

*Sylwia Gworek\**, *Arkadiusz Utrata\*\**

## WYKORZYSTANIE PREDYKTORÓW TYPU NEURAL NETWORK DO PROGNOZOWANIA SZEREGÓW CZASOWYCH\*\*\*

---

### 1. Wstęp

Obok klasycznych metod prognozowania zjawisk gospodarczych istnieje grupa metod nieklasycznych. Do grupy tych metod dołączają kolejne, wykorzystujące nowe techniki modelowania procesów roboczych i oddziaływania.

Jedną z metod zaliczanych do nieklasycznych jest prognozowanie realizowane za pomocą sztucznych sieci neuronowych. Analiza literatury przedmiotu w aspekcie dziedzin zastosowania sztucznych sieci neuronowych w prognozowaniu ekonomicznym pozwala stwierdzić, że przedmiotem tych zastosowań są te dziedziny, w których istnieje możliwość uzyskania bardzo licznego szeregu czasowego charakteryzującego wybrane cechy realizowanych procesów.

Powstaje pytanie: co z analizą szeregu czasowego opisującego dziedzinę (obszary) procesu o nielicznych wartościach realizowanej cechy? Chodzi tutaj o sytuację, gdy dany proces może zostać opisany tylko ograniczonym (nielicznym) szeregiem czasowym. Nie jest to zjawisko marginalne, ponieważ analiza stabilności szeregu czasowego narzuca często konieczność odrzucenia tych danych, które istotnie różnią się od pozostałych elementów tego szeregu. Czy sztuczne sieci neuronowe znajdują w takiej sytuacji swoje zastosowanie? Czy słuszna jest teza postawiona przez J. Mikusia, że „Błędem byłoby stosowanie ich do rozwiązywania problemu, dla którego istnieje rozwiązanie ściśle analityczne” [5]? Powyższe dylematy zostaną przedyskutowane w ramach analizy szeregu czasowego opisującego wielkość wydobycia węgla kamiennego w Polsce.

---

\* Absolwentka Wydziału Górnictwa i Geoinżynierii, Akademia Górniczo-Hutnicza, Kraków

\*\* Wydział Górnictwa i Geoinżynierii, Akademia Górniczo-Hutnicza, Kraków

\*\*\* Artykuł opracowano w ramach pracy statutowej 10.10.100.35, Wydział Górnictwa i Geoinżynierii AGH

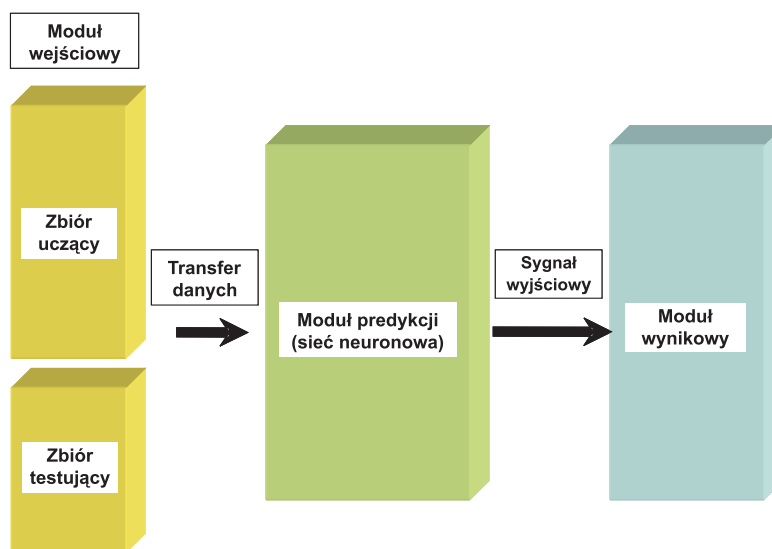
## 2. Modele sztucznych sieci neuronowych a prognozowanie szeregów czasowych

Neuron, czyli komórka nerwowa, jest podstawowym elementem biologicznej sieci neuronowej, natomiast sztuczny neuron jest elementem, którego własności odpowiadają wybranym własnościom neuronu biologicznego. Z kolei sztuczna sieć neuronowa to model prognostyczny, opisany odpowiednim algorytmem w postaci programu komputerowego. Program ten naśladuje zdolność ludzkiego umysłu do klasyfikacji wzorów i sporządzania prognoz lub podejmowania decyzji na podstawie analizy retrospektywnej. Istotną cechą sieci neuronowych jest fakt, że w wyniku procesu uczenia sieć może nabyć zdolność predykcji (przewidywania). Sieć dzięki swoim zdolnościom do uczenia, adaptacji i uogólnienia doświadczeń, pozwala zautomatyzować procesy wnioskowania retrospektywnego, tzn. wykorzystującego wcześniej zgromadzone dane.

Przy wykorzystywaniu sieci neuronowych należy podjąć decyzje dotyczące:

- przygotowania danych do przetwarzania przez sieć;
- wyboru modelu sieci, czyli tzw. architektury sieci (typ sieci, struktura połączeń, liczba warstw i neuronów w warstwach, funkcja aktywacji);
- wyboru metod treningu, czyli odpowiedniego procesu uczenia sieci;
- metod weryfikacji otrzymanych wyników.

Sztuczne sieci neuronowe wykorzystują do konstrukcji operatora predykcji szeregi czasowe. Przy czym w prognozowaniu szeregów czasowych dominują sieci neuronowe jednokierunkowe. Sam zaś proces prognozowania przy użyciu sieci neuronowych przebiega według schematu przedstawionego na rysunku 1.



Rys. 1. Schemat blokowy postępowania w prognozowaniu przy użyciu sieci neuronowych

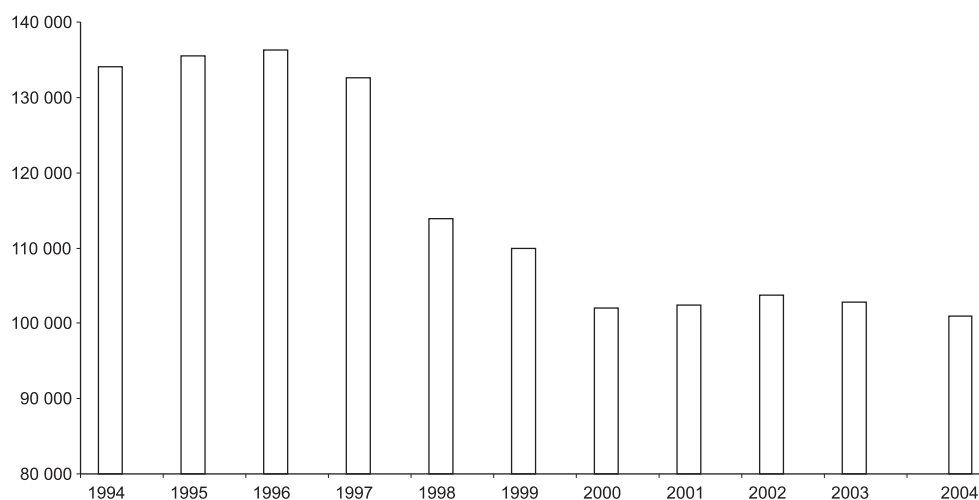
Źródło: opracowanie własne

Jeżeli chodzi o dziedzinę prognozowania ekonomicznego, to według literatury przedmiotu sztuczne sieci neuronowe mogą być wykorzystane m.in. do prognozowania [4, 5, 16]:

- zachowania rynku akcji,
- bankructw,
- zmian trendu rynkowego,
- kursów walutowych,
- ryzyka kredytowego,
- wyników finansowych,
- wielkości i wartości sprzedaży.

### 3. Analiza predykcji z wykorzystaniem sztucznych sieci neuronowych

Do oceny efektywności predykcji zjawisk za pomocą sztucznych sieci neuronowych wybrano szereg czasowy obejmujący 11 danych, opisujących wydobycie węgla kamiennego w Polsce w latach 1994–2004 (rys. 2). Jest to niezbyt liczny szereg czasowy, stanowiący próbę testową do weryfikacji skuteczności predykcji przy użyciu sztucznych sieci neuronowych.



**Rys. 2.** Wydobycie węgla kamiennego w Polsce w latach 1994–2004 [mln ton]

Źródło: opracowanie własne na podstawie danych z [17–19]

Dane szeregu czasowego, obejmującego okres 1994–2003 stanowiły podzbiór inicjalny, a wartość wydobycia z 2004 była daną kontrolną, będącą podstawą sprawdzenia dokładności predykcji wykorzystującej ww. metodę.

Obliczenia przy użyciu sieci neuronowych przebiegają według następujących czterech etapów:

1. Przygotowanie danych do przetwarzania przez sieć neuronową, w pełni charakteryzujące rozwiązywane zagadnienie.
2. Dobór odpowiedniej architektury sieci.
3. Wybór odpowiedniego procesu uczenia.
4. Weryfikacja otrzymanych wyników.

Należy podkreślić, że każdy z nich jest równie istotnym elementem składowym w całym procesie uczenia sieci neuronowych, a tym samym w procesie predykcji przy wykorzystaniu tej metody prognozowania.

Powyższe etapy są powtarzane tak długo, aż miara jakości końcowej sieci neuronowej nie przyjmie satysfakcjonującej wartości.

### **Przygotowanie danych do przetwarzania przez sieć neuronową**

Określenie właściwych zmiennych wejściowych dla zaproponowanej architektury sieci jest elementem dość skomplikowanym, ściśle zależy od wybranego programu do symulacji nauki sieci.

Ze względu na wybór programu QNET 2000 struktura danych jest następująca:

- dane uporządkowane są w wierszach;
- pierwsze dziesięć to dane wejściowe (IN), ostatnia jedenasta jest daną wyjściową (OUT);
- liczba wierszy zależy od liczby wektorów wejściowych dobranych według miary Vapnika–Chervonenkisa.

Do generowania danych stworzono program GENERATOR, który właśnie w taki sposób tworzy pliki wejściowe dla programu QNET 2000. Program QNET 2000 nie wymaga podawania plików podzielonych na dane uczące i testujące, gdyż jest to możliwe bezpośrednio w programie. W symulacjach przyjęto podział 70/30, tzn. 70% wszystkich danych z próby to dane uczące, a 30% to dane testujące dla sieci.

### **Architektura sieci**

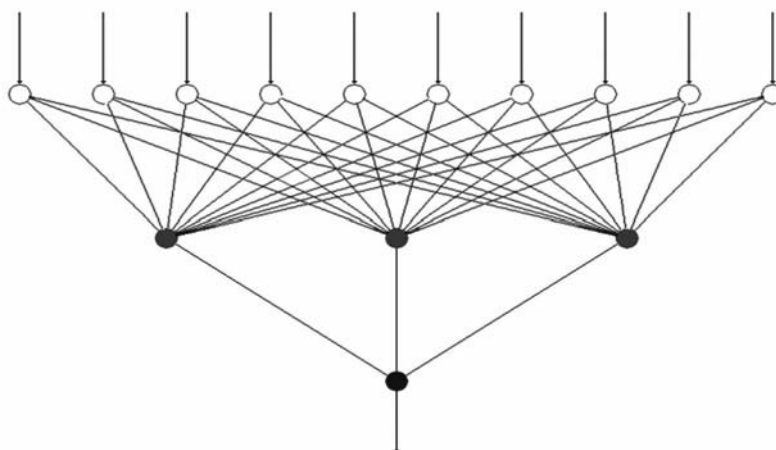
Kolejnym z elementów w konstrukcji sieci neuronowej jest odpowiednio dobrana liczba neuronów w warstwie ukrytej. Liczba ta pociąga za sobą odpowiednią liczbę wzorców podawanych sieci (liczba wektorów wejściowych). W celu dobrania odpowiedniej konfiguracji wykorzystano miarę Vapnika–Chervonenkisa ( $VCdim$ ) postaci

$$2 \left\lceil \frac{K}{2} \right\rceil N \leq VCdim \leq 2N_w (1 + \log N_n) \quad (1)$$

gdzie:

- [ ] — część całkowita liczby,
- $N$  — wymiar wektora wejściowego,
- $K$  — liczba neuronów w warstwie ukrytej,
- $N_w$  — całkowita liczba wag w sieci,
- $N_n$  — całkowita liczba neuronów w sieci.

Miara ta pozwala jedynie na oszacowanie wstępnej liczby wektorów, dlatego w praktyce przeprowadzono dużą liczbę testów, po których wybrano taką konfigurację sieci, dla której otrzymano najlepsze wyniki, tzn. wybrano taką architekturę sieci, dla której liczba neuronów w warstwie ukrytej jest najmniejsza, przy akceptowalnym błędzie uczenia i zdolności generalizacji. Optymalny dobór omawianych parametrów ma olbrzymi wpływ na szybkość generalizacji sieci. Minimalizacja błędu uczenia na zbyt małej, w stosunku do liczby wag sieci, liczbie danych uczących powodowała dowolność wartości wielu wag, co przy zmianie punktów testujących względem uczących jest przyczyną znacznego odchylenia aktualnej wartości od spodziewanej. Dalsze zmniejszanie liczby neuronów w warstwie ukrytej, przy niezmienionej liczbie próbek uczących, pozwoliło uzyskać zarówno mały błąd uczenia, jak również dobrą zdolność generalizacyjną. Należy podkreślić, że dalsze zmniejszanie liczby neuronów ukrytych powoduje niezdolność sieci do odwzorowania danych uczących. Sieć nie jest zdolna dokonać poprawnego odwzorowania danych uczących, gdyż liczba stopni swobody takiej sieci jest za mała w stosunku do wymagań odwzorowania.



Rys. 3. Architektura sieci wykorzystywanej do symulacji

Źródło: opracowanie własne

Wybrana konfiguracja sieci neuronowej posiadała następujące parametry (rys. 3):

- warstwa wejściowa: 10 neuronów,
- warstwa ukryta: 3 neurony,
- warstwa wyjściowa: 1 neuron,
- liczba wektorów wejściowych: 50.

### Proces uczenia

Po znalezieniu odpowiedniej architektury sieci można przystąpić do właściwego procesu uczenia, który ma prowadzić do otrzymania prognozy wydobycia węgla kamiennego w Polsce w 2004 roku. W przeprowadzanych symulacjach zdecydowano się oprzeć na algorytmie wstecznej propagacji błędów.

Algorytm ten można przedstawić w następujących powtarzalnych cyklach:

1. Na początku przeprowadzana jest analiza sieci neuronowej o zwykłym kierunku przepływu sygnałów, tj. od warstwy wejściowej poprzez ukrytą do wyjściowej. Po przeprowadzonej symulacji określone zostają sygnały wejściowe pomiędzy warstwami ukrytymi oraz sygnał wyjściowy.
2. Generacja sieci propagacji wstecznej. Dla tak skonstruowanej sieci są obliczane wartości odpowiednich różnic wstecznych, tzn. sieć propagacji wstecznej definiowana jest poprzez zainicjowanie procesu zmiany kierunku przepływu sygnału, od warstwy wyjściowej do warstwy wejściowej oraz zmianę funkcji aktywacji na jej pochodne.
3. Proces uczenia przeprowadzany jest na zbiorze wyników otrzymanych w etapach poprzednich, ale w sieci zwykłej, czyli o normalnym przepływie, oraz w sieci wstecznej propagacji o przepływie odwróconym.
4. Powyższe etapy powtarzane są dla całego zbioru danych aż do chwili spełnienia kryterium algorytmu, tzn. gdy spełniony jest warunek – norma gradientu będzie mniejsza od zadanej wielkości dokładności procesu uczenia.

W symulacjach za funkcje aktywacji przyjęto funkcje sigmoidalne dla wszystkich neuronów sieci.

#### Weryfikacja otrzymanych wyników

Po nauczaniu sieci przystąpiono do właściwego procesu predykcji, w wyniku którego otrzymano następującą prognozę wydobycia węgla kamiennego w Polsce dla 2004 roku

$$y_{2004}^p = 1008\,937,39 \text{ tys. ton.}$$

Wartość rzeczywistego wydobycia w tym roku wyniosła 100 995 tys. ton [19].

Do oszacowania celowości wykorzystania sieci neuronowej do prognozowania krótkiego szeregu czasowego (opisującego wydobycie węgla kamiennego w Polsce) postanowiono:

- porównać błędy predykcji prognozowania z wykorzystaniem sztucznych sieci neuronowych z błędami tego samego rodzaju, ale uzyskanymi w wyniku prognozowania szeregu czasowego za pomocą następujących metod:
  - logarytmicznego modelu tendencji rozwojowej,
  - średnich ruchomych prostych,
  - średnich ruchomych ważonych, wykorzystujących wagi liniowe,
  - średnich ruchomych ważonych, wykorzystujących wagi harmoniczne,
  - wyrównania wykładniczego,
  - wag harmonicznych;
- przedstawić ranking ww. metod uwzględniający zarówno dokładność predykcji, jak również szybkość obliczeń, skomplikowanie wybranej metody, zakres koniecznej wiedzy badacza oraz konieczność wykorzystania określonego warsztatu badawczego (maszyny liczące, programy komputerowe itp.); za każde wyróżnione kryterium metoda mogła otrzymać od 1 do 5 punktów; w sumie do zdobycia było 25 punktów (5×5 punktów).

Realizując pierwszy etap oceny celowości wykorzystania sieci neuronowej do prognozowania, obliczono następujące błędy predykcji:

— średni absolutny błąd MAE

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t| \quad (2)$$

— średni absolutny błąd procentowy MAPE

$$\text{MAPE} = \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% \quad (3)$$

— procentowy błąd średniokwadratowy MSPE

$$\text{MSPE} = \frac{1}{T} \sum_{t=1}^T \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2 \cdot 100\% \quad (4)$$

gdzie:

- $y_t$  — dane rzeczywiste (empiryczne) szeregu czasowego,
- $\hat{y}_t$  — dane teoretyczne szeregu czasowego,
- $T$  — długość szeregu czasowego.

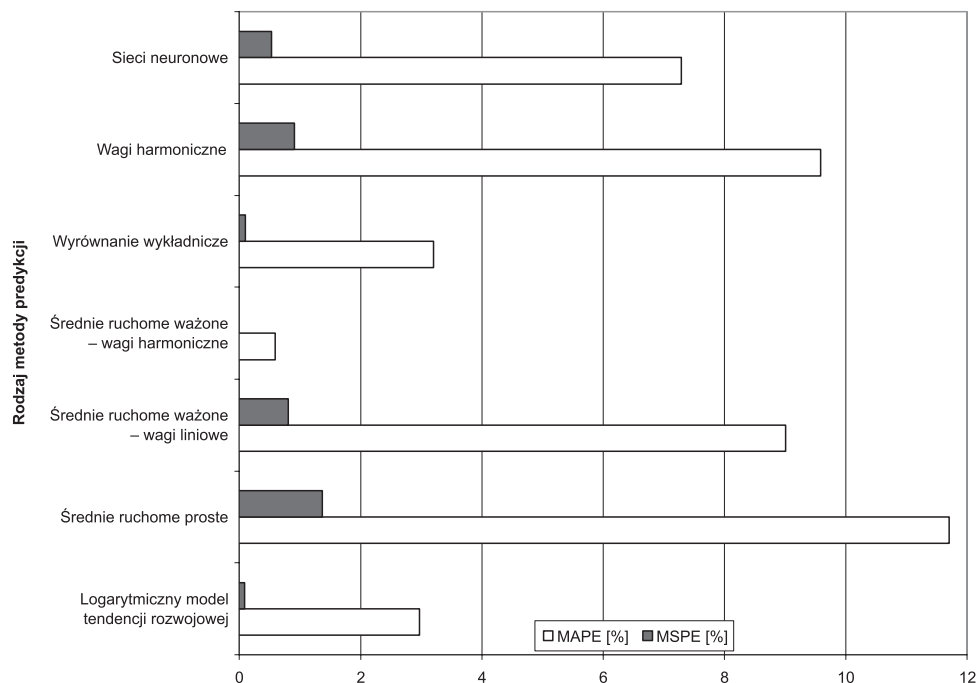
Wyniki obliczeń zestawiono w tabeli 1, a kształtowanie się błędów procentowych MAPE i MSPE przedstawiono na rysunku 4.

TABELA 1

**Wartości błędów predykcji dla wybranych metod prognozowania wydobycia węgla kamiennego**

| Lp. | Rodzaj metody                             | Rodzaj błędu   |          |          |
|-----|---|----------------|----------|----------|
|     |   | MAE [tys. ton] | MAPE [%] | MSPE [%] |
| 1   | Logarytmiczny model tendencji rozwojowej  | 2910,61        | 2,97     | 0,09     |
| 2   | Średnie ruchome proste                    | 13 379,90      | 11,70    | 1,37     |
| 3   | Średnie ruchome ważone – wagi liniowe     | 10 001,70      | 9,01     | 0,81     |
| 4   | Średnie ruchome ważone – wagi harmoniczne | 600,60         | 0,59     | 0,00     |
| 5   | Wyrównanie wykładnicze                    | 3134,60        | 3,20     | 0,10     |
| 6   | Wagi harmoniczne                          | 8828,64        | 9,58     | 0,91     |
| 7   | Sieci neuronowe                           | 7942,36        | 7,29     | 0,53     |

Źródło: opracowanie własne



**Rys. 4.** Kształtowanie się średniego absolutnego błędu procentowego (MAPE) oraz procentowego błędu średniokwadratowego (MSPE) dla analizowanych metod predykcji

Źródło: opracowanie własne na podstawie danych z tabeli 1

Analizując dane zamieszczone w tabeli oraz na rysunku 4, można stwierdzić, że:

- najlepszą, spośród analizowanych, metod predykcji szeregu czasowego opisującego wydobycie węgla kamiennego jest metoda wykorzystująca średnie ruchome ważone z wagami harmonicznymi;
- sztuczne sieci neuronowe nie dają co prawda najlepszych wyników, ale też są metodą charakteryzującą się w miarę poprawnymi wynikami (czwarte miejsce wśród siedmiu metod).

Realizując drugi etap oceny celowości wykorzystania sieci neuronowej do prognozowania, opracowano tabelę 2 zawierającą punktową ocenę ww. metod predykcji.

Wartościowania wyżej wymienionych metod predykcji dokonali autorzy publikacji, wykorzystując własne doświadczenia, nabyte podczas prowadzenia prac analitycznych nad wykorzystaniem klasycznych i nieklasycznych metod predykcji do prognozowania zjawisk gospodarczych. Metody z najwyższą liczbą punktów były najbardziej przyjazne dla użytkownika. Analizując dane zamieszczone w tabeli, można stwierdzić, że najbardziej przyjaznymi predyktorami są kolejno: średnie ruchome ważone wagami harmonicznymi, logarytmiczny model tendencji rozwojowej oraz średnie ruchome ważone wagami liniowymi. Sieci neuronowe zajmują w tym rankingu ostatnie miejsce, a to z uwagi na znaczne skomplikowanie predyktora i wymagane w związku z tym posiadanie przez badacza znacznej wiedzy i rozbudowanego warsztatu badawczego.



TABELA 2

**Punktowa ocena wybranych metod predykcji**

| Lp. | Rodzaj metody                             | Ocena punktowa | Ocena procentowa<br>25 pkt = 100% |
|-----|---|----------------|-----------------------------------|
| 1   | Logarytmiczny model tendencji rozwojowej  | 21             | 84%                               |
| 2   | Średnie ruchome proste                    | 19             | 76%                               |
| 3   | Średnie ruchome ważone – wagi liniowe     | 20             | 80%                               |
| 4   | Średnie ruchome ważone – wagi harmoniczne | 23             | 92%                               |
| 5   | Wyrównanie wykładnicze                    | 18             | 72%                               |
| 6   | Wagi harmoniczne                          | 16             | 64%                               |
| 7   | Sieci neuronowe                           | 16             | 64%                               |

Źródło: opracowanie własne

#### 4. Zakończenie

W artykule zaprezentowano rozważania dotyczące zastosowania wybranych modeli i metod prognostycznych do predykcji.

Na podstawie analiz przeprowadzonych w artykule można sformułować następujące uwagi i wnioski:

1. Sieci neuronowe to narzędzie o bardzo dużych możliwościach, jednak mało popularne wśród badaczy zajmujących się prognozowaniem zjawisk ekonomicznych i gospodarczych; przyczyną tego stanu rzeczy może być albo znaczne skomplikowanie tego predyktora, albo brak możliwości uzyskania szeregu czasowego o znacznej liczebności.
2. Znane w literaturze przykłady wykorzystania sieci neuronowych do prognozowania zjawisk gospodarczych dotyczą tych dziedzin, które można scharakteryzować licznym szeregiem czasowym.
3. Sztuczne sieci neuronowe, w przypadku analizy krótkiego szeregu czasowego, nie dają co prawda najlepszych wyników, ale też są predyktorem charakteryzującym się w miarę poprawnymi wynikami prognozowania.
4. Trzeba zgodzić się z tezą zgłoszoną przez J. Mikusia, że nie należy stosować skomplikowanego narzędzia badawczego (sieci neuronowych) do predykcji tych zjawisk ekonomicznych, dla których istnieje możliwość zastosowania w miarę prostych metod analitycznych.

#### LITERATURA

- [1] *Dittmann P.*: Prognozowanie w przedsiębiorstwie. Kraków, Oficyna Ekonomiczna 2003
- [2] *Greń J.*: Statystyka matematyczna. Modele i zadania. Wyd. VII. Warszawa, PWN 1982
- [3] *Gworek S.*: Ocena skuteczności predykcji wybranych modeli prognostycznych. Kraków, WGiG AGH KEZP 2005 (praca magisterska – promotor dr inż. A. Utrata)
- [4] *Kiełtyka L. i in.*: Inteligentny system prognozowania. Zasady funkcjonowania. Zastosowania. Częstochowa, Wydawnictwo Politechniki Częstochowskiej 2000

- [5] *Mikuś J.*: Prognozowanie w badaniach marketingowych. Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej 2003
- [6] *Pawłowski Z.*: Ekonometryczna analiza procesu produkcyjnego. Wyd. 2. Warszawa, PWN 1976
- [7] Praca zbiorowa pod redakcją M. Woźniaka: Statystyka ogólna. Wyd. II poprawione. Kraków, AE 1997
- [8] Praca zbiorowa pod redakcją M. Cieślak: Prognozowanie gospodarcze. Metody i zastosowania. Warszawa, PWN 2002
- [9] Praca zbiorowa pod redakcją M. Krzysztofiaka: Ekonometria. Warszawa, PWE 1984
- [10] Praca zbiorowa pod redakcją T. Szapiro: Decyzje menedżerskie z Excelem. Warszawa, PWE 2000
- [11] *Starzyńska W.*: Statystyka praktyczna. Warszawa, PWN 2000
- [12] *Tadeusiewicz R.*: Sieci neuronowe. Warszawa, Akademicka Oficyna Wydawnicza 1993
- [13] *Utrata A.*: Adaptacyjne metody predykcji mierników oceny pracy przedsiębiorstwa. Materiały konferencyjne Szkoły Ekonomiki i Zarządzania w Górnictwie 2004, Bukowina Tatrzańska 15–17 września 2004, 455–461
- [14] *Utrata A.*: Wpływ sposobu prezentacji zmiennej egzogenicznej na stopień dopasowania klasycznych modeli tendencji rozwojowej. Materiały konferencyjne Szkoły Ekonomiki i Zarządzania w Górnictwie 2003, Bukowina Tatrzańska 10–12 września 2003, 351–358
- [15] *Volk W.*: Statystyka stosowana dla inżynierów. Wyd. 1. Warszawa, WNT 1973
- [16] [www.aitech.com.pl](http://www.aitech.com.pl), 2005
- [17] [www.mgip.gov.pl](http://www.mgip.gov.pl), 2005
- [18] [www.pgi.gov.pl](http://www.pgi.gov.pl), 2005
- [19] [www.stat.gov.pl](http://www.stat.gov.pl), 2005
- [20] *Zeliaś A., Pawelek B., Wanat S.*: Prognozowanie ekonomiczne. Teoria. Przykłady. Zadania. Warszawa, PWN 2003
- [21] *Zeliaś A.*: Teoria prognozy. Warszawa, PWE 1997