

Monika Hadaś-Dyduch

Uniwersytet Ekonomiczny w Katowicach

e-mail: monika.dyduch@ue.katowice.pl

WPLYW ROZSZERZENIA PRÓBKII PRZY GENEROWANIU WSPÓLCZYNNIKÓW FALKOWYCH SZEREGU NA TRAFNOŚĆ PROGNOZY

Streszczenie: Celem badania jest ocena wpływu zaproponowanej metody rozszerzenia próbki na trafność prognozy szeregów, w tym przypadku indeksu WIG. Prognozę szeregu prezentującego WIG wykonano na podstawie modelu opartego na transformacie falkowej. Przed przystąpieniem do aplikacji rozszerzenia szeregu wejściowy szereg danych podzielono na próbki o parzystej liczbie obserwacji celem wyznaczenia dokładniejszych prognoz. W artykule skoncentrowano się tylko na dodatkowym rozszerzeniu próbki przy wyznaczeniu współczynników, omijając proces predykcji. Zatem nie opisano w artykule szczegółowo modelu zastosowanego do predykcji, ponieważ nie jest celem artykułu ocena zdolności predykcyjnych modelu, a jedynie wpływ na końcowy wynik prognozy metody rozszerzenia szeregu przy wyznaczeniu współczynników a_k .

Słowa kluczowe: transformata falkowa, analiza wielorozdzielcza, predykcja WIG, predykcja szeregów, współczynniki falkowe.

DOI: 10.15611/ekt.2014.4.06

1. Wstęp

Celem badania jest ocena wpływu zaproponowanej metody rozszerzenia próbki na trafność prognozy szeregów, w tym przypadku indeksu WIG. Badanie w tym kontekście jest innowacyjne, chociaż samo prognozowanie indeksów giełdowych różnorodnymi metodami jest i było w literaturze szeroko omawiane [Lach 2012; Hadaś-Dyduch 2013a; 2014; Guresen, Kayakutlu, Daim 2011].

Prognozę szeregu prezentującego WIG wykonano na podstawie modelu opartego na transformacie falkowej, ponieważ falki – z jednej strony – umożliwiają dokładniejszą analizę przez specyfikację osobnych zależności według pasm częstości, a następnie konstruowanie prognoz oryginalnych szeregów w postaci agregatów prognoz wyznaczonych dla poszczególnych komponentów procesów (skal czasu), z drugiej zaś – upraszczają analizę przez przekształcenie szeregu do postaci, dla której może być łatwiej dobrać odpowiedni predyktor [Bruzda 2013]. Oczywiście, jak

każda metoda, predykcja z użyciem falek ma swoje wady i zalety. Wadą takiego podejścia jest m.in. znaczna liczba parametrów podlegających estymacji i arbitralność wyboru falki czy poziomu dekompozycji.

Autorski model zastosowany do badania szeroko opisany w pozycji [Hadaś-Dyduch 2014] składa się z kilku zasadniczych etapów. W pierwszej kolejności wyznaczamy funkcję skalującą i jej falę, następnie odpowiednio dobrane statystycznie i merytorycznie szeregi czasowe dzielimy na podszeregi z podziałem na dwa dodatkowe podzbiory. Odpowiednio przygotowane dwie grupy podszeregów wprowadzamy do etapu sztucznej sieci neuronowej. Efektem działania sieci są współczynniki falkowe szukanych prognoz w postaci współczynników falkowych. W ostatnim etapie poprzez aplikacje algorytmu odwrotnej transformaty falkowej otrzymujemy wartości szukanych prognoz. Dla każdego etapu niezbędne jest wyznaczenie odpowiednich błędów i zastosowanie odpowiednich parametrów minimalizujących te błędy.

2. Założenia badania

Badanie wykonano na szeregu prezentującym WIG. Liczba obserwacji uwzględniona w badaniu wynosiła 4117. Szereg danych podzielono na podszeregi o parzystej liczbie obserwacji, otrzymując 2058 szeregów dwuelementowych bądź 1029 szeregów czteroelementowych itd. Przyjmując podział każdego szeregu na podszeregi dwuelementowe, otrzymano:

Podszereg 1: 1 obserwacja, 2 obserwacja
 Podszereg 2: 3 obserwacja, 4 obserwacja

 Podszereg 2058: 4115 obserwacja, 4116 obserwacja.

Prognozę wykonano na okres 6, 12 i 24 miesięcy.

Prognozę szeregu prezentującego WIG wykonano na podstawie modelu opartego na transformacie falkowej, szczegółowo opisanego w pracach [Hadaś-Dyduch 2013a; 2014]. Celem badania jest analiza wpływu zaproponowanej metody rozszerzenia próbki na trafność prognozy WIG.

Do predykcji szeregu WIG zastosowano m.in. analizę falkową opartą na falce Daubechies, wyrażonej wzorem:

$$\psi(r) = -\frac{1+\sqrt{3}}{4}\phi(2r-1) + \frac{3+\sqrt{3}}{4}\phi(2r) - \frac{3-\sqrt{3}}{4}\phi(2r+1) + \frac{1-\sqrt{3}}{4}\phi(2r+2) \quad \psi(r) = 0$$

dla $r < -1$ lub $r > 2$.

3. Rozszerzenie próbki

Mając dany szereg danych, nazwany w dalszej części również próbką:

$$\vec{s} = (s_0, s_1, s_2, \dots, s_{2^n-2}, s_{2^n-1}),$$

obliczamy, uwzględniając dodatkowe rozszerzenie próbki, współczynniki falkowe.

W tym celu postępujemy zgodnie z algorytmem:

1. Przyjmujemy, że:

$$\vec{s} = (s_0, s_1, s_2, \dots, s_{2^n-2}, s_{2^n-1}) = p_0, p_1, p_2, \dots, p_{2^n-2}, p_{2^n-1}.$$

2. Zapisujemy formalną postać rozszerzenia próbki \vec{s} :

$$\vec{s} = (s_{-(2^n-1)}, \dots, s_{-2}, s_{-1}; s_0, s_1, \dots, s_{2^n-2}, s_{2^n-1}; s_{2^n}, \dots, s_{2^{n+1}-1}).$$

3. Przypisujemy elementom:

$$s_{-(2^n-1)}, \dots, s_{-2}, s_{-1}$$

oraz elementom

$$s_{2^n}, \dots, s_{2^{n+1}-1}$$

odpowiednie wartości elementów wejściowych próbki:

$$\vec{s} = (s_0, s_1, s_2, \dots, s_{2^n-2}, s_{2^n-1}) = p_0, p_1, p_2, \dots, p_{2^n-2}, p_{2^n-1}.$$

4. Zapisujemy ostateczną postać rozszerzenia próbki \vec{s} :

$$(s_{-(2^n-1)}, \dots, s_{-2}, s_{-1}; s_0, s_1, s_2, s_3, \dots, s_{2^n-1}, s_{2^n-1}; s_{2^n}, \dots, s_{2^{n+1}-1}),$$

$$(p_{2^n-1}, \dots, p_1, p_0; p_0, p_1, p_2, \dots, p_{2^n-2}, p_{2^n-1}; p_{2^n-1}, p_{2^n-2}, \dots, p_0).$$

5. Wyznaczamy współczynniki falkowe a_k z zależności:

$$a_k = \sum_{r=k+0}^{k+3} \varphi(r-k) \cdot s_r, \quad (1)$$

gdzie:

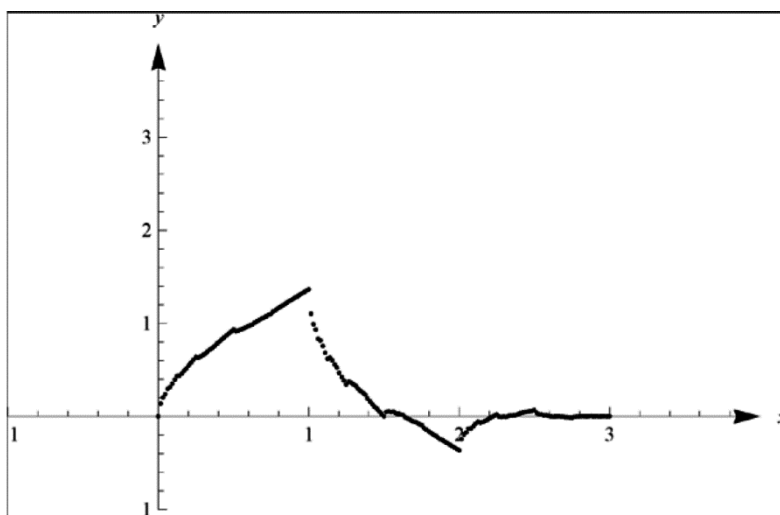
$\varphi: D \rightarrow R$ – funkcja skalująca falki Daubechies spełniająca następujące warunki:

- $\varphi(r) = h_0 \cdot \varphi(2r) + h_1 \cdot \varphi(2r-1) + h_2 \cdot \varphi(2r-2) + h_3 \cdot \varphi(2r-3),$

gdzie:

$$h_0 = \frac{1+\sqrt{3}}{4}, h_1 = \frac{3+\sqrt{3}}{4}, h_2 = \frac{3-\sqrt{3}}{4}, h_4 = \frac{1-\sqrt{3}}{4},$$

- $\sum_{k \in \mathbb{Z}} \varphi(k) = 1,$
- $\varphi(r) = 0$ dla $r \leq 0 \vee r \geq 3,$
- $D_j = \{k2^j : k \in \mathbb{Z}\},$
- $D = \bigcup_{j \in \mathbb{Z}} D_j = \bigcup_{j=0}^{\infty} D_j.$



Rys. 1. Funkcja $\varphi : D \rightarrow \mathbb{R}$

Źródło: opracowanie własne.

6. Aplikujemy algorytm transformaty falkowej polegający na zastąpieniu $2N = 2^{(n+1)}$ całkowitych przesunięć funkcji skalującej:

$$\tilde{f}(r) = \sum_{k=0}^{2^{(n+1)}-1} a_k \varphi(r-k)$$

przez równoważną kombinację liniowych całkowitych przesunięć funkcji skalującej $\varphi([r/2]-k)$ i falek $\psi([r/2-1]-k)$ z niższego poziomu rozdzielczości:

$$\tilde{f}(r) = \sum_{k=0}^{2^n-1} a_k^{(n-1)} \varphi([r/2]-k) + \sum_{k=0}^{2^n-1} c_k^{(n-1)} \psi([r/2-1]-k).$$

Współczynniki $a_k^{(n-1)}$ wskazują niższy poziom rozdzielczości niż początkowe współczynniki $a_k^{(n)}$.

4. Implementacja zaproponowanego dodatkowego rozszerzenia próbki do wyznaczenia współczynników falkowych

Wartości poszczególnych współczynników falkowych a_k wyznaczamy na podstawie algorytmu przedstawionego w rozdziale drugim z uwzględnieniem dodatkowego rozszerzenia próbki.

Opierając się na zależności (1), wyznaczamy dla poszczególnych wartości k współczynniki falkowe dla pierwszego podzeregu danych.

- Dla $k = -3$.

Wyznaczamy wartość funkcji skalującej w punkcie $r+3$ oraz w punkcie r dla $r = -(2^n - 1), \dots, 2^{n+1} - 1$. Wyniki obliczeń pomocniczych zestawiono w tab. 1.

Tabela 1. Obliczenia pomocnicze do wyznaczenia współczynnika a_k dla $k = -3$

r	$\phi(r-k)$	$\phi(r+3)$	s_r
$-(2^n - 1)$	$\phi(-2^n + 4)$	0	0
\vdots	\vdots	\vdots	\vdots
-2	$\phi(1)$	$\frac{1-\sqrt{3}}{2}$	0
-1	$\phi(2)$	$\frac{1+\sqrt{3}}{2}$	0
0	$\phi(3)$	0	967,7
1	$\phi(4)$	0	945,6
2	$\phi(5)$	0	0
3	$\phi(6)$	0	0
4	$\phi(7)$	0	0
\vdots	\vdots	\vdots	\vdots
$2^n - 1$	$\phi(2^n + 2)$	0	0
2^n	$\phi(2^n + 3)$	0	0
\vdots	\vdots	\vdots	\vdots
$2^{n+1} - 1$	$\phi(2^{n+1} + 2)$	0	0

Źródło: opracowanie własne.

Na podstawie obliczeń zawartych w tab. 1 oraz wzoru (1) otrzymujemy wartość współczynnika falkowego $a_{-3} = 0$.

- Dla $k = -2$. Odpowiednie obliczenia zestawiono w tab. 2.

Tabela 2. Obliczenia pomocnicze do wyznaczenia współczynnika a_k dla $k = -2$

r	$\phi(r-k)$	$\phi(r+2)$	s_r
$-(2^n - 1)$	$\phi(-2^n + 3)$	0	0
\vdots	\vdots	\vdots	\vdots
-2	$\phi(0)$	0	0
-1	$\phi(1)$	$\frac{1+\sqrt{3}}{2}$	0
0	$\phi(2)$	$\frac{1-\sqrt{3}}{2}$	967,7
1	$\phi(3)$	0	945,6
2	$\phi(4)$	0	0
3	$\phi(5)$	0	0
4	$\phi(6)$	0	0
\vdots	\vdots	\vdots	\vdots
$2^{n+1} - 1$	$\phi(2^{n+1} + 1)$	0	0

Źródło: opracowanie własne.

Na podstawie obliczeń zawartych w tab. 2 oraz wzoru (1) otrzymujemy $a_{-2} = -14366,2$.

Postępując analogicznie, otrzymujemy wartości współczynników falkowych dla pozostałych wartości k pierwszego podszeregu danych, tj.:

- Dla $k = -1$, $a_{-1} = 975,7892$.

Tabela 3. Obliczenia pomocnicze do wyznaczenia współczynnika a_k dla $k = 2$

r	$\phi(r-k)$	$\phi(r-2)$	s_r
0	$\phi(-2)$	0	967,7
1	$\phi(-1)$	0	945,6
2	$\phi(0)$	0	0
3	$\phi(1)$	$\frac{1+\sqrt{3}}{2}$	0
4	$\phi(2)$	$\frac{1-\sqrt{3}}{2}$	0

Źródło: opracowanie własne.

- Dla $k = 0$, $a_0 = 1291,713622$.
- Dla $k = 1$, $a_1 = 0$.
- Dla $k = 2$, $a_2 = 0$ (zob. tab. 3).

W sposób analogiczny do przedstawionego powyżej wyznaczamy współczynniki a_k dla pozostałych podszeregów utworzonych z szeregu głównego.

Ostatecznie w wyniku transformaty falkowej otrzymujemy macierz współczynników falkowych, która przy podziale wejściowego szeregu na szeregi dwuelementowe, jednym poziomem rozdzielczości falki i zaprezentowanej metodzie rozszerzenia próbki ma postać:

$$C = \begin{pmatrix} 5080,5622 & 5110,2607 & -25,7196 & 25,7196 \\ 5368,7082 & 5482,5524 & -98,5920 & 98,5920 \\ 5095,0579 & 5173,5468 & -67,9733 & 67,9733 \\ \vdots & \vdots & \vdots & \vdots \\ 28166,9057 & 28395,3294 & -197,8208 & 197,8208 \\ 28022,0584 & 27503,6006 & 448,9976 & -448,9976 \\ 27165,4975 & 27144,6166 & 18,0834 & -18,0834 \end{pmatrix}.$$

5. Wyniki prognozy z uwzględnieniem rozszerzonej próbki

Otrzymana macierz współczynników falkowych jest niezbędna do kolejnego etapu modelu prognozy (zob. [Hadaś-Dyduch 2013a; 2014]), tj. do inicjalizacji sztucznej sieci neuronowej oraz wyznaczenia współczynników odwrotnej transformaty falkowej¹. W ostateczności otrzymujemy następujące błędy dla prognozy:

- dwuletniej:
 - bezwzględny błąd procentowy: 0,009%,
 - odchylenie od wartości rzeczywistej: 4,
- rocznej:
 - bezwzględny błąd procentowy: 0,055%,
 - odchylenie od wartości rzeczywistej: 21,
- półrocznej:
 - bezwzględny błąd procentowy: 0,179%,
 - odchylenie od wartości rzeczywistej: -70.

¹ Nie opisano w artykule szczegółowo modelu zastosowanego do predykcji, ponieważ nie jest celem artykułu ocena zdolności predykcyjnych modelu, a jedynie wpływ na końcowy wynik prognozy metody rozszerzenia szeregu przy wyznaczaniu współczynników a_k .

Tabela 4. Zestawienie wyników prognozy z podziałem szeregów

Metoda	Bezwzględny błąd procentowy dla prognozy o horyzoncie (w miesiącach)		
	24	12	6
Bez dodatkowego rozszerzenia próbki i z podziałem szeregu na mniejsze jednostki	0,03%	0,11%	0,21%
Dodatkowe rozszerzenie szeregu i podział szeregu na mniejsze jednostki	0,01%	0,06%	0,18%

Źródło: opracowanie własne.

Tabela 5. Zestawienie wyników prognozy bez podziału szeregów

Metoda	Bezwzględny błąd procentowy dla prognozy o horyzoncie (w miesiącach)		
	24	12	6
Bez dodatkowego rozszerzenia szeregu i bez podziału szeregu na mniejsze jednostki	0,12%	0,29%	0,29%
Dodatkowe rozszerzenie szeregu i bez podziału szeregu na mniejsze jednostki	0,06%	0,17%	0,19%

Źródło: obliczenia własne.

Należy zauważyć, że prognoza WIG modelem z aplikacją zaproponowanego rozszerzenia podszeręgów szeregów czasowych jest w każdym przypadku mniejsza od prognozy, tym samym modelem bez dodatkowego rozszerzenia podszeręgów (tab. 4).

6. Zakończenie

Celem badania była ocena wpływu zaproponowanej metody rozszerzenia podszeręgów wyodrębnionych z głównego szeregu czasowego na trafność prognozy indeksu WIG. Całościową prognozę indeksu WIG z aplikacją zaproponowanej metody rozszerzenia podszeręgów wykonano na podstawie modelu opartego na transformacji falkowej. Jednakże przed przystąpieniem do aplikacji modelu oraz do rozszerzenia odpowiednich podszeręgów wejściowe szeregi danych podzielono na próbki o parzystej liczbie obserwacji celem wyznaczenia dokładniejszych prognoz.

W badaniach przedstawionych w artykule skoncentrowano się tylko na dodatkowym rozszerzeniu podszeręgów szeregów głównych przy wyznaczaniu współczynników falkowych, świadomie pomijając proces predykcji. Sam proces predykcji opisano zwięźle we wstępie do artykułu. Zatem nie opisano w artykule szczegółowo modelu zastosowanego do predykcji, ponieważ celem artykułu nie była ocena zdol-

ności predykcyjnych modelu, a jedynie wpływ na końcowy wynik prognozy WIG metody rozszerzenia szeregu przy wyznaczaniu współczynników a_k . Otrzymane prognozy poprzez zastosowanie algorytmu predykcji M. H-D, opisanego szeroko w [Hadaś-Dyduch 2014], oraz zaproponowane rozszerzenie podszeregów uwzględnionych w badaniu szeregów czasowych, jak pokazują otrzymane wyniki, obarczone są małym błędem. Na wartości błędów otrzymanych prognoz WIG mogą mieć zasadniczy wpływ dwa czynniki, tj. podział szeregu danych na mniejsze jednostki, tzw. podszeregi, oraz dodatkowe rozszerzenie próbki przy wyznaczaniu współczynników falkowych. W artykule nie opisano szczegółowo modelu zastosowanego do predykcji, skoncentrowano się tylko na przedstawieniu idei dodatkowego rozszerzenia próbki oraz ocenie otrzymanych błędów predykcji na podstawie skuteczności rozszerzenia, ponieważ badania przedstawione w artykule są kontynuacją badań przedstawionych w [Hadaś-Dyduch 2014].

Literatura

- Bruzda J., 2013, *Prognozowanie metodą wyrównywania falkowego*, Acta Universitatis Nicolai Copernici Zarządzanie.
- Guresen E., Kayakutlu G., Daim T., 2011, *Using a rtificial neural network models in stock market index prediction*, Expert Systems with Applications, vol. 38, s. 10389-10397.
- Hadaś M., 2006, *Zastosowanie sieci falkowo-neuronowej do predykcji ekonomicznych szeregów czasowych*, [w:] Prace Naukowe UE we Wrocławiu nr 1112, *Prognozowanie w zarządzaniu firmą*, praca zbiorowa pod redakcją naukową Pawła Dittmanna i Joanny Krupowicz, Wydawnictwo UE we Wrocławiu, Wrocław, s. 69-80.
- Hadaś M., 2008, *Sieć falkowo-neuronowa jako skuteczne narzędzie do analizy i predykcji szeregów czasowych*, [w:] Prace Naukowe UE, *Metody matematyczne, ekonometryczne i komputerowe w finansach i ubezpieczeniach 2006*, praca zbiorowa pod redakcją Piotra Chrzana i Tadeusza Czernika, Wydawnictwo AE, Katowice, s. 175-185.
- Hadaś M., 2008a, *Wykorzystanie metod sztucznej inteligencji do wspomagania decyzji inwestycyjnych*, [w:] *Inwestowanie na rynku kapitałowym*, Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania nr 10, Uniwersytet Szczeciński, Szczecin, s. 446-457.
- Hadaś M., 2009, *Prognozowanie szeregów czasowych w oparciu o współczynniki transformaty falkowej, optymalizowane przez sztuczną sieć neuronową*, [w:] *Metody matematyczne, ekonometryczne i komputerowe w finansach i ubezpieczeniach 2009*, praca zbiorowa pod redakcją Andrzeja Stanisława Barczaka, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice, s. 59-69.
- Hadaś-Dyduch M., 2013, *Efektywność inwestycji kapitałowych mierzona modelem opartym na analizie falkowej w niestabilnym otoczeniu gospodarczym*, [w:] *Innowacje w bankowości i finansach. Studia ekonomiczne*, Zeszyty Naukowe Wydziałowe Uniwersytetu Ekonomicznego w Katowicach numer 174, red. naukowy J. Harasim, B. Frączek, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2013, s. 307-313.
- Hadaś-Dyduch M., 2013a, *Prognozowanie szeregów czasowych w oparciu o współczynniki transformaty falkowej, optymalizowane przez sztuczną sieć neuronową*, [w:] *Metody matematyczne, ekonometryczne i komputerowe w finansach i ubezpieczeniach 2009*, praca zbiorowa pod redakcją Andrzeja Stanisława Barczaka, Wydawnictwo Uniwersytetu Ekonomicznego w Katowicach, Katowice 2011, s. 59-69.

- Hadaś-Dyduch M., 2014, *Non-classical algorithm for time series prediction of the range of economic phenomena with regard to the interaction of financial market indicators*, Chinese Business Review 13 (4), New York, s. 221-231.
- Lach A., 2012, *Predykcja indeksu WIG przy użyciu neuronowego i neuronowo-rozmytego systemu klasyfikującego*, Zeszyty Naukowe UEP, nr 242, *Matematyka i informatyka na usługach ekonomii. Metody – analizy – prognozy*.

IMPACT OF SAMPLE EXTENSION IN THE GENERATION OF WAVELET COEFFICIENTS SERIES ON THE ACCURACY OF FORECASTS

Summary: The purpose of the study is to investigate the effect of the proposed method of sample extensions to the accuracy of the forecast series, in this case the WIG index. The forecast of series presenting the WIG was made on the basis of a model based on wavelet transform. Prior to the application of the extension number, the number of input data samples was divided into an even number of observations to define a more accurate forecasts. The article focuses only on the sample at an additional extension sets encapsulation of skipping the process of prediction coefficients. Therefore the article does not describe in detail the model used to predict because the aim of this article is not to evaluate the ability of prediction model and only the effect on the final outcome of prediction of six-rule extension method when determining the a_k coefficients.

Keywords: wavelet transform, multi-resolution analysis, prediction of WIG, prediction series, wavelet coefficients.