordinates Systems on the Localization of Filled Shafts

For this article the author made a transformation based on several dozens of common relief features (joining points – *punkty łączne* – PŁ) that were identified based on *Erzkarte* maps and contemporary topographic maps made with a scale of 1:10 000 in system1942. The calculated co-ordinates of the shafts were used to bringing them on a topographic map (in system 1942, the transformation can also be performed in any national systems). All the work was carried our according to the following scheme:

1) scanning of archive and topographic maps;
2) calibration of archive maps in the USG-S system;
3) calibration of topographic maps in system 1942;
4) finding common terrain details on both maps and finding approximate transforming formulae in C-GEO program based on 76 points (Helmert transformation);
5) preliminary calculation of the corners of the archive maps into system 1942 based on the determined transformation co-ordinates and repeating the calibration archive maps in this system;
6) identifying the same situation points (based on archive and topographic maps) so that on each section of the archive maps there are several joining points, more or less equally distributed. In the case of a smaller number of points the tops of the grid of squares of system 1942 was used;
7) choosing 21 best joining points and determining the combined with the analysis of the deviations on the joining points and other points having co-ordinates in both systems) – testing;
8) defining coefficients of the transformation from USG-S into system 1942 and geographic system and calculating geographic co-ordinates for the examined shafts;
9) making a supplement with old shafts for the topographic map (the fragment of a new map is presented in fig. 5).

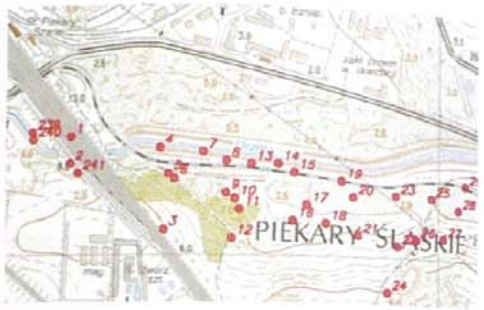


Fig. 5. The fragment of a topographic map with the analyzed and missing shafts (shafts numbers compliant whith Tab. 1)

Rys. 5. Fragment mapy topograficznej z analizowanymi i brakującymi szybami (numery szybów zgodne z tab. 1)

Below the comments resulting from the carried out works and analyses are presented:

1. The calibration of maps was made in CADRaster PRO version 7.1 of the Tessel Systems. The transformation of raster map (also called calibration) means their adjustment to the system of co-ordinates. It eliminates irregularities occurring in the raster picture, making the map cartometric. There are many calibration models. In the paper a two-cubic model was chosen (min. 16 vectors) getting mean-square errors (i.e. difference between the final destination point and the point calculated by the program) ranging from 0.568 to 1,645 m.
2. Preliminary Helmert transformation a (preserving angles) of the USG-S system into system 1942 for 76 joining points allowed the determination of the following co-ordinate of transformation:
   u = – 0.0286157496, v = -0.9991479827 with the mean transformation error m_T = 5.535 m and with adjustments on the adjustment points reaching: for co-ordinate X: 0.14÷14.66 m, and for co-ordinate Y: -0.20÷15.42 m. This transformation allowed choosing best 21 adjustment points for tests, because the errors of the co-ordinates of adjustment points are best seen in general (affine) transformation of the 1st degree of equal-angle transformation of the 1st degree (Helmert transformation). As we already know, the co-ordinates of joining points were determined graphically based on the 1:10 000 maps, thus they are loaded with considerable errors.
3. Test studies were to compare the accuracies of the directions in different kinds of the transformation (m_T) and their influence on the error in the point situation (m_P) after transformation. For this purpose, Helmert transformation was carried out (multi-angle conformal) and the affine transformation (multinomial, non- conformal). The problem appearing during the transformation of co-ordinates is establishing the degree of the transforming polynomial. With Geonet program of 2006 the transformations with combined polynomials were carried out, getting the following adjustment errors:
   – 1st degree polynomial: m_T = 3.92 [m]
   – 2nd degree polynomial: m_T = 3.58 [m]
   – 3rd degree polynomial: m_T = 3.62 [m]
   – 4th degree polynomial: m_T = 3.58 [m].

   The best adjustment and the smallest errors of transformed points were obtained by affine transformation of the 2nd degree. Polynomials of higher degrees can slightly change the transformation error, but increase the errors of transforming points. So the conclusion can be made: *there is no need to apply transformations of higher degrees, when the co-ordinates of adjustment points are not very accurate.*

   • Using the determined coefficients of affine transformation of the 2nd time from USG into system 1942, the co-ordinates of shafts were determined in this system and the shafts were marked in the topographic map;

   • In program C-GEO the co-ordinates of shafts were also calculated in a geographic system, but because the author has been continuing her research on marking many USG-S parameters, therefore Table 1 does not contain their[6], co-ordinates.

   • The author carried out the comparison of the co-ordinates of 8 points for which the co-ordinates in USG-S were known and on which the GPS measurements were made obtaining geographic co-ordinates on the ellipsoid WGS-84 (*Sprawozdanie…*). The mean error of the situation of 8 comparative points was 5.46 m, which in the map with a scale of 1:10 000 (used for the studies) is a satisfactory result.

[6] Further studies on USG-S, include making a supplement with the points of old triangulations, reading geographic co-ordinates from the catalogues, and co-ordinates in USG-S from archive maps, measurements will be the topic of further publication of the author.

4. Closing Remarks and Conclusions

Making new numeric mining topographic maps and the basic map applying modern technologies as well as according to the Directive by the Council of Ministers (2000) and proper standards ISO/CEN is a good opportunity to make them reflect a real state of phenomena and processes taking place in a given area, also in the past. New techniques of collecting, processing and visualization of cartographic data allow their easy up-dating, transfer into the directions of processing systems, and consequently the possibility of their use by many interested people and business enterprises. The data presented in archive mining maps have not only historical value, as the carried out studies showed. They are good in the supplementation of present mining maps (with old, usually filled mining shafts of different size), and even topographic maps of Poland (by making a proper thematic adjustment).

The information on historic shallow exploitation should be in spatial databases of each mining commune and help in making up-to-date plans of land management. This will allow avoiding many investment mistakes and will contribute to defining the regions of possible impact of old mining exploitation. The necessity of making a register of old shafts can be illustrated by the quoute from an article of a daily newspaper "Trybuna" no. 229 of 30/09/2009 – *"Kłopotliwy szyb"* (A Troublesome Shaft): *„The research and security work is going on at an old 150-m mining deep shaft, accidentally discovered by the workers during the construction of the western section of the road in Zabrze. It makes a part of a key*

communication route for the whole urban region of the Upper Silesia.

The brick shaft has a diameter of about 6 m. This construction has never been marked on any contemporary maps or sketches. It can only be seen on Prussian mining maps issued on the turn of 19th and 20th centuries. They are (among others) in the archives of the Zabrze Coal Mining Museum. It also turned out that an identical shaft was located in tha area of the car parki of the near-by supermarket. It was only there where the ground collapsed a few years ago."

Based on the author's studies (exatcly these "Prussian maps" from this museum) the following detail conclusions can be made:

1. The accuracy of determining the location of the filled shafts (and boreholes) is, based on archive maps, about 10 m (1mm in 1:10 000 maps).
2. It is recommended to supplement present maps based on historic maps, because on present maps many liquidated shafts are missing. It is also possible to make a special mining map of the supplement for the topographic map showing old mining and give it to the interested mining communes.
3. It is recommended to make "topographic descriptions of filled shafts" based on the available maps and pass this information to a proper institution (e.g. to the archives of the Higher Mining Office in Katowice).
4. There are works being continued on getting more accurate parameters of the local system "Sucha Góra". It should be assumed that soon it will be possible to make transformations from a local system "Sucha Góra" into a geographic system, which gives base for the localization of every shaft in the area with GPS receivers and making detail studies, e.g. georadar, before a new investment is made.

Literature:

Jura J., Sikorska M., 2006, *O odtworzeniu układu odniesienia na mapach eksploatacji rudnych*. Zeszyty Naukowe Politechniki Śląskiej. Seria: Górnictwo z. 278. s. 181–190, Gliwice.

Kubajak A., Przywara Z., Komor J., Duda A., Skiba R., Mój H., 2007, *Górnictwo i tradycje górnicze w Piekarach Śląskich*, Wydawnictwo *Kubajak*.

Maciaszek J., 2009, *Zasady tworzenia i wyszukiwania metadanych w systemie informatycznym MICARIS (Mine Cartography Information System)*. Przegląd Górniczy, październik 2009, Katowice.

Maciaszek J., Szewczyk J. 2007, *„Archival mining maps of Upper Silesia and the Mine Cartography Information System MICARIS"*, [w:] Společnosti důlních měřiču a geologu. XIV Konferencja mierniczych górniczych, s. 1–18, Ostrava.

Maciaszek J., Szewczyk J. 2008, *Historyczne pola górnicze węgla kamiennego rejonu Dąbrowy Górniczej (w świetle zbiorów archiwalnych map górniczych z muzeum miejskiego „Sztygarka")*. Materiały VII Konferencji Naukowo-Technicznej Ochrona Środowiska na terenach górniczych, s. 143–154, Szczyrk.

Maciaszek J., Szewczyk J. 2009, *Historyczne pola górnicze w rejonie Wałbrzycha*, [w:] Materiały X Dni Miernictwa Górniczego i Ochrony Terenów Górniczych, s. 273–292, Kraków.

Michałowski J., Sikorski T. 1932, *Katalog punktów trygonometrycznych w granicach Rzeczpospolitej Polskiej. Biblioteka służby geograficznej*, tom 8, Warszawa.

Niemczyk O., *Bergmännisches Vermessungswesen – ein Handbuch des Markscheidewessens: Matematisch – markscheiderische Grundlage.*

Rajnich R., Siembab J., Sosna A., 1986, *Mapy górnicze. Część 1. Zasady teoretyczne oraz część opisowa*. Wyd. Śląsk, Katowice.

Szafarczyk A., Szymczyk M., 2003, *Wykorzystanie Giromatu 2000 w przeliczeniach między układami współrzędnych na Górnym Śląsku*. Geodezja, tom 9, zeszyt 2/1, Kraków.

Warchał B., *Transformacje płaskich współrzędnych Gaussa – Krügera w układzie Borowej Góry na sferoidalne Soldnera w układzie Suchej Góry i odwrotnie*. Praca niepublikowana.

Sprawozdanie z pomiarów nawiązawczych linii obserwacyjnych za pomocą techniki GPS w rejonie Piekar Śląskich. 2006. Praca niepublikowana. AGH – Kraków.

Zbiorcze zestawienie szybów oraz płytkiej eksploatacji w OG „ZG Piekary". Praca niepublikowana. Kompania Węglowa S.A. ZG Piekary.

Rozporządzenie Rady Ministrów z dnia 8.08.2000 w sprawie państwowego systemu odniesień przestrzennych. Warszawa. 2000.

Streszczenie

W pracy zaprezentowano możliwość wykorzystania archiwalnych map górniczych do uzupełniania aktualnych map górniczych i topograficznych o ważne szczegóły sytuacyjne, np. zasypane szyby górnicze. Porównano lokalizację 98 szybów górniczych w rejonie Piekar Śląskich wykonaną na podstawie map archiwalnych z inwentaryzacją wykonaną przez kopalnię. Scharakteryzowano układ współrzędnych *Sucha Góra-Soldnera (USG-S)* w którym wykonane są liczne mapy górnicze (aktualne i archiwalne dla Śląska) i wyznaczono przybliżone jego parametry. Przeprowadzono analizę różnych rodzajów transformacji z *USG-S* na układ państwowy. Wskazano na konieczność opracowania specjalnej mapy zasypanych szybów górniczych, która będzie pomocna w planach zagospodarowania przestrzennego terenów dawnego górnictwa kruszcowego na ziemiach polskich. Informacje o archiwalnych mapach górniczych łatwo można znaleźć w Internecie w systemie MICARIS (Mine Cartography Information System – www.micaris,gis.edu.pl).