

## PRÓBA WYKORZYSTANIA SIECI NEURONOWYCH W KOMPLEKSOWEJ INTERPRETACJI DANYCH ELEKTROMAGNETYCZNYCH NA PRZYKŁADZIE ZŁOŻA GRABOWNICA

**An attempt to apply artificial neural networks in integrated  
interpretation of electromagnetic data – the Grabownica hydrocarbon  
deposit case study**

**Wojciech MASTEJ**

*AGH Akademia Górniczo-Hutnicza,*

*Wydział Geologii, Geofizyki i Ochrony Środowiska; al. A. Mickiewicza 30, 30-059 Kraków;*

*e-mail: wmastej@agh.edu.pl*

**Treść:** W pracy przedstawiono próbę zastosowania metody rozpoznawania obrazów z użyciem sztucznych sieci neuronowych w kompleksowej interpretacji danych z sondowań magnetotellurycznych AMT (*Audio-frequency Magnetotellurics*) i CSAMT (*Controlled Source Audio-frequency Magnetotellurics*) ze złoża ropy i gazu Grabownica (dipol podłużny i poprzeczny względem linii profili pomiarowych). Obiektami klasyfikacji były pionowe pasy z przekrojów magnetotellurycznych, o szerokości 50 m, utworzone na bazie siatek interpolacyjnych. Strukturę obiektów formalnie niestrukturalnych wymuszono częściowo poprzez ich podział na: przypowierzchniową, wysokooporową strefę do 200 m n.p.m., strefę niskooporowych utworów uszczelniających oraz słabych maksimów – kominów dyfuzyjnych, od 200 do –100 m n.p.m. i najniższą, wysokooporową strefę złoża. Prawidłowe wskazanie rzeczywistych stref złożowych na bazie informacji pochodzących z często niejasnych i rozbieżnych między sobą obrazów z przekrojów magnetotellurycznych, świadczą o przydatności tej metody. Wskazano także dwie nieznanne strefy złożowe w NE części przekroju G4 i w SW części przekroju G2.

**Słowa kluczowe:** rozpoznawanie obrazów, sztuczne sieci neuronowe, magnetotelluryka, złoża węglowodorów

**Abstract:** The paper presents an attempt to apply the pattern recognition method using artificial neural networks in integrated interpretation of the magnetotelluric data AMT (*Audio-frequency Magnetotellurics*) and CSAMT (*Controlled Source Audio-frequency Magnetotellurics*), acquired from the Grabownica oil and gas deposit (longitudinal and transversal dipole in relation to measurement profile lines). Vertical belts of magnetotelluric cross-sections, each 50 m wide, obtained from interpolation grids, were the objects of classification. The structure of formally non-structural objects was partly imposed through division into: near-surface high-resistivity zone, to 200 m a.s.l., low-resistivity screens and weak maxima – diffusion chimneys, from 200 to –100 m a.s.l., and the lowest, high-resistivity deposit zone. The usefulness of the method was proven by correct identification of real deposit zones based on data often coming from ambiguous and incompatible magnetotelluric cross-sections. In addition, two unknown deposit zones were discovered in NE part of cross-section G4 and SW part of cross-section G2.

**Key words:** pattern recognition, artificial neural networks, magnetotellurics, hydrocarbon deposits

## WSTĘP

Prognozowanie stref złóż węglowodorów za pomocą wielu różnych atrybutów geofizycznych, uzyskanych podczas sondowań magnetotellurycznych 2D, narzuca konieczność łączenia informacji o złożu, zawartych w wartościach każdego parametru, poprzez żmudną interpretację przekrojów magnetotellurycznych. Interpretacja taka napotyka często trudności, zwłaszcza gdy na przekrojach pojawiają się zakłócenia, a między obrazami z przekrojów, uzyskanych różnymi metodami, brak jest zgodności (Wojdyła *et al.* 2009, Stefaniuk *et al.* 2010a, b, Stefaniuk 2011). Zastosowanie metody rozpoznawania obrazów, o ile zakończyłoby się sukcesem, usunęłoby w pewnym zakresie te przeszkody, ponieważ na podstawie wielowymiarowej informacji możliwe byłoby jednoznaczne zaklasyfikowanie stref jako perspektywicznych bądź nieperspektywicznych. Wiarygodne wskazywanie stref perspektywicznych jako miejsc, które należałoby poddawać dalszej analizie, pozwoliłoby traktować zaproponowaną metodę jako przydatną w kompleksowej interpretacji danych elektromagnetycznych. Standardowo (Fukunaga 1972) podczas rozpoznawania obrazów klasyfikuje się obiekty po uprzedniej ekstrakcji cech. W pracy przeprowadzono klasyfikację nadzorowaną (wzorcową) wcześniej zdefiniowanych obiektów za pomocą sprawnego narzędzia, jakim są sztuczne sieci neuronowe (SSN).

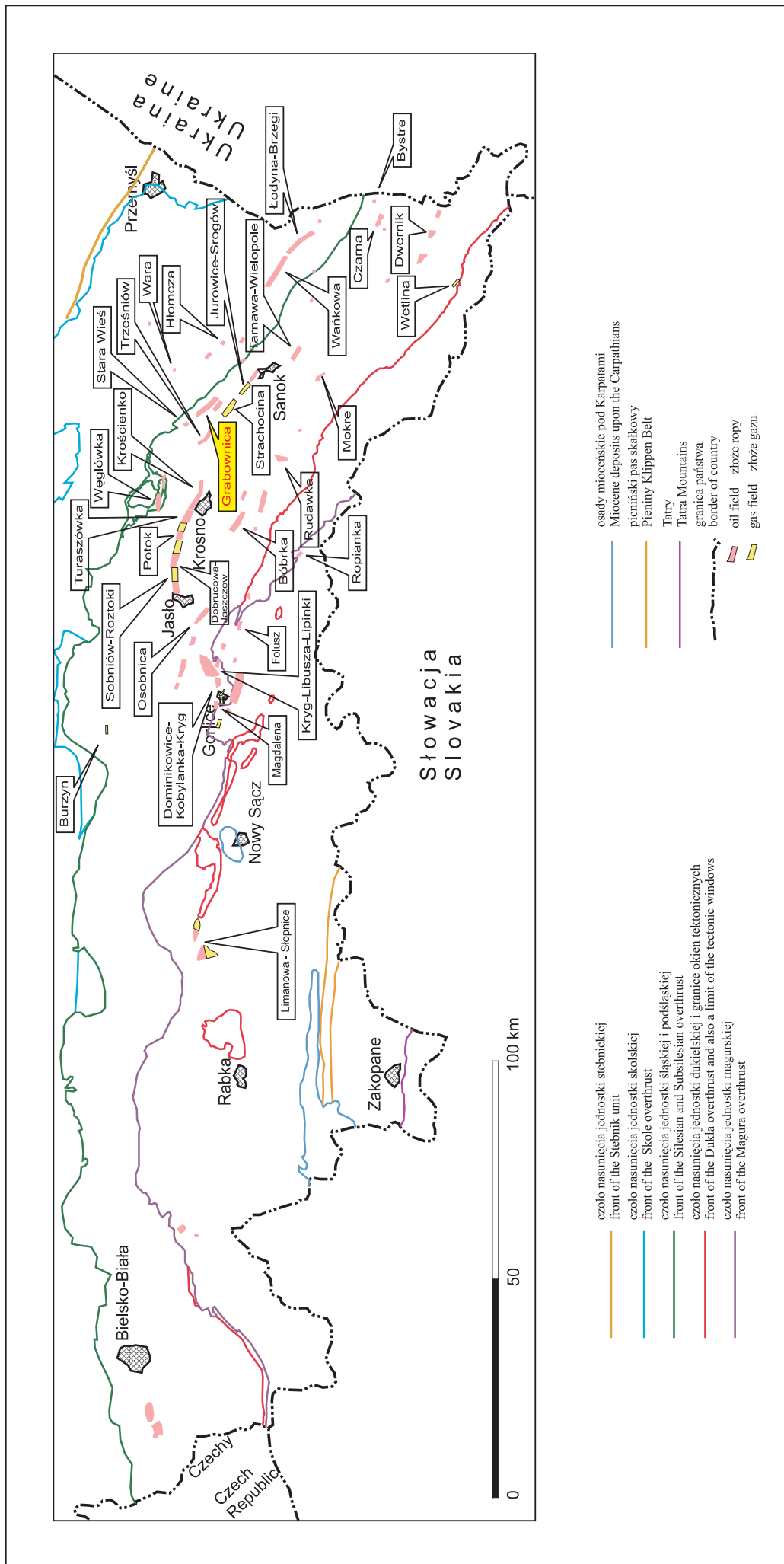
Najważniejszym celem niniejszej pracy było zweryfikowanie przydatności zaproponowanej metodyki. Trudnością, której pokonanie stanowiło zadanie o fundamentalnym znaczeniu, był fakt, że dane elektromagnetyczne oddają przede wszystkim strukturę (tektonikę ośrodka skalnego, kolektory nasycone węglowodorami, poziomy wodonośne), natomiast za pomocą SSN prowadzi się standardowo rozpoznawanie obrazów o charakterze niestrukturalnym. Podejście strukturalne jest bardziej złożone i pracochłonne – wymagałoby zbudowania odpowiednich gramatyk w celu oddania struktury obiektów. Zatem korzystne byłoby, jeśli udało się z sukcesem zastosować podejście niestrukturalne.

Oczywiście, proponowany sposób prognozowania stref złożowych jest możliwy pod warunkiem skorelowania wartości atrybutów geofizycznych z produktywnością (wielkością złoża) bądź przynajmniej z obecnością lub brakiem obecności złoża węglowodorów.

## MATERIAŁ BADAWCZY

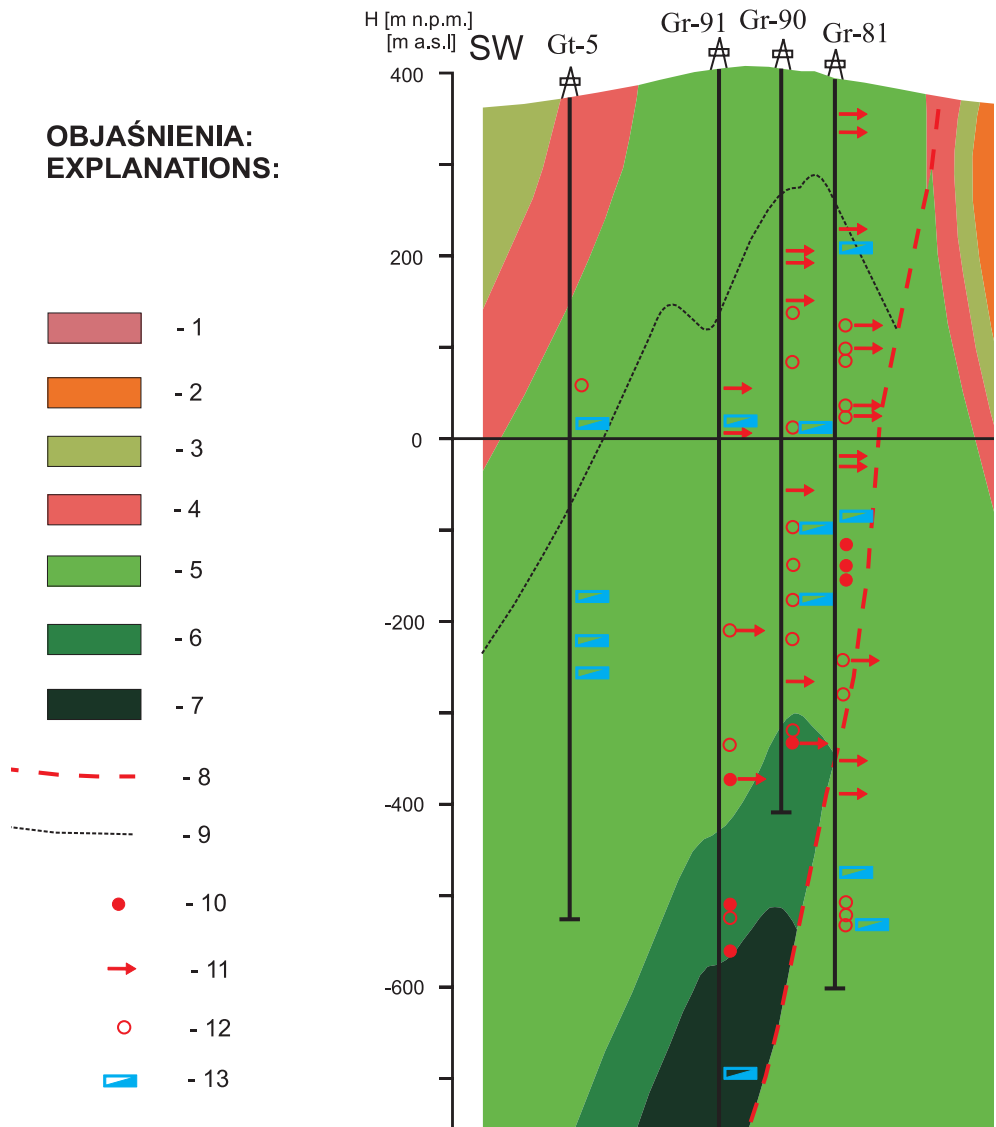
Zaproponowaną metodę przetestowano na materiale badawczym ze strefy złoża ropy naftowej i gazu ziemnego Grabownica oraz z okalających je obszarów niezłożowych. Eksploatowane od końca XIX w. złożo Grabownica (Karnkowski 1993) (Fig. 1 na wklejce) leży na SE od Brzozowa, w jądrze skomplikowanego fałdu antyklinalnego zbudowanego z kredowych warstw wierzowskich, lgockich i godulskich (Fig. 2) serii śląskiej. Jest to złożo typu warstwowego, z poziomami roponośnymi na głębokościach 300–1070 m.

Złożo Grabownica zostało przecięte kilkoma przekrojami magnetotellurycznymi AMT i CSAMT w wariantach z zastosowaniem dipola podłużnego i poprzecznego względem linii profili pomiarowych. Punkty pomiarowe AMT są oddalone od siebie średnio o 50 m, CSAMT – o 50 lub 25 m; każdy punkt to kilkadziesiąt (ok. 20) wartości każdego z parametrów w funkcji głębokości (mierzonych co ok. 30 m). Zasięg głębokościowy sondowań w metodzie CSAMT wynosi ok. 900 m p.p.t., a w metodzie AMT – ok. 1150 m p.p.t.



**Fig. 1.** Lokalizacja złoża ropy Grabownica na tle wybranych złóż węglowodorów w polskich Karpatach Zewnętrznych (wg Dziadzia *et al.* 2006, zmienione)

**Fig. 1.** Location map of the Grabownica oil deposit against a background of the selected hydrocarbon deposits in Polish Outer Carpathians (after Dziadzia *et al.* 2006, modified)



**Fig. 2.** Przekrój geologiczny przez złożo Grabownica opracowany przez Lejkowską (1975) (wg Wojdyły *et al.* 2009): 1 – warstwy menilitowe, 2 – łupki pstre i piaskowce ciężkowickie (eocen), 3 – warstwy istebniańskie, 4 – piaskowce godulskie, łupki pstre godulskie, radiolaryty, 5 – warstwy łgockie, 6 – warstwy wierzowskie, 7 – warstwy grodziskie, 8 – uskoki i wewnętrzne nasunięcia, 9 – strop poziomu DK 3-1, 10 – przyływ ropy, 11 – ślady ropy, 12 – ślady gazu, 13 – przyływ solanki

**Fig. 2.** Geological cross-section of Grabownica deposit made by Lejkowska (1975) (after Wojdyła *et al.* 2009): 1 – Menilite Beds, 2 – Variegated Shales (Eocene), 3 – Istebna Beds, 4 – Godula Sandstones, Godula Variegated Shales, Cherts, 5 – Lgota Beds, 6 – Verovice Beds, 7 – Grodziszcze Beds, 8 – faults and internal thrusts, 9 – roof of the DK 3-1 level, 10 – oil influx, 11 – oil tracs, 12 – gas tracs, 13 – brine influx

W badaniach wykorzystano dane z trzech przekrojów magnetotellurycznych: G2, G3 i G4 (łącznie 7.76 km przekrojów). Środkowe części przekrojów G2 i G4 przecinają złożę, natomiast przekrój G3 znajduje się poza złożem (poza obszarem górniczym). Dane zobrażowano w formie regularnej siatki kwadratowej o oczku  $50\text{ m} \times 50\text{ m}$ , za pomocą metody Occama 1D (Constable *et al.* 1987, de Groot-Hedlin & Constable 1990). Każdy węzeł tej siatki jest reprezentowany matematycznie przez wektor wartości parametrów AMT/CSAMT charakteryzujących skały zawierające węglowodory, w wariantach z dipolem podłużnym i poprzecznym.

## METODYKA BADAŃ

Problematyka rozpoznawania obrazów i tematyka sztucznych sieci neuronowych ma bogatą literaturę. W nowszych pracach, np. w monografii Kosińskiego (2007), można znaleźć wykaz literatury, w tym wiele monografii; wśród literatury polskojęzycznej należy wymienić monografię Tadeusiewicza & Flasińskiego (1991). Wśród pozycji literatury z zakresu stosowania metod rozpoznawania obrazów i sztucznych sieci neuronowych do prognoz złożowych można wymienić np. prace Darlak & Włodarczyk (2003) i Juchy *et al.* (2004).

Klasyfikacja nadzorowana wymaga od nadzorcy klasyfikacji (nauczyciela) uprzedniego zdefiniowania obiektów, poddawanych klasyfikacji, wyznaczenia klas oraz wskazania obiektów wzorcowych dla tych klas. Nie jest oczywiste, co uznamy za obiekty, które będą podlegały klasyfikacji, w sytuacji gdy mamy do dyspozycji dane 2D z przekroju magnetotellurycznego. Założono, że obiekty powinny mieć charakter bloków przestrzeni 2D ośrodka skalnego, z punktami pomiarowymi wewnątrz. W podejściu niestrukturalnym do rozpoznawania obrazów zakłada się, że struktura obiektu, jeśli takowa istnieje, nie jest ujawniana, a zapis obiektu to wektor wartości cech. Mimo to pośrednio dążono do uwzględnienia struktury, choćby w ograniczonym zakresie, gdyż wiadomo, że profilowania AMT/CSAMT niosą informację o strukturze ośrodka geologicznego. Być może, mimo zagubienia informacji o strukturze, poprawna klasyfikacja byłaby możliwa, ale nie uznano tego za podejście optymalne. Dążąc do uwzględnienia struktury w obiekcie, stwierdzono wstępnie, że opisanie obiektów tylko wektorami wartości czterech cech (AMT i CSAMT podłużne i poprzeczne) spowoduje kompletną utratę informacji o strukturze. Jak się wydaje, jednym ze sposobów częściowego uwzględnienia struktury w obiekcie niestrukturalnym jest wprowadzenie do wektora obiektu informacji o obiektach mu najbliższych, np. sąsiadujących z nim, w formie dodatkowych cech. Operacja taka spowoduje w praktyce, że obiekty będą zachodzić na siebie – nie będą rozłączne, co może wywołać spadek rozdzielczości prognozy złożowej.

Innym, prostszym sposobem, zastosowanym ostatecznie w niniejszej pracy, było zdefiniowanie większych obiektów rozłącznych i uwzględnienie w ich opisie parametrów AMT i CSAMT odnoszących się do elementów składowych tych obiektów. W stosunku do pierwszego sposobu następuje tu większe uśrednienie wartości parametrów (większe bloki), co również może mieć wpływ na spadek rozdzielczości prognozy złożowej. Aby temu zapobiec, zachowano małe rozmiary obiektów w poziomie (większa rozdzielczość), wydłużając je jednocześnie maksymalnie w pionie.

Konstrukcja obiektów wynikała z charakterystycznych cech przekrojów magnetotelurycznych ze złoża Grabownica. W górnej części złoża, od powierzchni terenu (około +400 m n.p.m.) do rzędnej około +200 m n.p.m. pojawia się wysokooporowa strefa, być może poziomy wodonośny lub nieporowate piaskowce. W środkowej, niższej części (od +200 do -100 m n.p.m.) widać wydłużone pionowo, naprzemienne maksima i minima, natomiast w najniższej części, poniżej rzędnej -100 m n.p.m., widać wysokooporowe struktury fałdów antyklinalnych, najprawdopodobniej z obecnością węglowodorów (Fig. 2).

Po wielu próbach obiekt zdefiniowano zatem jako prostokąt, okalający każdy (co 50 m) kompletny pionowy rząd węzłów siatki interpolacyjnej, przy czym podzielono go na trzy części, według zdefiniowanych wyżej granic, a następnie opisano oddzielnie każdą część zestawem czterech parametrów (AMT i CSAMT podłużne i poprzeczne).

Obiekty ze złoża Grabownica zostały opisane wartościami łącznie dwunastu parametrów. Są to średnie wartości z węzłów siatki interpolacyjnej, znajdujące się w odpowiedniej części obiektu (bloku). Na przekroju G2 wydzielono w ten sposób 50 prostokątów o szerokości 50 m i długości równej głębokości przekroju magnetotelurycznego. Na tej samej zasadzie zdefiniowano 55 prostokątów na przekroju G3 i 52 – na przekroju G4. Specyficzny kształt obiektów powodował, że informacja uśredniana była w liniach pionowych, w wyniku czego z dwuwymiarowego przekroju otrzymano jednowymiarową linię.

W celu maksymalnego skontrastowania wzorców klas zdefiniowano tylko dwie klasy: złożową i niezłożoną, a także wskazano obiekty, które leżą na pewno w strefach złożowych oraz takie, które leżą na pewno w strefach niezłożowych. Nadmienić tu należy, że klasa złożowa reprezentuje zarówno złożo węglowodorów (wysokie wartości AMT/CSAMT, zwłaszcza CSAMT), jak również skały o cechach charakterystycznych dla stref złożowych. Podobnie klasa niezłożowa reprezentuje brak złoża (niskie wartości AMT/CSAMT) lub skały o cechach charakterystycznych dla stref niezłożowych.

Spśród obiektów o znanej przynależności klasowej wydzielił się wzorce klas, tzw. ciąg uczący (U). Na bazie wzorców możliwa jest późniejsza klasyfikacja obiektów o nieznanym przynależności klasowej. Celem zweryfikowania klasyfikacji klasyfikuje się próbnie również obiekty niewzorowe, których przynależność klasowa jest znana, tzw. obiekty walidacyjne (V). Obiekty walidacyjne nie biorą udziału w procesie uczenia sieci, zatem niskie błędy w ich klasyfikacji oddają cenną właściwość sieci – jej zdolność do generalizacji, czyli prawidłowego klasyfikowania nieznanymi obiektów. Pozwala to oszacować jakość klasyfikacji nieznanymi obiektów. Użycie tzw. „Automatycznego projektanta sieci” zakłada jednak testowanie wielu sieci, a przy wyborze kolejnych sieci, wykorzystywana jest informacja o błędach klasyfikacji obiektów walidacyjnych. Konieczne było więc wykorzystanie dodatkowego zbioru obiektów testowanych (T), o znanej przynależności klasowej, do których „Automatyczny projektant sieci” nie ma dostępu.

Obiekty U, V i T wyznaczono oddzielnie dla klasy złożowej i niezłożowej. Wybierano tylko te obiekty, które nie wykazywały braków danych. Środkowe części przekrojów G2 i G4 przecinają złożo, z nich zatem pobrano obiekty należące do klasy złożowej. Należy zaznaczyć, że w środkowych częściach tych przekrojów znajdują się złożowe (lub były złożowe) otwory wiertnicze: Wanda-106, Wanda-103 i Gr-104 w linii G2 oraz: Genpeg-65, Genpeg-66, a przede wszystkim HB-7. Obiekt klasy niezłożowej pobrano z krańców

przekrojów G2, G4 oraz z krańców przekroju G3 znajdujących się poza złożem. Obiekty części środkowej G3 potraktowano jako obiekty o nieznanym przynależności klasowej i nie włączano ich do zbioru  $UVT = U \cap V \cap T$ . Choć znajdują się tam dwa otwory negatywne: Gr-108 i Gr-109, nie było stuprocentowej pewności, że nie ma tu węglowodorów. Ta teza została później zweryfikowana w prognozie złożowej. Ostatecznie zdefiniowano 30 obiektów uczących, 26 walidacyjnych oraz 23 testowe o znanej przynależności klasowej.

Problem właściwego doboru obiektów obu zdefiniowanych klas wymaga jeszcze dalszych badań. Dobór ten można by przeprowadzić na podstawie szczegółowych informacji o tektonice fałdów niższego rzędu z jądrowej strefy antyklinalnej, należałoby również zbadać wpływ, jaki może mieć zaszumienie wywołane nieprecyzyjnym doбором obiektów, zarówno z jądra strefy antyklinalnej, jak i spoza niego.

Obiekty spoza zbioru UVT to obiekty, których przynależność klasowa była niepewna lub/i były opisane niepełnym zestawem parametrów geofizycznych. Nie uczestniczyły one w etapie uczenia sieci i nie były podstawą do oceny jakości klasyfikacji, natomiast były one poddane klasyfikacji.

Obiekty ze złoża Grabownica zostały opisane pełnym zestawem dwunastu parametrów:

AMT_r_d	CSAMT_r_d
AMT_r_s	CSAMT_r_s
AMT_r_g	CSAMT_r_g
AMT_pp_d	CSAMT_pp_d
AMT_pp_s	CSAMT_pp_s
AMT_pp_g	CSAMT_pp_g

gdzie r oznacza sondowania dipolem równoległym, pp – sondowania dipolem poprzecznym, d – dolną część obiektu, s – środkową część obiektu, g – górną część obiektu.

Wyniki klasyfikacji wyrazić można w wartościach prawdopodobieństwa przynależności obiektu do klasy złożowej (klasyfikacja probabilistyczna), które mogą być interpolowane na mapach prognostycznych, albo poprzez wybór klasy dla obiektu (klasyfikacja dyskretna).

W razie pozytywnej weryfikacji proponowanej metody prognozy złożowej, możliwa jest ocena wykorzystywanych atrybutów geofizycznych pod względem ich przydatności do tejże prognozy. Zadanie to starano się rozwiązać przez badanie informatywności atrybutów, tzn. określenie przydatności każdego atrybutu do rozklasyfikowania obiektów, co wiąże się z ilością informacji niesionej przez dany atrybut. Wpływ poszczególnych parametrów na jakość klasyfikacji nie jest jednakowy i zależy od wag przypisanych odpowiednim wejściom neuronów. Informacja o wagach jest w zasadzie niejawna, a jeśli już to trudno interpretowalna. Ocena informatywności przeprowadzono więc poprzez tzw. analizę wrażliwości sieci na utratę określonego atrybutu – duża wrażliwość wskazywała na przydatność atrybutu. Należy jednak zwrócić uwagę na niebezpieczeństwo związane z tym, że wyniki badania informatywności pojedynczych cech nie mogą być jednoznaczne w sposób obiektywny: klasyfikacja ma charakter wielowymiarowy i informacja jest niesiona nie tylko w pojedynczych atrybutach, ale również w podzbiorach atrybutów. Zatem niska informatywność atrybutu nie przesądza ostatecznie jego braku przydatności, gdyż atrybut taki

może należeć do podzbioru informatywnego, a jego odrzucenie może się wiązać z istotnym spadkiem informatywności podzbioru. Na przykład, analiza wrażliwości wyławia ważne atrybuty, ale w odniesieniu do sytuacji zastanej, gdy w sieci „pracują” pozostałe atrybuty. Standardowo wyniki analizy wrażliwości dla konkretnego parametru są przedstawiane oddzielnie w zbiorze U i V jako błąd sieci po stracie parametru, tzn. stosunek (iloraz) tego błędu do błędu sieci przed próbnym usunięciem parametru oraz miejsce parametru w ranking. Im wyższy iloraz błędów, tym wyższa informatywność, z zastrzeżeniem, że powinien on przekraczać 1. Jeśli tak się nie dzieje, to parametr wywiera szkodliwy wpływ na jakość klasyfikacji, gdyż błąd zmniejsza się po jego usunięciu. Parametr taki powinien być usunięty.

## WYNIKI

Za pomocą „Automatycznego projektanta sieci” przetestowano wiele sieci o różnych architekturach, rozmiarach (niektóre zmienne były wyłączane podczas badania), i wartościach wag przypisanych wejściom neuronów. Najlepsze sieci zestawiono w tabeli 1. Błędy uczenia najlepszych sieci osiągają w zbiorach U i V bardzo niski poziom, rzędu kilku setnych. Przy wyborze sieci, użytej później do klasyfikacji, kierowano się nie tylko niskimi wartościami błędów, ale również wielkością sieci (ilością wejść i neuronów ukrytych) – preferowano mniejsze sieci, o ile potrafiły one osiągnąć zadowalająco niski poziom błędów. Jak jednak wspomniano w opisie metodyki, w przypadku użycia „Automatycznego projektanta sieci” błędy w zbiorach U i V nie opisują rzeczywistej zdolności sieci do prawidłowego klasyfikowania nieznanymi obiektów. Rolę taką spełniają błędy w zbiorze T, które są zdecydowanie wyższe, rzędu 0.30, choć jeszcze akceptowalne. Niewątpliwie jednym z możliwych powodów są zbyt małe liczebności zbiorów U i V. Niestety, zwiększenie liczebności tych zbiorów powoduje również zwiększenie niepewności co do przynależności klasowej dobieganych obiektów. Problem ten może być w przyszłości złagodzony przez zwiększenie gęstości siatki interpolacyjnej, na bazie której definiowano obiekty.

Błędy we wszystkich trzech zbiorach, U, V i T, uwzględniają sumaryczne odchyłki od wskazań nauczyciela w klasyfikacji probabilistycznej. Efekty klasyfikacji dyskretnej z użyciem optymalnego progu do rozdzielenia obiektów obu klas (Tab. 2) pokazują, że oszacowanie błędów w zbiorach T jest mniej wiarygodne i najprawdopodobniej zawyżone, z powodu zbyt małej liczby tych obiektów. Dzieje się tak szczególnie w klasie złożowej (dwa obiekty, jeden rozpoznany poprawnie, drugi – nie). Niewiarygodne oszacowanie błędów w zbiorach T nie wpływa na jakość klasyfikacji, choć nie pozwala na jej w pełni wiarygodną ocenę. Nie zmienia to jednak faktu, że liczebności w zbiorach U i V są zbyt niskie.

Wyniki klasyfikacji dyskretnej, traktowane jako prognoza złożowa, przedstawiono na mapie złoża Grabownica (Fig. 3) i na badanych przekrojach magnetotellurycznych przez to złożo: G2 (Fig.4), G4 (Fig. 5) i G3 (Fig. 6).

Na przekroju G2 (Fig. 3, 4) złożo zostało wskazane dokładnie wewnątrz konturu złoża, skąd zostały pobrane obiekty UVT, wskazane przez nauczyciela (nadzorcę klasyfikacji) jako złożowe. Jeden z nich, leżący nad strefą niskich oporności, został rozpoznany jako niezłożowy.



**Tabela (Table) 1**

Wykaz dziesięciu najlepszych spośród tysięcy sieci neuronowych testowanych w jednej sesji obliczeniowej dla złoża Grabownica

*Ten best networks from among thousands neural networks that were tested in one computation session for the Grabownica deposit*

Lp. No.	Typ Type	Wejścia Inputs	Ukryte Hidden	Błąd uczenia Learning error	Błąd wal. Validation error	Błąd tst Testing error	Uczenie Learning
1	RBF	3	2	0.39	0.38	0.32	KM, KN, PI
2	Liniowa Linear	3	–	0.40	0.38	0.30	PI
3	Liniowa Linear	4	–	0.40	0.38	0.30	PI
4	Liniowa Linear	5	–	0.39	0.38	0.30	PI
5	RBF	4	2	0.38	0.38	0.31	KM, KN, PI
6	RBF	5	2	0.37	0.38	0.28	KM, KN, PI
7	MLP	5	6	0.21	0.26	0.26	BP50, CG50, CG88b
8	MLP	6	3	0.18	0.17	0.30	BP50, CG97b
9	MLP	6	11	0.08	0.13	0.27	BP50, CG50, CG143b
10*	MLP	8	18	0.05	0.06	0.30	BP50, CG50, CG139b

Objaśnienia:

\* Sieć użyta do klasyfikacji.

Typ (sieci): RBF – sieć oparta na radialnych funkcjach bazowych (neurony o radialnej funkcji agregującej i liniowej funkcji aktywacji); MLP – perceptron wielowarstwowy; liniowa – jednowarstwowa sieć z neuronami liniowymi.

Uczenie (= trzy ostatnie algorytmy uczenia): KM – K-średnich; KN – K-najbliższych sąsiadów; PI – pseudoinwersja (liniowa optymalizacja najmniejszych kwadratów); BP – wsteczna propagacja błędów; CG – gradienty sprzężone; liczby przy oznaczeniach BP i CG oznaczają liczby epok uczenia; b – kod sposobu zakończenia: zachowano sieć o najmniejszym błędzie walidacyjnym.

Liczbę warstw ukrytych w sieciach ograniczono do jednej.

*Explanations:*

\* *The network used in the classification.*

*Type of network: RBF – radial basic function (neurons with radial aggregating function and linear activation function); MLP – multi layer perceptron; Linear – 1-layer network with linear neurons.*

*Learning (= the three last algorithms): KM – K-Means; KN – K-Nearest Neighbours; PI – Pseudoinversion (the least square linear optimization); BP – (Back Propagation of Errors); CG – Conjugate Gradient Descent; numbers at the BP and CG – numbers of learning epochs; b – code of way of termination: the network with the least validation error was saved.*

*Number of hidden layers was limited to one.*

**Tabela (Table) 2**

Liczby poprawnie i niepoprawnie sklasyfikowanych obiektów (klasyfikacja dyskretna) dla złoza Grabownica (sieć 10)

*Numbers of correctly and incorrectly classified objects (discrete classification) for the Grabownica deposit (network 10)*

	Klasy w zbiorze U <i>Classes in set U</i>		Klasy w zbiorze V <i>Classes in set V</i>		Klasy w zbiorze T <i>Classes in set T</i>	
	Złoże <i>Deposit</i>	Nie-złoże <i>Non-deposit</i>	Złoże <i>Deposit</i>	Nie-złoże <i>Non-deposit</i>	Złoże <i>Deposit</i>	Nie-złoże <i>Non-deposit</i>
Razem <i>Total</i>	8	22	6	20	2	21
Poprawnie <i>Correctly</i>	8	22	6	20	1	20
Błędnie <i>Incorrectly</i>	0	0	0	0	1	1
Nieokreślone <i>Ambiguous</i>	0	0	0	0	0	0
Złoże <i>Deposit</i>	8	0	6	0	1	1
Nie-złoże <i>Non-deposit</i>	0	22	0	20	1	20

**Tabela (Table) 3**

Analiza wrażliwości parametrów wejściowych dla złoza Grabownica (sieć 10)

*Sensitivity analysis of input parameters for the Grabownica deposit (network 10)*

		AMT_r_s	AMT_r_g	AMT_p_d	AMT_p_s	AMT_p_g	CSAM T_r_s	CSAM T_pp_s	CSAM T_pp_g
Zbiór U <i>Set U</i>	Ranga <i>Rank</i>	8	2	7	3	4	6	5	1
	Błąd <i>Error</i>	0.10	0.47	0.19	0.46	0.45	0.23	0.26	0.48
	Iloraz <i>Quotient</i>	1.80	8.50	3.40	8.30	8.15	4.24	4.71	8.60
Zbiór V <i>Set V</i>	Ranga <i>Rank</i>	8	4	5	1	3	7	6	2
	Błąd <i>Error</i>	0.08	0.39	0.27	0.47	0.44	0.17	0.18	0.46
	Iloraz <i>Quotient</i>	1.23	5.63	4.03	6.86	6.43	2.58	2.63	6.69

Ten przypadek sugeruje, że aby uczynić prognozę bardziej precyzyjną, należałoby pobrać nie wszystkie obiekty z wnętrza konturu złoża, ale tylko te, które leżą w pobliżu otworów złożowych. Taki zabieg zmniejszy jednak liczebność obiektów U, V i T, co jest równie niekorzystne. W odległości 750 m na SW za konturem złoża (poza rysunkiem przekroju) pojawiły się obiekty mniej wiarygodne (braki danych), sklasyfikowane jako złożowe.

Podobny obraz jest na przekroju G4 (Fig. 3, 5) – wszystkie obiekty z wnętrza konturu złoża zostały rozpoznane jako złożowe, a dodatkowo, 500 m na NE za konturem złoża, ujawniła się strefa złożowa (jeden obiekt bardziej wiarygodny obok dwóch mniej wiarygodnych z brakami danych). Strefa ta jest na granicy rysunku przekroju, ale widać, że zaczynają się tam maksima oporności, co potwierdzałoby prognozę.

Przekrój G3 (Fig. 3, 6) okazał się całkowicie płonny, choć można było podejrzewać, że w osi struktury antyklinalnej – w przedłużeniu złoża na NE, mogła się pojawić strefa złożowa. Aby wykryć tę ewentualną strefę w sposób jak najbardziej wiarygodny, wszystkie obiekty z jej wnętrza traktowano jako obiekty o nieznanym przynależności klasowej. Uzyskany wynik wskazuje, że najprawdopodobniej jest to rzeczywiście strefa płonna.

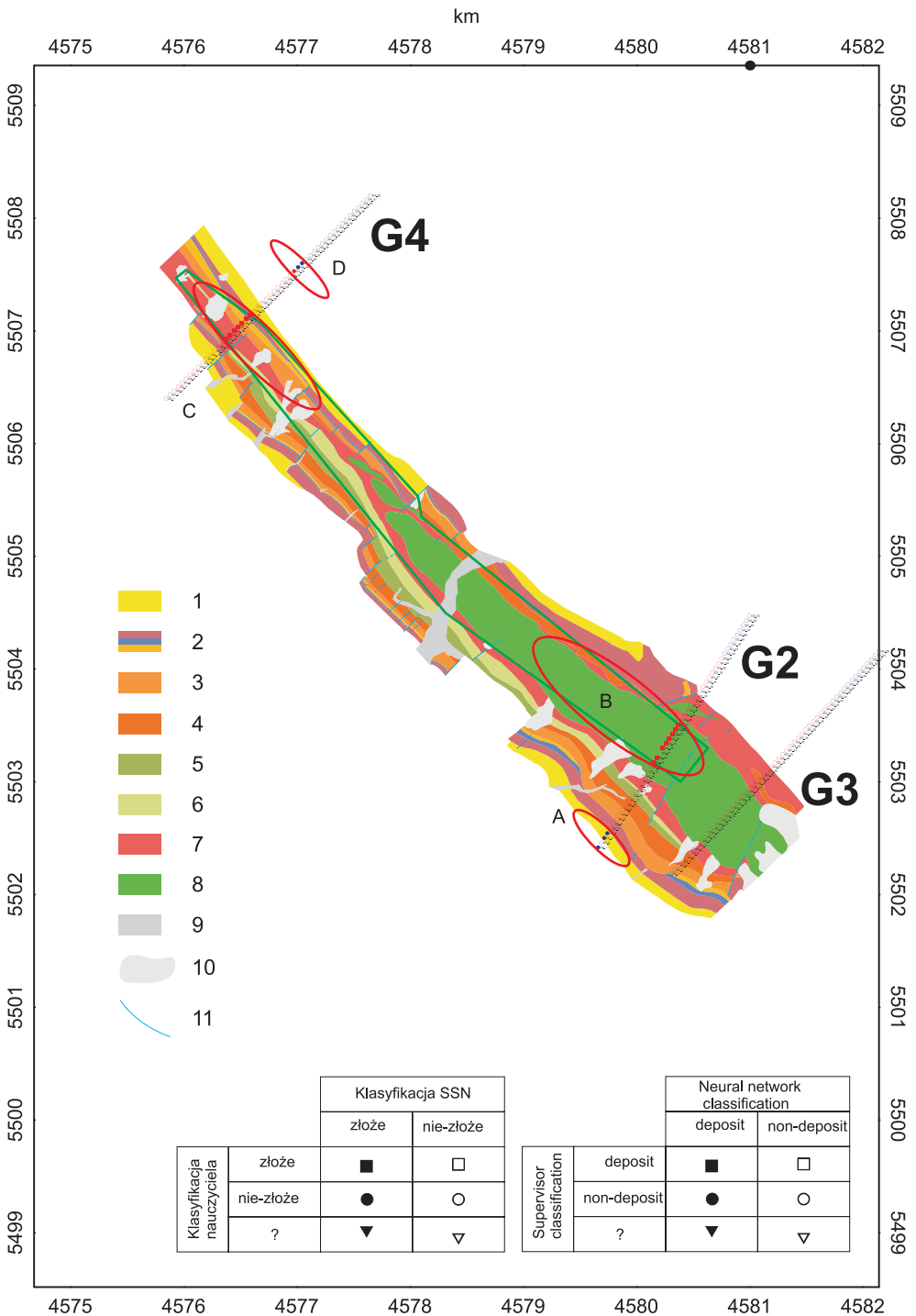
Stosunkowo dobra jakość klasyfikacji upoważnia do oceny, które parametry miały na to szczególnie duży wpływ. Jak już jednak wspomniano w rozdziale metodycznym, trzeba tu zastrzec, że, taka ocena zwykle nie obejmuje pełnej wiedzy o przydatności parametrów w prognozie złożowej. Analiza wrażliwości (Tab. 3) pokazuje, że wszystkie osiem uwzględnionych parametrów ma korzystny wpływ na jakość klasyfikacji (iloraz błędów powyżej 1). Cztery parametry zostały wyłączone przez „Automatycznego projektanta sieci”, co oznacza, że są one nieprzydatne w sieci nr 10. Nie znaczy to jednak, że są one w ogóle nieprzydatne, gdyż dobór parametrów nie odbywał się metodą optymalizacji o charakterze globalnym.

Największe znaczenie (w połączonym rankingu dla zbiorów U i V) mają kolejno: CSAMT\_pp\_g, AMT\_pp\_s, AMT\_r\_g i AMT\_pp\_g.

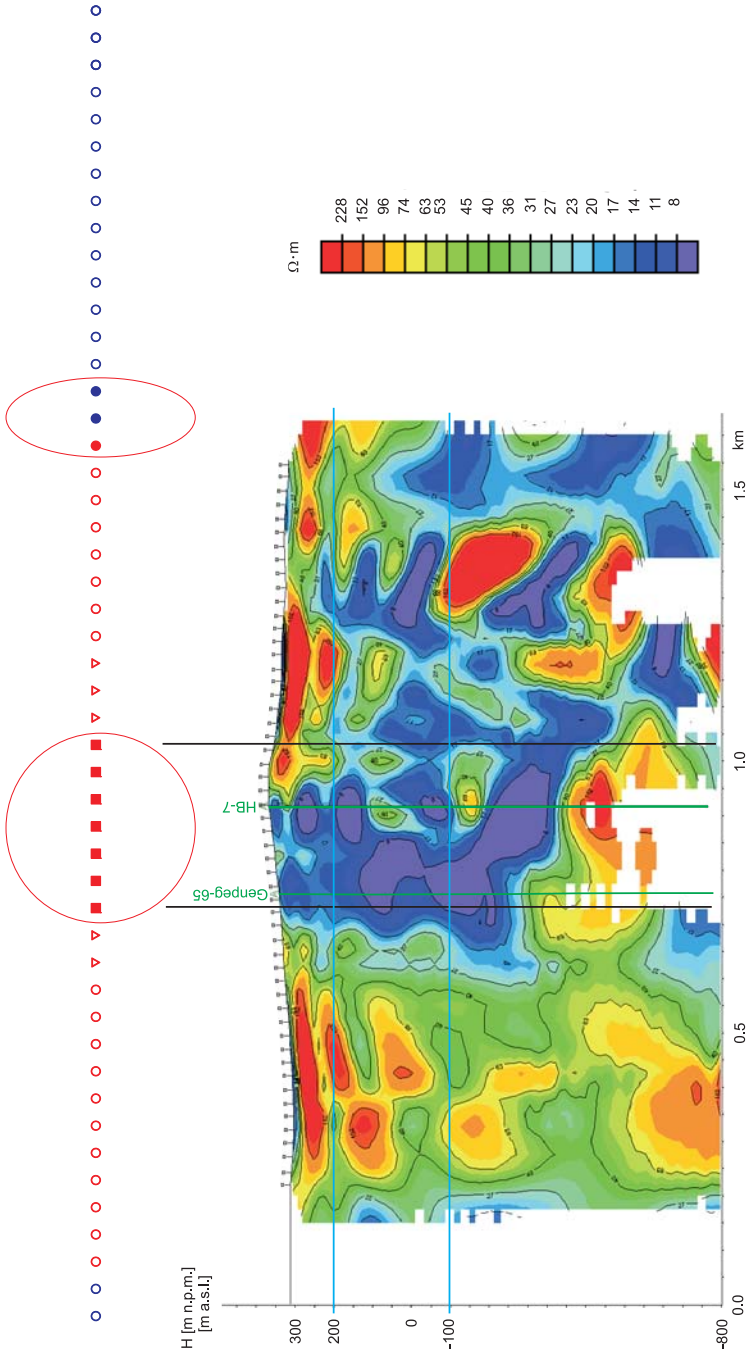
---

**Fig. 3.** Prognoza wykonana za pomocą sieci neuronowych dla złoża Grabownica na tle modelu geologicznego złoża opracowanego przez S. Szewczyka (wg Wojdyła *et al.* 2009): 1 – warstwy krośnieńskie, 2 – łupki menilitowe z rogowcami i piaskowcami podrogowcowymi (eocen), 3 – warstwy hieroglifowe (eocen), 4 – łupki pstre (eocen), 5 – łupki czarnorzecze (kreda górna), 6 – piaskowce czarnorzecze (kreda górna), 7 – warstwy godulskie (kreda górna), 8 – warstwy lgockie (kreda dolna), 9 – aluwia, 10 – tereny osuwiskowe, 11 – dyslokacje. Odpowiednimi symbolami oznaczono klasy obiektów, wskazane przez nadzorcę klasyfikacji (nauczyciela) oraz wynik klasyfikacji za pomocą SSN. Czerwony kolor symboli oznacza obiekty opisane pełnym zestawem parametrów, niebieski – obiekty z niekompletnymi danymi. Czerwone elipsy wskazują na potencjalne złożo

**Fig. 3.** Prognosis made by means of neural networks for Grabownica deposit against a background of a geological model made by S. Szewczyk (after Wojdyła *et al.* 2009): 1 – Krosno beds, 2 – Menilite Shales with cherts and Subchert Sandstones (Eocene), 3 – Hieroglyphic Beds (Eocene), 4 – Variegated Shales (Eocene), 5 – Czarnorzecze Shales (Upper Cretaceous), 6 – Czarnorzecze Sandstones (Upper Cretaceous), 7 – Godula Beds (Upper Cretaceous), 8 – Lgota Beds (Lower Cretaceous), 9 – alluvia, 10 – landslide areas, 11 – dislocations. The classes pointed by supervisor and neural network classification are marked with symbols. Red marks – objects possessing full set of the parameters, blue marks – objects with incomplete set of the parameters. Red ellipses point to potential deposit

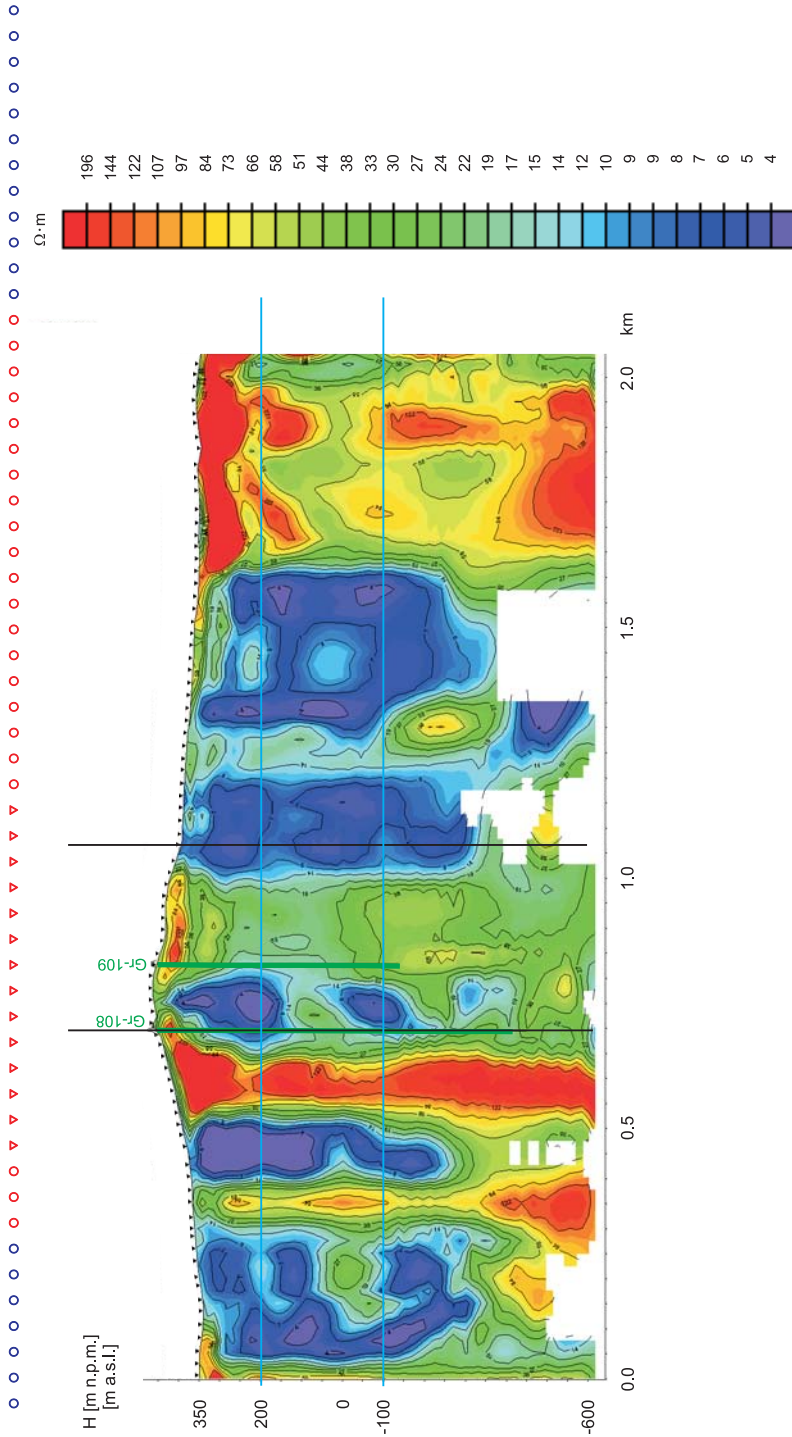






**Fig. 5.** Prognoza złożowa za pomocą sieci neuronowych na tle przekroju magnetotellurycznego G4 parametru CSAMT na złożu Grabownica (dane magnetotelluryczne wg Wojdyła *et al.* 2009): małe kwadraty z kropkami w środkach – punkty pomiarowe; pozostałe objaśnienia jak na figurze 4

**Fig. 5.** Prognosis made by means of neural networks against a background of magnetotelluric cross-section G4 for the CSAMT parameter for Grabownica deposit (magnetotelluric data after Wojdyła *et al.* 2009): small squares with dots inside – measurement points; other explanations as in figure 4



**Fig. 6.** Prognoza złożowa za pomocą sieci neuronowych na tle przekroju magnetotellurycznego G3 parametru CSAMT na złożu Grabownica (dane magnetotelluryczne wg Wojdyła *et al.* 2009); objaśnienia jak na figurze 4

**Fig. 6.** Prognosis made by means of neural networks against a background of magnetotelluric cross-section G3 for the CSAMT parameter for Grabownica deposit (magnetotelluric data after Wojdyła *et al.* 2009); explanations as in figure 4

## WNIOSKI

Dowodzono przydatności SSN do prognozy złożowej. Sieci uczą się bardzo łatwo, osiągając bardzo niskie błędy w zbiorze uczącym i walidacyjnym. Rzeczywiste błędy, które może popełniać sieć, ujawnione w zbiorze testowym są rzędu 0.30.

Błędy te można by było najprawdopodobniej obniżyć poprzez zwiększenie liczebności ciągu uczącego, natomiast zwiększając liczebność zbiorów V i T, można by poprawić jakość oszacowania błędów. Przy ograniczonej liczbie obiektów U, V i T, których przynależność do klasy złożowej czy niezłożowej nie budzi wątpliwości, zwiększanie liczebności tych zbiorów może wręcz prowadzić do gorszej prognozy. Dlatego rozsądnym rozwiązaniem mogłoby być wykorzystywanie gęstszych siatek interpolacyjnych, na bazie których można tworzyć więcej obiektów. Przy dużej ich liczbie można by nawet zastosować ostrzejsze kryterium ich doboru – zamiast pobierania obiektów złożowych z wnętrza konturu złoża, a niezłożowych z zewnątrz, można by uwzględniać tylko te obiekty, które leżą w pobliżu otworów wiertniczych, wskazujących jednoznacznie na obecność bądź brak złoża (choć otworów tych jest mało).

Mimo tych zastrzeżeń należy stwierdzić, że błędy sieci są na poziomie możliwym do zaakceptowania. Wynika to zapewne z dobrej korelacji między parametrami wejściowymi a obecnością bądź brakiem złoża. Dobra jakość prognozy złożowej jest potwierdzona przez prawidłowe rozpoznanie stref złożowych i niezłożowych na analizowanych przekrojach magnetotellurycznych (G2, G3, G4). Prognoza wskazuje dodatkowo małe strefy złożowe w NE części przekroju G4 i w SW części przekroju G2. Wyniki dotychczasowych badań nad przydatnością zaproponowanej metodyki do prognozy złożowej wskazują jednoznacznie, że to podejście jest wysoce perspektywiczne.

*Autor dziękuje dr. hab. inż. M. Stefaniukowi oraz pracownikom Oddziału Krakowskiego Przedsiębiorstwa Badań Geofizycznych: mgr inż. Elżbiecie Maj, mgr inż. Joannie Figule i dr. inż. Markowi Wojdyle za udostępnienie materiałów i cenne uwagi.*

*Prezentowana praca została wykonana w ramach realizacji projektu badawczo-rozwojowego nr 13102 pt. „Opracowanie nowatorskich metod wykrywania złóż węglowodorów oraz rozpoznawania struktury i zmienności złóż w trakcie ich eksploatacji za pomocą głębokich i średniozasięgowych badań elektromagnetycznych”, finansowanego przez Ministerstwo Nauki i Szkolnictwa Wyższego w ramach Inicjatywy Technologicznej I (umowa 162/R/T02/2007/IT1).*

## LITERATURA

- Constable S.C., Parker R.L. & Constable C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, 52, 289–300.
- Darłak B. & Włodarczyk M., 2003. *Niektóre problemy oceny właściwości skał zbiornikowych z wykorzystaniem sieci neuronowych*. IGNiG, Kraków.



- Dziazio P.S., Borys Z., Kuk S., Masłowski E., Probulski J., Pietrusiak M., Górka A., Moryc J., Baszkiewicz A., Karnkowski P., Karnkowski P.H. & Pietrusiak M., 2006. Hydrocarbon resources of the Polish Outer Carpathians – Reservoir parameters, trap types, and selected hydrocarbon fields: A stratigraphic review. W: Golonka J. & Picha F.J. (Eds), *The Carpathians and their foreland: Geology and hydrocarbon resources, AAPG Memoir*, 84, 259–291.
- Fukunaga K., 1972. *Introduction to Statistical Pattern Recognition*. Academic Press, New York & London.
- Groot-Hedlin C. de & Constable S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics*, 55, 1613–1624.
- Jucha S., Kotlarczyk J. & Mastej W., 2004. Cyfrowe mapy prognozy potencjalnych skał zbiornikowych węglowodorów. *Geoinformatica Polonica*, 6, 2004, 59–65.
- Karnkowski P., 1993. *Złoża gazu ziemnego i ropy naftowej w Polsce. T. 2: Karpaty i zapadlisko przedkarpackie*. Towarzystwo Geosynoptyków „Geos”, Kraków.
- Kosiński R.A., 2007. *Sztuczne sieci neuronowe. Dynamika nieliniowa i chaos*. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Lejkowska S., 1975. *Przekrój geologiczny przez złożę Grabownica* (mat. niepubl.).
- Stefaniuk M., Maj E., Sito Ł., Słyś M. & Wojdyła M., 2010a. Anomalie elektromagnetyczne związane z akumulacjami węglowodorów na przykładzie wybranych złóż karpackich i zapadliska przedkarpackiego. *III Konferencja Naukowo-Techniczna „Ropa i gaz – złoża konwencjonalne i niekonwencjonalne”, Czarna 11–14.04.2010*, 53–54.
- Stefaniuk M., Wojdyła M., Maj E., Sito Ł. & Słyś M., 2010b. Interpretacja danych elektromagnetycznych w środkowej i północno-zachodniej części złoża ropy naftowej i gazu ziemnego „Grabownica”. *Materiały seminaryjne pod redakcją dra hab. inż. Michała Stefaniuka, Iwkowa 2010*, 11–12.
- Stefaniuk M., 2011. Metody elektromagnetyczne w prospekcji naftowej. *Geologia* (kwartalnik AGH), 37, 1, 5–36.
- Szewczyk S., 1988. Model złoża Grabownica. W: Łuczejko J., Nycz R., Szewczyk S. & Szewczyk E., *Dodatek nr 2 do dokumentacji geologicznej złoża ropy naftowej Grabownica* (mat. niepubl.).
- Tadeusiewicz R. & Flasiński M., 1991. *Rozpoznawanie obrazów*. Państwowe Wydawnictwo Naukowe, Warszawa.
- Wojdyła M., Maj E., Słyś M. & Stefaniuk M., 2009. Zastosowanie metody magnetotellurycznej do rozpoznawania płytkich złóż węglowodorów w Karpatach na przykładzie złoża „Grabownica”. *Konferencja Naukowo-Techniczna „Ropa i gaz a skały klastyczne Polski”, Czarna 17–20.03.2009*, 59–60.

## Summary

Electromagnetic data are multidimensional in character because of several parameters that can be measured during magnetotelluric survey. This makes it difficult to obtain synthetic information on economic reserves of oil and gas or their absence, especially when the

data are ambiguous. As it is proposed in the paper, such a synthesis can be done if the pattern recognition method employing artificial neural networks to supervised classification is applied. The Grabownica oil deposit case study (Figs 1, 2) proves that the proposed method is useful.

The AMT and CSAMT survey in two variants (longitudinal and transverse dipole) were used in neurocomputing. Vertical belts of 50 m width, cut off from interpolating grids of magnetotelluric cross-sections, were objects of classification. Three cross-sections were used: G2 (Figs 3, 4), G4 (Figs 3, 5) and G3 (Figs 3, 6). The structure of the formally non-structural objects was partly imposed through its division into three parts (Figs 4–6): 1) near-surface high-resistivity aquifers, to 200 m a.s.l., 2) low-resistivity screens and weak maxima – diffusion chimneys, from 200 to –100 m a.s.l., and 3) the lowest high-resistivity deposit zone. Therefore, data dimensionality increased three times, from 4 to 12. Unfortunately, not all objects were described by full set of twelve parameters. The objects coming from a known deposit or non-deposit zones (respectively deposit or non-deposit classes) were divided into three sets: U (learning set), V (validation set) and T (tested set). The validation set, which is usually used for validation cannot be taken into consideration for fully credible assessing of network errors in the case when “Automatic designer of networks” was used. This role was played by the set T. The objects  $U \cup V \cup T$  from the deposit class were taken from inner parts of magnetotelluric cross-sections G2 and G4, which cut the deposit range, whereas such objects representing the non-deposit class come from the sides of G2 and G4.

Moreover, the fourth set was established for objects from areas where no information about the two classes was available. These objects were taken from those parts of G2 and G4 sides where one cannot have certainty whether they belong to the distinguished classes, as well as from the whole cross-section G3.

“Automatic designer of networks” turned off four input parameters from among twelve ones, because in the network 10 they were not helpful. However, it does not mean that they were not helpful at all. The ranking of the eight remaining parameters was done by sensitivity analysis, i.e. through increasing network error, when one selected parameter was temporarily removed (Tab. 3). The analysis showed that all eight parameters had good influence on network 10 because the increase was within a range of 1.2 to 8.6.

The first aim of investigations was the verification of the method in known deposit or non-deposit zones. In each magnetotelluric cross-section, such zones were properly found by the method (Figs 3–6). Moreover, two unknown deposit zones were identified in NE part of the cross-section G4 (Figs 3, 5) and in the SW part of cross-section G2 (Figs 3, 4). Such the results were due to very low errors of the best network model (number 10) that was selected for final classification (Tab. 1), in the  $U \cup V$  sets, and higher error, but still acceptable – in the T set. The discrete classification errors were low too (Tab. 2). All described errors could have been lower if one could increase quantities of the sets U, V and T, through increase of density of interpolation grids, which were the base to produce objects.